UNIVERSITE DE BOURGOGNE FRANCHE-COMTE

ECOLE DOCTORALE SOCIETES, ESPACES, PRATIQUES, TEMPS

LABORATOIRE THEMA – UMR 6049 CNRS

THESE

Pour obtenir le grade de

Docteur de l'Université de Bourgogne Franche-Comté

Discipline : Géographie

par

CHRISTOPHE MIMEUR

Soutenue le 9 décembre 2016

LES TRACES DE LA VITESSE ENTRE RESEAU ET TERRITOIRE

APPROCHE GEOHISTORIQUE DE LA CROISSANCE DU RESEAU FERROVIAIRE FRANÇAIS

Sous la direction de

THOMAS THEVENIN

Devant un jury composé de :

Corinne BLANQUART, Directrice de recherche, *IFSTTAR*	*Examinateur*
Anne BRETAGNOLLE, Professeur, *Université Paris 1 Panthéon-Sorbonne*	*Rapporteur*
Valérie FACCHINETTI-MANNONE, Maître de conférences habilitée à diriger des recherches, *Université de Bourgogne Franche-Comté*	*Examinateur*
Jordi MARTI-HENNEBERG, Professeur, *Université de Lleida, Espagne*	*Rapporteur*
Thomas THEVENIN, Professeur, *Université de Bourgogne Franche-Comté*	*Directeur*

REMERCIEMENTS

Puisqu'on rapproche souvent l'exercice de la thèse d'un long voyage, mon penchant pour les métaphores filées m'a fait prendre un train dont le parcours a duré trois ans, et qui a mis à l'épreuve l'ensemble d'un système pour arriver aujourd'hui à bon port. Depuis mon plus jeune âge, j'entends un TGV ou un train toutes les cinq minutes depuis ma fenêtre, il faut croire qu'il n'y a pas de hasards. Alors j'en ai pris un en marche… Et dans un train, les passagers sont nombreux : tout aussi nombreux sont ceux qui m'ont entouré, de près, de loin, qui m'ont aidé, de près, de loin.

Si certains diront que c'est moi qui pilote la machine, alors cela n'aurait pas été possible sans un régulateur, véritable garant de l'ensemble. Alors je remercie *Thomas,* mon directeur de thèse, pour toute la confiance qu'il n'a cessé de m'accorder. Pour toutes les opportunités que tu m'as données d'échanger avec un vaste réseau de recherche, me laissant volontiers embarquer dans un train ou un avion. Tu as terminé de forger mon âme de géographe et de globe-trotter. Tes conseils toujours avisés, nos conversations au détour du couloir du troisième étage et ton soutien indéfectible m'ont été indispensables et ont contribué au parcours scientifique mais aussi profondément humain de la thèse, depuis notre première présentation à Lisbonne jusqu'aux derniers instants de relecture. J'associe *Bob* à ces remerciements dont la passion pour notre pays, l'espace et la vitesse ont contribué à nos réflexions, à de nouvelles rencontres, dont j'espère qu'il y en aura encore plein d'autres.

Si je me suis trouvé aux commandes du train, c'est sans doute aussi grâce à *Valérie Mannone-Fachinetti*, qui est la première à m'avoir fait monter dans le train. Je te dois mes premières incursions dans le monde fascinant de la recherche. Depuis ma première année de licence, tu as contribué à ma passion pour la géographie. Que tu acceptes aujourd'hui d'évaluer mon travail est un véritable honneur pour moi qui a pu aussi compter sur ton expertise et ta profonde humanité.

Dans une telle aventure, il doit y avoir des contrôleurs. Je remercie *Anne Bretagnolle*, *Jordi Marti-Henneberg* qui ont accepté d'évaluer mon travail. J'exprime aussi ma gratitude à *Corinne Blanquart* d'avoir accepté de faire partie de mon jury. Je veux également remercier l'association *Rails & Histoire*, dont le soutien a été important pendant mon voyage, m'a permis d'en faire de nouveaux, et d'échanger avec des passionnés du monde ferroviaire, en même temps qu'ils sont experts.

Mon voyage m'a amené à m'arrêter auprès de personnes qui ont contribué à ce travail, et dont l'aide a été précieuse. Je remercie très vivement *Martin Koning*, dont l'accompagnement m'a

conduit sur les voies de l'économétrie. Nos discussions ont toujours été riches de conseils et d'échanges qui nous pousseront sans aucun doute à continuer de travailler ensemble. Aussi, je remercie *François Queyroi* et *Arnaud Banos* pour m'avoir, eux, conduit sur les voies du graphe et de la simulation. Nul doute que nos collaborations continueront de se concrétiser prochainement. Je remercie également *Gilles*, qui m'a lancé sur les pentes du SQL, dont j'aurais désormais du mal à me passer, et à *Ludovic* et la plate-forme géomatique de la MSH de Dijon, qui ont toujours su démêler mes requêtes et héberger mes données. J'adresse mes remerciements aux équipes du BRGM dont les échanges nous ont permis d'acquérir de nouvelles données, encore et toujours.

Mes escales m'ont aussi porté vers les gares de l'enseignement. J'exprime toute ma reconnaissance à une vraie équipe pédagogique dans laquelle il a été apaisant et stimulant de travailler. Beaucoup ont été mes profs avant qu'ils ne deviennent des collègues : parmi eux, les échanges avec *Pierre, Pascal, Yves, Nadège, Albin* ont contribué à ma passion d'enseigner, mais aussi auprès de collègues qui sont montés à bord depuis moins longtemps, comme *Cécilia, Julie, Jean-Philippe*. Un remerciement tout particulier à *Anne* pour son enthousiasme tout aussi indispensable qu'inépuisable, et pour son profond soutien pendant toute la phase de rédaction, qui a été très important et qui porte aujourd'hui ses fruits. Je remercie *Dany*, dont nos discussions ont été riches, stimulantes. Parmi ceux qui ont contribué à mon apprentissage de la recherche, j'associe aussi tous les bisontins du laboratoire ThéMA, dont l'enthousiasme pour mon projet a toujours stimulé nos rencontres, je pense surtout à Yohan, Nicolas, Medhi, Maxime, Vincent et Valentine avec qui j'ai partagé de longues heures de routes italiennes. Je remercie également *Robert* dont l'expérience et la bienveillance ont compté, de même que *Patrick* et *Agnès* avec qui j'ai partagé tant de choses depuis ma première année de licence, au détour d'un bureau ou d'un café. Nos discussions ont été d'un réconfort important, encore plus dans les derniers moments.

Et puis, il y a les compagnons de voyage, avec qui j'ai partagé le même compartiment, pour tout ou partie du voyage, qui ont fait que mon voyage a été effectué en première classe. Parmi eux, *Guillaume* a été un prof, un encadrant, un collègue mais surtout un véritable ami, sur lequel j'ai toujours pu compter pour éviter les déraillements. Et puis *Justin* bien sûr, un camarade de promo, mais surtout lui aussi un véritable ami, avec qui j'ai partagé le voyage, le bureau mais aussi plein d'autres moments indispensables. J'associe aussi *Rebecca*, qui est arrivé à destination il y a deux ans, ainsi que ceux qui viennent de prendre les commandes de leur train : *Florian*, avec qui les échanges à l'heure du déjeuner ont toujours été riches et remplis de détente ; *Larissa*, dont la gentillesse et les permanentes attentions n'ont d'égal que le soleil qu'elle a importé du Brésil ; *Emilie*, avec qui j'ai démêlé plus d'une requête ; *Thibaut*, qui vient de monter à bord mais aussi *Maud*. Je pense aussi à ceux qui ont contribué à la bonne humeur du bureau : *Séverine, Natalie*, avec qui les contacts n'ont pas été rompus et démontrent une véritable amitié. Nous nous sommes tous retrouvés à un moment au wagon-bar pour se ressourcer et faire que le voyage soit moins long. Je remercie aussi les étudiants de master avec qui j'ai partagé des projets pédagogiques mais aussi des moments de relâche salvateurs.

Enfin, il y a la famille et les amis qui ont pris le train avec moi. Ceux qui font que quand j'avais l'impression d'avancer comme un « tacot », ont contribué à me relancer à « grande vitesse ».

Ils n'ont jamais quitté le quai pour m'attendre et être à mes côtés. Parmi eux, il y a les anciens potes du TMEC, *Nico* et *Biber*, avec qui j'ai partagé tant de choses à Dijon, mais aussi à Paris, Luzy, Marseille (l'un des trois fait moins rêver, je vous laisse deviner !). Il y a les amis de longue date, *la Voland, Ju*, sur lesquels je peux toujours compter.

Et puis, il y a des personnes indispensables, qui ont assuré et assurent la bonne marche du pilote. *Lucile* y occupe une place toute particulière, une amie de la première heure, depuis 27 ans, mais dont la présence a été indispensable : tu as su me faire parfois arrêter mon train pour faire le plein, nos soirées et journées passées ensemble m'ont été indispensables, du cinéma jusqu'à la relecture de mon manuscrit, et surtout jusqu'aux derniers instants de ce long voyage. A ce propos, je remercie vivement mes *parents*, qui ont lu avec attention mon travail, qui ont contribué à la collecte de l'information géohistorique (!!!) et dont le soutien indéfectible me poursuit depuis de longues années : c'est un confort hors classe qui n'a pas de prix, de même que celui de mes *grands-parents*, dont le soutien permanent m'a apporté et m'apporte toujours une grande force. Une autre force, c'est celle inépuisable d'*Océane*, ma petite nièce et filleule qui a enchanté mon voyage et celle de ses parents, *Thierry* et *Hélène*, qui m'ont épaulé pendant trois ans, ont nourri mes week-ends, mes moments de doute et aussi souvent mon estomac, merci à vous ! Je remercie finalement toute ma *famille*, tous mes *amis*, dont les attentions ont été d'une grande aide : il y a surtout mes cousines, mais aussi mes cousins, mes amis et voisins de la campagne. Quelque part et à chaque moment, vous avez tous occupés mes pensées pendant l'écriture du manuscrit – ce ne sont pas des pensées qui distraient, ce sont des pensées indispensables pour avancer.

Alors, bonne route sur les traces de la vitesse !

SOMMAIRE

INTRODUCTION GENERALE 10

Réseau et territoire : une dualité d'actualité différenciée par la vitesse 11
Trace et tracé dans une approche par la donnée géohistorique 12
La base de données FRANcE au cœur des humanités numériques 14
Le cheminement de la thèse : une modélisation croissante 15

PREMIERE PARTIE : EFFET STRUCTURANT ET ECHELLES SPATIO-TEMPORELLES 18

INTRODUCTION DE LA PREMIERE PARTIE 19
CHAPITRE 1. L'EFFET STRUCTURANT DES INFRASTRUCTURES DE TRANSPORT : DE L'ENGOUEMENT A LA CONTROVERSE 22
Introduction 22
1.1. Les vertus de la mise en réseau 23
1.2. Qualifier les effets des infrastructures de transport 36
1.3. De la remise en cause à la controverse 51
Conclusion 62
CHAPITRE 2. LA CONSTRUCTION DE NOUVELLES ECHELLES DANS UNE PERSPECTIVE GEOHISTORIQUE 64
Introduction 64
2.1. Une géohistoire pour des réflexions à de larges échelles 65
2.2. Le « spatial turn », entre questionnement et instrument 76
2.3. Vers la définition de cadres d'exploration du lien réseau/territoire 87
Conclusion 100
CHAPITRE 3. LES TRACES DU LIEN RESEAU / TERRITOIRE PAR LA BASE DE DONNEES FRANCE 102
Introduction 102
3.1. L'information géohistorique face à l'hétérogénéité des sources 103
3.2. De l'exploitation à la complexification de l'information 119
Conclusion 131
CONCLUSION DE LA PREMIERE PARTIE 132

DEUXIEME PARTIE : L'ACCESSIBILITE DANS UN CONTEXTE GEOHISTORIQUE : UNE MODELISATION CROISSANTE DE LA DONNEE, DE L'EFFET DU RESEAU A SA DYNAMIQUE DE CROISSANCE 136

INTRODUCTION DE LA DEUXIEME PARTIE 137
CHAPITRE 4. CAPITALISER L'INFORMATION GEOHISTORIQUE : ENTRE PRECISION ET GENERALISATION 140
Introduction 140
4.1. Introduire la base de données FRANcE dans la Geocomputation 141
4.2. La théorie des graphes pour le réseau ferroviaire 151
4.3. L'accessibilité des 36 000 communes dans FRANcE : entre précision et généralisation 164
Conclusion 176
CHAPITRE 5. MESURER DES RELATIONS DE LONG TERME ENTRE INFRASTRUCTURE ET TERRITOIRE 178
Introduction 178
5.1. Caractériser la dotation en réseau par l'accessibilité 179
5.2. Un rapprochement vers l'économie historique 191
5.3. Mobiliser des informations géohistoriques de long terme pour la base de données FRANcE 201

Conclusion ... *217*
CHAPITRE 6. MODELISER LES INTERACTIONS DANS UN CONTEXTE GEOHISTORIQUE ... 220
Introduction ... *220*
6.1. La modélisation des interactions entre transport et territoire ... *221*
6.2. La construction d'un modèle de croissance d'un réseau de transport pour FRANcE ... *230*
6.3. Explorer le modèle d'évolution de réseau de transport ... *240*
Conclusion ... *256*
CONCLUSION DE LA DEUXIEME PARTIE ... 257

PARTIE 3 : GEOHISTOIRE DE L'INTERACTION ENTRE RESEAU ET POPULATION : DE LA MESURE DE L'EFFET DU RESEAU A LA SIMULATION DE SA CROISSANCE ... 262

INTRODUCTION DE LA TROISIEME PARTIE ... 263
CHAPITRE 7. DE L'ACCELERATION DU RESEAU A LA DEFINITION DES EFFETS SUR LES DYNAMIQUES DEMOGRAPHIQUES ... 266
Introduction ... *266*
7.1. Expansion et accélération : une réelle homogénéisation ? ... *267*
7.2. Les effets structurants du réseau sur la structure démographique ... *278*
7.3. Portée spatiale et temporelle de l'effet du réseau ... *293*
Conclusion ... *309*
CHAPITRE 8. DES CAPACITES CIRCULATOIRES A LA SIMULATION DE LA CROISSANCE DU RESEAU FERROVIAIRE ... 312
Introduction ... *312*
8.1. L'émergence géohistorique de corridors par les capacités circulatoires du réseau ... *313*
8.2. De l'efficacité radiale à la suggestion de transversales ... *332*
8.3. Enrichir les comparaisons par des explorations locales ... *346*
Conclusion ... *358*
CHAPITRE 9. LA BASE DE DONNEES FRANCE, UN PROJET ENTRE ELARGISSEMENT ET APPROFONDISSEMENT. 360
Introduction ... *360*
9.1. L'élargissement : une théorie des réseaux « à la française » ? ... *361*
9.2. L'approfondissement : méthodes, outils et publics ... *375*
Conclusion ... *390*
CONCLUSION DE LA TROISIEME PARTIE ... 391

CONCLUSION GENERALE ... 394

Une trajectoire de l'accumulation à la capitalisation ... *395*
Le géographe, et l'informaticien, l'économiste, l'historien… ... *396*
Le géographe, un vecteur pour penser le réseau par l'interaction spatiale ... *398*

BIBLIOGRAPHIE ... 402

TABLE DES MATIERES ... 428

TABLE DES FIGURES ... 432

TABLE DES TABLEAUX ... 436

ANNEXES ... 438

INTRODUCTION GENERALE

« Du tortillard au TGV, en passant par le corail et l'omnibus,
les trains changent, la vache reste. »
Vincent ROCA[1]

« Lorsqu'il est question de réseau et de territoire,
réalité et métaphore sont en interface. »[2]
Claude RAFFESTIN, 1996, p. 5

Si la citation de Vincent Roca se veut doter d'un sens du second degré, il n'en demeure pas moins que parmi les transformations de la société, celle initiée dans le monde des transports s'est imposée par des temporalités exceptionnelles (Studeny, 2009). Alors, l'essor des transports motorisés a permis le passage d'un rapport de vitesse de 1 à 44 entre la marche à pied et le TGV (Ollivro, 2000), tandis que le tramway, le corail puis le TER se trouvent à des situations intermédiaires. De nombreuses *traces* de cette évolution demeurent jusqu'à aujourd'hui, quoiqu'une large part appartienne désormais au passé. Si le bitume avait durablement dissimulé les rails du tramway qui serpentaient l'avenue Foch à Dijon jusqu'au début des années 1960, les premiers travaux de terrassement du nouveau tramway bourguignon les ont fait ressurgir : ces traces dissimulées sont désormais effacées. Aussi, le viaduc de Montsarin, en Saône-et-Loire, porte les vestiges de l'omnibus qui le traversait de manière quotidienne au XIXème siècle : il est aujourd'hui accolé à un autre viaduc qui lui est parallèle et qui doit supporter la charge des caténaires et les vibrations des TGV qui le traversent à près de 300 kilomètres à l'heure. La pratique quotidienne, depuis près de trente ans, de ces deux situations n'est sans doute pas étrangère à l'intérêt que nous portons au travail que nous développons ici. Ces exemples de *traces* d'une révolution des transports ont vocation à se multiplier dans l'Hexagone et interrogent avec force la première citation : les trains changent donc, mais que changent-ils ?

[1] Vincent ROCA est un écrivain et chroniqueur français, qui a reçu le Grand Prix Raymond Devos en 2011 et dont les textes sont régulièrement joués au Festival d'Avignon.
[2] RAFFESTIN Claude (2007). Préface, in OFFNER Jean-Marc, PUMAIN Denise (dir.) (2007). *Réseaux et territoires, Significations croisées*, Editions de l'Aube, La Tour d'Aigues, 284p.

Réseau et territoire : une dualité d'actualité différenciée par la vitesse

Le rail est porteur d'une infrastructure lourde, dont les effets ont été maintes fois étudiés et qui suscitent toujours aujourd'hui une étonnante actualité scientifique (Offner, 1993a, 2014) et fait même l'objet de rubriques spécifiques dans les revues scientifiques ("Les controverses de l'Espace Géographique", 2014). Elle se double d'une forte demande économique et sociale, parce que l'Homme est fasciné par la vitesse (Ollivro, 2000 ; Studeny, 1995), persuadé qu'il faut aller le plus vite possible, afin de toujours élargir son horizon spatio-temporel. L'étude d'un système de transport s'intéresse donc de manière inhérente aux rapports qu'entretiennent espace et temps (Volvey et al., 2005). Quand il est organisé, on lui prête le vocable de *« réseau »*, dont on soupçonne qu'il porte des effets qui se font sentir sur le *« territoire »* : l'essor de la vitesse les dissocie de manière intrinsèque par l'apparition de deux métriques différentes (Lévy et Lussault, 2003), mais la citation de Claude Raffestin montre comment ces deux concepts sont tout autant imbriqués.

Aujourd'hui, les problématiques de transport se manifestent partout, des sphères académiques jusque dans les sphères politiques, largement diffusées par les discours techniques et médiatiques. Elles se sont également multipliées avec l'essor de nombreux modes de transport, parmi lesquels l'automobile constitue un transport de masse depuis la fin de la seconde guerre mondiale (Dupuy, 1999) et l'essor de la grande vitesse ferroviaire depuis les années 1980 a remis au goût du jour l'attrait pour le chemin de fer (Auphan, 1989) : l'extension d'une offre de transport est alors perçue comme *« sinon un gage de progrès, du moins un facteur explicatif des modes d'urbanisation et d'aménagement »* (Offner, 1993a). La permanence des études portant sur les interactions entre réseau et territoire tient également du fait que le réseau est devenu le support d'un attribut de la modernité (Bavoux et al., 2005), où la mobilité qu'il engendre est pourvue d'une connotation très positive, au centre du fonctionnement de la société, parce que *« l'espace étant hétérogène, les transports fournissent le moyen de s'adapter par le déplacement à toutes les distributions différenciées »* (Bavoux et al., 2005), faisant de l'accessibilité une condition décisive. Le réseau devient donc un *« instrument du bien collectif »* (Dupuy, 1989), dans lequel la puissance publique voit *« un soutien non géographiquement discriminant à des opportunités de développement »* (Beaucire, 2014) : les politiques d'aménagement semblent s'inscrire dans *« une longue tradition de pratiques administratives diverses mises en œuvre par le ministère des Travaux publics »* (Caron, 2002). C'est pourquoi cette thématique de recherche rencontre un vif succès permanent en France. Parallèlement à cette fièvre des discours vantant les bienfaits de l'infrastructure, de nombreuses contributions scientifiques se sont attachées à atténuer cette bienveillance en montrant que l'effet n'est sans doute pas automatique (Plassard, 1977) et résulte d'une convergence de facteurs. Alors la dualité de ces deux concepts se mue en une interaction, où le pendant de ces effets supposés du réseau sur les dynamiques territoriales se concrétise quand certains concluent que l'installation d'une infrastructure vient *« conforter les positions économiques acquises »* (Wolkowitsch, 1980) et qu'elle s'en accommode pour se développer.

Ce travail de recherche se situe au cœur de cette interaction entre réseau et territoire, en plaçant l'espace et le temps au centre de l'analyse, en en faisant de potentiels facteurs d'explication. De ce point de vue, de nombreuses recherches ont été conduites pour démontrer les effets de grandes infrastructures de transport (Bérion, Joignaux et Langumier, 2007 ; Plassard, 1977), souvent à partir d'études de cas précis et contextualisés (Delaplace, 2014), dont les discours politiques et

médiatiques se sont souvent emparées à des fins exagérées de généralisation. Pour autant, les effets constatés relèvent finalement d'un processus qui est historique et socialisé (Bazin et al., 2011), dont les effets avant et après l'infrastructure peinent à être mesurés en raison d'un manque de données très précises (Bretagnolle, 2014). L'objectif de la thèse est donc de réinvestir les composantes de l'interaction entre réseau et territoire, dans un contexte plus large, qui nécessite d'intégrer les *traces de la vitesse* dans le processus d'analyse, par l'exploitation de larges bases de données spatio-temporelles. Cette analyse s'inscrit dans un contexte où l'apparition du chemin de fer au XIXème siècle a permis de multiplier les vitesses de circulation par quatre ou par cinq. Alors, *« la notion d'effets structurants n'est pas à proscrire et se révèle extrêmement importante lorsqu'on travaille sur les transports à des échelles spatiales et temporelles larges »* (Bretagnolle, 2014). En quoi la croissance du réseau ferroviaire français participe-t-elle à un développement des territoires qu'il dessert ? Quelles caractéristiques de l'infrastructure influencent davantage les dynamiques territoriales ? Ces questions interrogent le premier sens de la boucle d'interaction souvent mobilisé dans la littérature de long terme (Kasraian et al., 2016), alors que le second l'est plus rarement. Ainsi, dans quelle mesure le réseau ferroviaire français s'est-il adapté à la conjoncture des dynamismes locaux ? Dans quelle mesure les concepteurs du réseau ont intégré les dynamiques territoriales existantes dans les choix de la diffusion du réseau ? Cet ensemble de questions constituent le cœur de cette recherche, à partir de larges échelles.

Trace et tracé dans une approche par la donnée géohistorique

La réflexion par les échelles spatiales et temporelles nous paraît naturelle pour mener notre travail pour deux raisons. La première est que même inconsciemment, nous associons espace et temps quand nous parlons de nos déplacements (Ollivro, 2000) : *« je suis à Paris en l'espace de deux heures », « c'est à une heure de marche à pied »*. La seconde nous pousse à justifier ces rapprochements dans la mesure où l'historien Fernand Braudel affirme que le transport favorise l'échange en étant un formidable outil pour toute société économique en progrès (Braudel, 1980). Il faut alors envisager l'étude des effets d'une infrastructure de transport en prenant en compte trois entrées complémentaires : sa localisation, ses effets à d'autres échelles ainsi que les délais associés à ces effets (Banister et Berechman, 2004). Ce contexte appelle à interroger les *traces de la vitesse entre réseau et territoire*, à partir de la polysémie du premier terme. Elles n'évoquent pas seulement la *« trace »* laissée par le passé mais aussi le *« tracé »*, celui qui permet un cheminement et qui relie deux points. La *trace* est porteuse d'une marque spatiale, parce qu'elle s'inscrit de manière physique mais aussi d'une marque temporelle, parce qu'elle fait référence à une situation passée, qui n'est pas forcément révolue. Le *tracé* porte tout autant ces deux marques : il suscite un cheminement, qui traduit un déplacement physique, qui lui-même est forcément associé à une durée, traduite de plus en plus souvent dans les indicateurs d'accessibilité.

Cette double dimension prend corps dans cette thèse par une approche géohistorique de l'interaction entre réseau et territoire. Si la dimension temporelle est au cœur de la recherche fondamentale en géographie depuis maintenant plusieurs décennies, c'est le nécessaire recours à l'archive historique qui justifie notre approche. Conscients des fortes marques laissées par les

« traces du temps » (Durand-Dastès, 1990), nous misons sur leur exploitation pour éclairer des processus de long terme (Kasraian et al., 2016 ; Nijkamp, 1986), en adoptant une granularité spatiale et temporelle fine mais extensive dans l'utilisation d'une base de données spatio-temporelles. La longue durée dévolue à l'approche géohistorique (Grataloup, 2015) est celle que nous impose l'innovation ferroviaire dont les prémices remontent en France à la première moitié du XIXème siècle et que nous confrontons à la situation des quelques 36 000 communes. Appliquée aux interactions entre réseau et territoire, *« la géohistoire est un champ de recherche : il génère des questions qui n'auraient pas été soulevées autrement, il révèle des relations historiques qui n'auraient jamais vu le jour et il justifie les histoires sur lesquels nous construisons nos versions du passé »* (White, 2010, traduction de l'auteur). Il s'agit d'une scène qui nous permet de *« produire une méthode »* (Delaplace, 2014) qui place l'espace au cœur des raisonnements, au côté des dimensions temporelles. Par une approche par la géographie théorique et quantitative pour analyser une base de données spatio-temporelles, notre démarche vient se positionner dans les applications traditionnelles du cadre géohistorique (Gregory et Geddes, 2014) qui utilisent le Système d'Information Géographique (SIG). Le géographe a le droit ou plutôt *« le devoir de sélectionner dans le passé ce qui est nécessaire à l'interprétation du monde actuel et de son espace »* (Durand-Dastès, 1990), même s'il faut reconnaître qu'il n'y a recours que rarement, tant la dimension historique a souvent été mise à l'écart de la discipline géographique depuis plus d'un demi-siècle (Boulanger et Trochet, 2005 ; Volvey et al., 2005).

Pourtant, par notre formation en géographie, avec quelques heures de détour par l'histoire pendant le premier cycle universitaire, de même que par un attrait certain pour le monde des transports qui fascine tant les plus jeunes que les plus nostalgiques, nous misons sur les enseignements de la profondeur temporelle pour étudier les interactions entre le réseau et le territoire. A l'intersection de l'espace et du temps, de même qu'à l'intersection des transports, de l'analyse spatiale et de l'histoire, cette thèse poursuit l'objectif d'une meilleure connaissance des mécanismes qui ont conduit à la fièvre ferroviaire pendant près d'un siècle. Elle vise aussi à la sauvegarde d'une mémoire des *traces de la vitesse*, dans la mesure où, à notre connaissance, il n'existe pas un tel outil interactif rendant compte de la croissance du réseau ferroviaire française et de l'évolution technique de son infrastructure entre 1830 et aujourd'hui.

Ainsi, la caractérisation des objets de la recherche mise donc sur une couverture exhaustive du territoire et de la période de l'étude, dans la mesure du possible (Knowles, 2005). Les dynamiques territoriales sont envisagées par la taille des communes françaises, parce qu'elle porte en elle un *« résumé de très nombreuses propriétés fonctionnelles »* (Pumain, 1997). La croissance du réseau ferroviaire français se doit de prendre en compte l'évolution des tracés sur le territoire français, assorti d'attributs tant quantitatifs que qualitatifs, justifiés par l'intégration de la vitesse dans nos problématiques de recherche. Dans la ligne d'une double inscription de cette thèse dans une approche géohistorique, théorique et quantitative, les approches utilisées dans les analyses de l'interaction s'inscrivent dans l'analyse spatiale et la modélisation, afin d'envisager une *« représentation simplifiée de la réalité en vue de la comprendre et de la faire comprendre »* (Haggett, Frey et Cliff, 1977). La donnée géohistorique se retrouve au cœur d'un processus de connaissances qui concilie les théories d'organisation de l'espace, les contextes dans lesquelles les analyses prennent corps, de même que les nouvelles informations construites par l'exploration des données spatio-temporelles, qui renseignent tant sur les *traces* du réseau que sur ses *tracés*.

La base de données FRANcE au cœur des humanités numériques

Par l'utilisation d'une large base de données spatio-temporelles, rendant compte d'attributs spatiaux et temporels du réseau et du territoire, nous défendons par ailleurs l'inscription de cette approche géohistorique dans le cadre plus large des humanités numériques. Bien que les initiatives demeurent encore discrètes dans la géographie française, le double tournant digital et spatial a opéré dans les sciences humaines et sociales dans les années 1980 (Gregory et Geddes, 2014). Ce sont surtout des considérations techniques qui occupent les premières années de développement, et qui nécessitent un croissant dialogue entre le chercheur en sciences humaines et sociales et l'informaticien, de sorte à intégrer le matériau du premier dans les structures du second. Alors que le premier apporte son expertise sur l'archive et la source historique, le second met en œuvre une conceptualisation des données. L'interdisciplinarité est le gage de tels projets par la diversité et l'hétérogénéité des sources mobilisées, au rang desquelles la carte, la photographie et le texte figurent en bonne place (Dacos et Mounier, 2014). Alors, le géographe s'y insère progressivement quand le SIG devient un outil privilégié pour prendre en compte l'espace : assorti d'attributs spatiaux et d'attributs temporels (Gregory et Ell, 2007), il devient alors un Système d'Informations GéoHistoriques (SIG-H).

Cette première phase s'est concrétisée dans un projet européen menée au début des années 2000 – *« A GIS for the History of European Integration (1825-2005) : the European road and rail infrastructure »*[3] –, qui vise à numériser le réseau ferroviaire européen, à partir de cartes et de sources originales. Pour l'instant, les principales publications s'attachent à soulever les challenges de la phase de collecte (Morillas-Torné, 2012), ainsi que des études de cas nationales ou binationales (Martí-Henneberg, 2011). Les principaux enjeux d'une telle entreprise résident dans la collaboration des disciplines pour identifier les sources nécessaires, les dispositifs de traitement au sein d'une pratique qui se veut ouverte et qui défend le libre accès aux données et aux moyens de leur construction et de leur analyse. Pour autant, la fastidieuse phase de collecte concentre souvent le chercheur sur les moyens à mettre en œuvre pour constituer des bases de données spatio-temporelles.

L'objectif de cette thèse est de dépasser cette phase pour se projeter dans la construction de connaissances géohistoriques à partir du matériau collecté, en s'appuyant sur les défis des humanités numériques, parmi lesquelles le recours intensif à l'informatique et le recours diversifié aux méthodes occupent une place de premier ordre. Deux moyens sont envisagés dans la thèse pour se concentrer sur les phases de traitement et d'analyse d'une large base de données spatio-temporelles : l'enrichissement des données par une conceptualisation et une structuration spécifiques ainsi que la diversification du corpus à partir d'informations collectées par ailleurs et qu'il suffit d'intégrer au modèle existant. Ainsi, nous construisons une base de données FRANcE (French RAilway NEtwork) pour envisager *les traces de la vitesse entre réseau et territoire*, par une approche géohistorique de la croissance du réseau ferroviaire français, à travers la qualité de

[3] Le projet a été soutenu par *European Science Fondation*, à travers le programme *EUROCORES : Inventing Europe, technology and the making of Europe, 1850 to the present*. La numérisation du réseau ferroviaire français fait l'objet d'un package du projet

l'infrastructure et l'accessibilité qu'elle permet, en la confrontant à la structure démographique des 36 000 communes françaises.

Le cheminement de la thèse : une modélisation croissante

L'objectif de ce travail de thèse est d'interroger l'interaction entre réseau et territoire en agrandissant la focale classiquement à l'œuvre dans les analyses : comment la phase de la croissance du réseau ferroviaire français et la distribution des populations dialoguent dans un projet de recherche géohistorique, dont la granularité spatiale et temporelle nous permet d'envisager ces deux types d'effets sur le long terme ?

Ce questionnement s'intègre dans une démarche globale de modélisation, visant l'étude *« de phénomènes, qui possèdent entre eux certaines relations »* (Langlois et Reguer, 2005). Cette démarche globale se décompose néanmoins à travers l'adoption de postures différentes, qui participe à une complexification progressive de l'approche, de la définition des objets de recherche jusqu'à la recherche des relations qu'ils entretiennent. Ces postures traduisent le cheminement de la thèse.

La première partie se concentre sur la définition des objets de recherche et des relations qui les lient. Cette posture compréhensive vise à étudier le processus de conceptualisation du réseau, à partir des postulats théoriques et empiriques qui en sont fait. Elle vise à replacer la littérature abondante sur la longue durée, selon les approches disciplinaires et méthodologiques, afin d'analyser comment les échelles spatiales et temporelles sont envisagées voire articulées dans le couple transport et territoire. Cette première étape participe à la justification du recours à de larges échelles spatiales et temporelles, qu'il s'agit de replacer dans le contexte épistémologique de la discipline géographique, afin de positionner ce travail de recherche dans une approche géohistorique, favorisant les allers et retours entre théorie et empirie à partir d'une large base de données spatio-temporelles. La dernière étape consiste en l'identification des défis inhérents au projet de SIG-H quand il est confronté à l'archive pour la constitution de la base de données FRANcE, afin de récupérer *les traces de la vitesse*.

Une fois le matériau collecté, la seconde partie passe par une posture formalisatrice pour envisager les interactions entre réseau et territoire. Dans ce cadre, l'enjeu d'enrichissement de la donnée passe par la construction d'une accessibilité géohistorique dont l'objectif est d'enrichir les approches de la dotation du territoire en infrastructures. L'insertion dans les humanités numériques se concrétise par le nécessaire rapprochement avec les mathématiques et l'informatique pour formaliser le réseau de transport par la théorie des graphes. Ensuite, nous nous attacherons à envisager les conditions de la confrontation de nos deux objets de recherche dans des perspectives explicatives et interprétatives. Notre positionnement dans la valorisation de la donnée pour de nouvelles analyses prend corps dans la diversification des approches. Après avoir identifié dans les techniques économétriques des opportunités de valoriser les nouvelles informations, l'enjeu de

diversification de l'information passe par la sophistication de l'analyse causale de l'effet du réseau. Enfin, nous nous focaliserons sur les moyens à entreprendre pour envisager l'inverse de la boucle d'interaction, où la modélisation dynamique, largement utilisée en géographie, servira à la construction d'un modèle d'évolution d'un réseau de transport pour comprendre les mécanismes élémentaires, vidés des complexités politiques et économiques, de la croissance du réseau, à partir de structures stylisées.

Enfin, la dernière partie est le fruit du challenge identifié dans les humanités numériques, qui vise à partager de nouvelles connaissances géohistoriques assorties du contexte dans lesquelles elles prennent corps, à partir des postures envisagées dans la seconde partie. Nous nous attacherons alors à mesurer l'effet du réseau dans les dynamiques territoriales de long terme, à partir de l'évolution de l'accessibilité mais aussi par la prise en compte de contextes locaux, en diversifiant l'information et en s'assurant de la véracité des modèles produits. Il s'agira de mesurer la portée de l'effet du réseau, dans des perspectives à la fois spatiales et temporelles. Alors, nous verrons comment la mise en œuvre empirique du modèle d'évolution de transport participe à l'interprétation de l'autre sens de la boucle d'interaction, en réintégrant systématiquement des éléments de contexte et participant à la narration de versions du passé, nous permettant d'évaluer les forces de la rencontre des différents facteurs participant à l'établissement des tracés des chemins de fer sur le temps long, entre diffusion et hiérarchisation du réseau selon des règles d'efficience économique et démographique. Finalement, à la lumière de l'approche adoptée et des résultats obtenus, nous envisagerons des perspectives de recherche dont les propositions défendent la largeur des échelles et le regard géohistorique dans les humanités numériques.

PREMIERE PARTIE :

EFFET STRUCTURANT ET ECHELLES SPATIO-TEMPORELLES

INTRODUCTION DE LA PREMIERE PARTIE

« Aimons la vitesse, qui est le merveilleux moderne,
mais vérifions toujours nos freins. »
Paul MORAND

Quand la *Rocket* de Georger et Robert Stephenson atteint une pointe de 48 km/h en 1829 près de Manchester et que la chaudière Flaman à double-corps bat le record du monde de vitesse sur rail entre Sens et Montereau à 144 km/h en 1890, l'historiographie des techniques évoque le plus souvent les records de vitesses, parce qu'ils éveillent le rêve et la performance mais entretient aussi *« l'illusion d'une société se déplaçant d'un seul tenant »* (Flonneau et Guigueno, 2009). Pour autant, *« celui qui court plus vite avance-t-il davantage ? »* (Ollivro, 2009). Les nombreuses réponses à cette question semblent avoir oublié pendant longtemps l'éminent conseil de Paul Morand, tant l'effet structurant suscite l'adhésion du plus grand nombre. Quand les élus franciliens se sont battus dans les années 1950 et 1960 pour être desservis par le Réseau Express Régional, quand les élus ruraux se battent pour obtenir un échangeur d'autoroute ou une gare à grande vitesse, aucun frein ne vient nuancer cette croyance dans l'*« effet structurant »* des infrastructures de transport (Offner, 1993a ; Plassard, 1977).

Définis comme les *« modifications et adaptations de comportement des individus et des groupes qui se manifestent par des changements dans les structures économiques et qui résultent de la création et de la mise en service de l'infrastructure de transport »* (Plassard, 1977), les effets structurants ont été largement investigués tant par la littérature académique que par la presse et les discours politiques (Bazin et al., 2011). Le contexte français est selon nous propice à la permanence des contributions, parce que le développement de l'aérien a succédé à la démocratisation de l'automobile et surtout parce que la très grande vitesse ferroviaire a revivifié l'intérêt pour le train (Auphan, 1989 ; Flonneau et Guigueno, 2009). Certains déplorent même une *« littérature devenue trop abondante sur les réseaux »* (Dupuy, 1997). Les débats réguliers démontrent l'actualité scientifique qui caractérise l'étude des interactions entre transport et territoire (Offner, 1993a, 2014). Ils montrent aussi l'importance de la définition des objets de la recherche ainsi que des méthodes développées. Ces questionnements occupent tout autant le politiste, le sociologue, l'ingénieur, l'économiste que le géographe, malgré un vocabulaire qui peine à trouver des caractéristiques communes. Pour le géographe, *« lorsqu'ils s'inscrivent dans l'espace géographique, les réseaux organisent le territoire »* (Dupuy, 1989). Jean-Marc Offner y décèle une *« paresse intellectuelle des experts et des chercheurs »* (Beaucire, 2014), affirmation que de nombreux géographes ont récusée. De notre point de vue, derrière cette formulation à la portée quasi-universelle, notre raisonnement nous amène à préciser les cadres d'analyse de notre travail. Ainsi, *« le profil spatial, rural ou urbain, régional ou international, se modifie selon la perspective de la vitesse »* (Studeny, 2009). Pour autant, *« observés […] à l'aune du temps long et de l'ensemble du réseau, en utilisant des bases de données historiques construites pour permettre la comparaison à ces vastes échelles, les effets structurants sont bien là »* (Bretagnolle, 2014).

Cette première partie vise alors à saisir les contours du concept d'*« effet structurant »* des

infrastructures de transport, et questionne les échelles spatio-temporelles dans lesquelles notre travail vise à l'étudier. Quelles définitions donne-t-on aux objets de notre recherche pour revisiter les effets structurants des infrastructures de transport ? Quels bagages théoriques sont régulièrement mobilisés pour les étudier ? Quelles échelles temporelles et spatiales sont à privilégier ? Quel outil de recherche est le plus adapté pour notre travail ? Ces premiers chapitres proposent pour cela une approche historique et épistémologique du concept, qui aboutit sur une démarche de projet géohistorique.

Sans être exhaustif, nous proposons d'abord une approche par l'histoire du champ de recherche afin de comprendre quels sont les contextes d'émergence des objets de notre recherche. Ainsi, nous verrons comment l'effet structurant participe à la construction d'un mythe technico-social, même si l'objet « réseau » n'a pas été conceptualisé de manière concomitante (Dupuy, 1987a ; Dupuy et Offner, 2005). Si, deux siècles plus tard, de nombreux discours viennent alimenter ce mythe, alors les nombreuses contributions scientifiques sont le témoin d'une grande pluridisciplinarité (Dupuy, 1989), dans laquelle la place de l'espace est croissante et envisagée de manières différentes. Pour autant, l'ensemble des discours participe à l'entretien d'une controverse scientifique (Chapitre 1).

La large revue de la littérature scientifique sur les effets structurants nous conduit alors à mener une réflexion sur les échelles spatio-temporelles qui caractériseront notre travail de recherche. Dans ce cadre, les travaux de Christian Grataloup sur l'émergence d'une démarche géohistorique en France retiennent notre attention (Grataloup, 2015), alors que les rapports entre espace et temps sont encore ténus dans l'hexagone, là où les anglo-saxons s'en sont affranchis depuis longtemps. Dès lors, nous aurons à préciser notre posture épistémologique dans la construction de connaissances géohistoriques de même que notre positionnement méthodologique, qui saisit l'opportunité de l'émergence des humanités numériques dans le *« spatial turn »* des sciences sociales (Chapitre 2).

Enfin, la double réflexion sur le concept des effets structurants et sur notre posture de recherche nous amène à mobiliser l'outil SIG, appliqué à l'information géohistorique. Après avoir présenté les conditions de la construction de la base de données FRANcE (French RAilway NEtwork), nous verrons quels enjeux notre question de recherche pose en termes de collecte de données. De ce point de vue, nous replaçons la vitesse au cœur du dispositif, apportant une information inédite sur le patrimoine ferroviaire français, où la valorisation géolocalisée n'est encore que peu représentée dans l'histoire des chemins de fer en France. Par la suite, nous identifions les défis opérationnels liés à la mobilisation d'un tel outil, dans une perspective géohistorique. Ils montrent comment cette démarche peut difficilement être intégrée et automatisée, alors même que les premières exploitations montrent les premières limites de l'utilisation d'un unique outil (Chapitre 3).

Chapitre 1. L'EFFET STRUCTURANT DES INFRASTRUCTURES DE TRANSPORT : DE L'ENGOUEMENT A LA CONTROVERSE

Introduction

Ce premier chapitre vise à montrer comment le concept d'*« effet structurant »* des grandes infrastructures de transport s'est peu à peu installé, voire institutionnalisé, tant dans les sphères académiques que non-académiques. Dans la perspective géohistorique que nous adoptons dans ce travail de recherche, il nous semble plutôt naturel de recourir à la genèse de ce concept tout autant structurant de notre propos que des liens entre le réseau et le territoire. Les principales difficultés d'un état de l'art sont dans les contours qu'on leur donne : aborder l'histoire de son champ de recherche est une démarche récente des sciences humaines, parce qu'elle nécessite entre autres de sortir de son champ disciplinaire. Il s'agit d'aborder la diversité des méthodes, d'exposer des interprétations, de prendre en compte différentes temporalités, et de les manier. Faire un état de l'art nécessite de se confronter à plusieurs illusions. Toutes les dates n'ont pas de signification en soi. Tous les personnages et auteurs impliqués ne le sont pas au même degré. Aussi, aborder son champ de recherche sous l'angle (géo-) historique mène à deux ambiguïtés. La première est le défaut d'historicité : il faut tout aussi être prudent dans les dangers anachroniques que dans l'interdiction exagérée de rapprocher des phénomènes de temporalités différentes. La seconde est le défaut de rationalité : une démarche positiviste voudrait que la science se construise seule.

Une fois ces difficultés intégrées, ce chapitre vise à montrer un certain décalage entre la construction du « mythe de l'effet structurant » qui apparaît de manière intense dans la première moitié du XIXème siècle et la construction de l'objet scientifique « réseau ». Dès lors, les saint-simoniens se sont attachés à théoriser les vertus de la mise en réseau avant que le réseau ne devienne un objet de recherche à lui seul (1.1.). De plus, si nous avons replacé l'émergence de l'*« effet structurant »* dans le contexte français, nous verrons comment l'évolution structurelle des sociétés a fait du lien entre le réseau et le territoire un champ de recherche international, pluridisciplinaire, dans lequel les géographes ont peiné ou peinent encore à trouver leur place (1.2.). Pourtant, le contexte français semble particulièrement propice pour les géographes, mais pas seulement, à une remise en cause de l'effet causal du réseau sur le territoire en laissant place à davantage de nuances (1.3.). Bien que la notion d' « effet structurant » ne soit surtout pas vouée à disparaître, elle devient le socle d'une controverse intemporelle, qui la fait demeurer d'une permanente actualité scientifique.

1.1. Les vertus de la mise en réseau

L'emploi du terme de *« réseau »* est aujourd'hui largement diffusé tant dans la sphère académique que dans la vie quotidienne. Ainsi, l'analyse des réseaux en sciences sociales serait pour certains un effet de mode (Lemercier, 2005), s'attachant à l'analyse des réseaux de famille, d'alliance, des espaces de connivence, de pouvoirs mais aussi plus largement aux réseaux professionnels ou aux incontournables réseaux sociaux. Assez naturellement, un essai de définition se heurte aux difficultés de consensus (Parrochia, 2005). Il est au départ statique, outil capable de capturer des objets, parce que son étymologie « *reseuil* » évoque le filet. William Harvey le projette pourtant au XVIIème siècle dans une nouvelle dimension dynamique, où ce concept lui permet d'établir une des premières véritables théories de la circulation sanguine. Depuis, le réseau est largement associé à ses capacités circulatoires, où le sang circule dans les artères, et se diffuse dans les organes, dessinant par la même occasion les lignes et les points, objets bien connus des géographes.

Pour autant, dans notre discipline, il s'agit d'un concept non seulement récent, mais certaines définitions souffriraient encore de leur caractère « inachevée » (Lévy et Lussault, 2003). En tout cas, force est de constater que cette notion fait l'objet d'une *« polysémie problématique »* parce qu'il peut être défini indépendamment par des propriétés techniques, économiques, sociologiques, spatiales et territoriales (Offner, 1996).

Pour exemple, issu des Hautes Etudes Commerciales, Maurice-François Rouge s'est vite consacré à l'aménagement et a évoqué dès 1953 l'intérêt qu'il y aurait à transformer l'étude de réseaux en une véritable discipline scientifique (Rouge, 1953 ; Rouge et Dupuy, 1989). Son appel n'a non seulement pas été suivi d'effets, mais cette thématique disparût de ses publications suivantes, et il faudra attendre la fin des années 1980 et la constitution du groupe Réseaux autour de Gabriel Dupuy.

Cette partie vise à montrer la trajectoire du concept de réseau tel qu'il est approprié par les ingénieurs et les scientifiques, mais aussi les intellectuels, depuis le début du XIXème siècle. Il est alors vu comme un instrument indispensable pour justifier et légitimer la construction de grandes infrastructures de transport, gage de stabilité économique, politique et sociale. Cette partie éclaire la mise en place progressive d'une théorie des réseaux, qui sert encore aujourd'hui de *« rhétorique des effets structurants »* afin d'en montrer la permanence sur la longue durée.

1.1.1. **Le saint-simonisme : structurer par la connexité**

Dès 1832, Michel Chevalier déclare avec force que *« les chemins de fer changeront les conditions de l'existence humaine »* (Chevalier, 1832), alors que l'on se déplace péniblement à pied, au mieux à cheval. Sans parler d'*« effets structurants »*, notion sur laquelle nous aurons à revenir dans la suite de notre propos, les bienfaits des infrastructures de transports sont institutionnalisés très tôt au XIXème siècle en les érigeant comme un *« mythe légitimant, une notion complexe dans laquelle se mélangent de manière indiscernable, indissociable et indissoluble le réel et le mythe »*

(Chevallier, 2012)[4]. Michel Chevalier, Ingénieur du Corps des Mines, appartient au courant de pensée hérité de Saint-Simon. Claude-Henri de Rouvroy, comte de Saint-Simon (1760-1825) est parfois qualifié comme celui dont l'œuvre a inspiré la formation des « sciences humaines » au début du XIXème siècle (Musso, 1999). Il est surtout l'inspirateur de l'école saint-simonienne, dont les principaux leaders furent précisément Michel Chevalier (1806-1879) mais aussi Barthélémy-Prosper Enfantin (1796-1864). Ce dernier fut à l'origine du canal de Suez, mais surtout l'un des créateurs de la Compagnie des chemins de fer de Paris à Lyon, avec le banquier Laffitte.

Instigateurs d'une interprétation de l'œuvre de Saint-Simon, les saint-simoniens sont porteurs d'une doctrine sociale, spirituelle, politique et religieuse, issue des Lumières, où l'objectif est *« de développer et de répandre les principes d'une philosophie nouvelle […], [qui] reconnaît que la destination de l'espèce […] est d'exploiter et de modifier à son plus grand avantage, la nature extérieure »* (*Le Producteur*, II ; 159). A l'issue de la révolution de 1830, leur doctrine se transforme en pratique. Ainsi, les premiers écrits qui évoquent la transformation des infrastructures de transport en un réseau de transport sont assez naturellement le fait d'hommes politiques, ingénieurs et proches du pouvoir. Hérités de l'Ancien Régime, ils font partie des « Grands Corps de l'Etat ». Loin d'être sous le joug du pouvoir politique centralisé, ils *« participent d'un même système de pouvoir […] conquis au cours d'un long processus d'autonomisation, un jeu d'influences externes »* (Kessler, 1986), se faisant alors *« des hommes politiques à titre au moins égal à celui des raisonneurs »* (*Le Globe*, 16 décembre 1831). Une part importante des disciples du courant saint-simonien a joué un rôle majeur dans le développement des chemins de fer en France.

L'appropriation du concept de réseau par les saint-simoniens leur permet ainsi de réunir deux courants de pensée parmi les plus répandus au XIXème siècle : l'économie de la circulation et l'utopie du lien universel (Mattelart, 1994). Cette union permet à Gabriel Dupuy de résumer l'ensemble de la doctrine saint-simonienne en affirmant que *« tout peut être mis en relation avec tout et engendrer un mouvement de circulation des personnes, des biens, des capitaux, des idées, du savoir »* (Dupuy, 1989). La liaison de ce tout est permise *« par un lien relativement matériel, c'est-à-dire par des voies de transport, et par un lien relativement spirituel, c'est-à-dire par des banques »* (Chevalier, 1832). Si de nombreux travaux taxent la doctrine saint-simonienne d'historiciste, elle a parfois été qualifiée d' *« idéologie simpliste »* et de *« vision naïve du monde »* (Gras, 1998), il reste à préciser les fondements de la théorie saint-simonienne qui institutionnalise le réseau comme *« premier pas vers l'association universelle »* (Chevalier, 1832) :

- la doctrine sociale saint-simonienne consiste en un principe d'égalité parfaite entre les Hommes, permise par le développement des activités agricoles, industrielles et commerciales ;
- la doctrine politique saint-simonienne consiste en une association des Hommes par l'établissement de liens fédérant l'ensemble du territoire national ;
- la doctrine spirituelle et religieuse saint-simonienne consiste en une organisation de la société qui pousse l'Homme vers le meilleur et améliore l'existence de son semblable.

[4] On veillera bien entendu à distinguer Jacques Chevallier, l'auteur de l'ouvrage *Le service public* en 2012, de l'économiste et politique Michel Chevalier, leader du mouvement saint-simonien au XIXème siècle.

A partir de ces différentes composantes de la doctrine saint-simonienne, on lit facilement la volonté d'établir *« un système général de communications et de transport »* où l'association universelle permet aux Hommes d'exploiter leurs facultés, de les mettre au service de la transformation de leur environnement (il s'agit de la *« nature extérieure »* évoquée dans la revue *Le Producteur*), et ainsi de mettre en œuvre de nouveaux outils. La société visée est celle du progrès dans un contexte politique dominé par la puissance de l'Etat et un contexte économique par le fort développement industriel, où le réseau devient un *« instrument du bien collectif »* (Dupuy, 1989) et la condition d'une expansion économique et sociale. D'après ce courant saint-simonien, réfléchir au lien entre le réseau et le territoire revient à étudier la manière par laquelle le réseau permet de connecter les Hommes. On touche là à une composante primaire des réseaux, tels qu'ils vont être conceptualisés par la suite, la connexité *(cf. 1.1.3, p. 30)*.

Ainsi, plus tard, en 1928, Mark Jefferson, professeur de géographie dans l'université du Michigan affirme que *« l'Europe de l'Ouest est indubitablement l'une des régions avec le plus haut niveau de culture au monde. La preuve est la couleur blanche* (Figure 1. 1) *qui révèle que l'Angleterre, les Pays-Bas, la Belgique, le Danemark, l'Allemagne, la Tchécoslovaquie et la Hongrie comme n'ayant aucun point éloigné de plus de dix miles du chemin de fer. Et il n'y en a d'ailleurs pas plus dans la plus grande partie de la France, la Suisse, l'Autriche ainsi que la moitié sud de la Suède. Cela signifie que chaque champ, chaque usine, chaque forêt, chaque mine, chaque village, chaque maison de ces régions bénéficie d'un accès facile au transport ferroviaire. L'agencement hérité du siècle dernier a permis plus que tout autre invention de l'Homme de transformer la vie humaine, et en particulier d'élever la société vers l'avant, de même que les classes les moins aisées »* (Jefferson, 1928).

Chef de la cartographie de la délégation américaine pour la Conférence de Paix de Paris en 1919, ce fin connaisseur de l'Europe est l'un des premiers géographes à reprendre le « mythe » unificateur du réseau de transport. Pour lui, la mobilité permise par le chemin de fer *« transforme et ennoblit les Hommes »* (Jefferson, 1928), de la même manière que la mobilité permise le long du Nil a fait de l'Egypte un pays qui a compté dans l'Histoire. Ainsi, le chemin de fer, tel qu'il est vanté par les ingénieurs des Ponts et Chaussées est le gage de l'association universelle, par la nouvelle connexité qu'elle permet, au-delà de la question de la rapidité qui n'était pas au cœur de la pensée de l'époque (Musso, 1999). Jefferson se contente d'étudier la répartition du réseau, indépendamment de la vitesse à laquelle on y circule. Mais au-delà de lier la population, passant de la traction animale à la traction mécanique, le chemin de fer a permis de *« bouleverser les cercles locaux d'existence »* par la *« révolution de la vitesse »* (Studeny, 1995), même si nous reviendrons sur des considérations chiffrées plus tard dans notre propos.

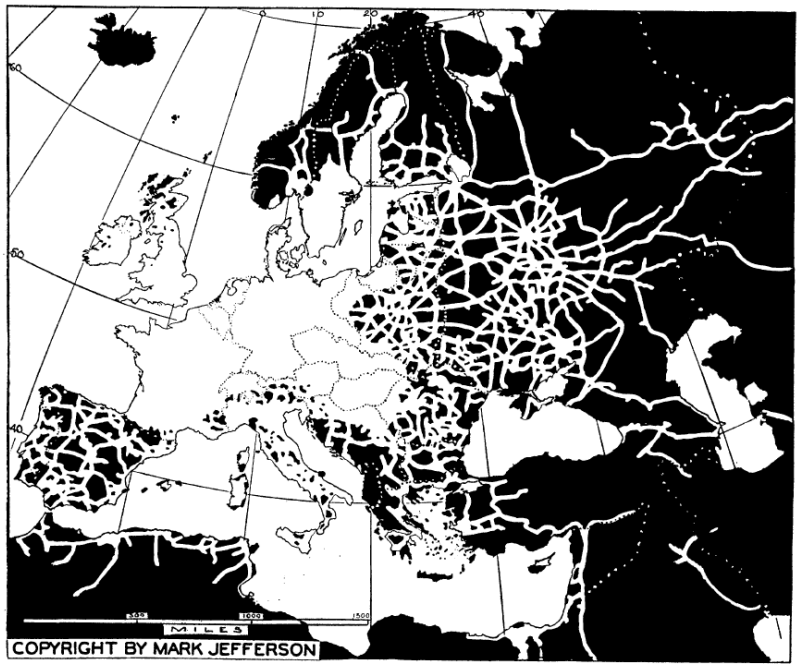


Figure 1. 1. La distance au chemin de fer en Europe dans le premier tiers du XXème siècle (Jefferson, 1928)

La doctrine saint-simonienne voit donc l'apparition d'une première définition du réseau de transport. En 1862, Léon Lalanne, ingénieur des Ponts et Chaussées, présente à l'Académie des Sciences un *Essai d'une théorie des réseaux de chemins de fer fondée sur l'observation des faits et sur les lois primordiales qui président au groupement des populations*. Il y donne une des premières définitions du réseau, comme ensemble des *« voies de communication qui sillonnent les pays civilisés, s'infléchissent dans leurs directions et s'entrecoupent mutuellement de manière à recouvrir le territoire d'une sorte de tissu, de filet à mailles inégales, à contours irréguliers. L'expression de réseau caractérise donc d'une manière heureuse l'ensemble des voies de communication d'un Etat »* (Lalanne, 1863).

De notre point de vue, la visée saint-simonienne de connexité se rapproche de théories de la complexité développées d'abord dans les sciences « dures » mais de plus en plus utilisées dans les sciences humaines et sociales (Pumain, 2003). Ainsi, la fièvre du chemin de fer viserait la construction d'un réseau, que l'on pourrait qualifier de *« structure émergente »*, où la mise en complexité serait le gage d'un *« système général de communications et de transport »* (Chevalier, 1832). Cette analogie est d'autant plus justifiée aujourd'hui que les défenseurs de la complexité pensent que *« plus les relations internes d'un corps sont nombreuses, subtiles et complexes, mieux ce corps est organisé et plus il agit sur son environnement »* (Musso, 1987). Dans ce cadre, cette première acception du concept de réseau le fait quitter la scène abstraite de la métaphore circulatoire vers une inscription durable et physique dans l'espace.

1.1.2. **L'unité nationale : structurer le contrôle politique**

Si la construction du « mythe » de l'effet structurant du réseau de transport passe par la totalité et l'exhaustivité, alors sa visée « connexioniste » fait que *« lorsqu'ils s'inscrivent dans l'espace géographique, les réseaux organisent le territoire »* (Dupuy, 1987b). De leur côté, les versants politique, économique et social de la philosophie des réseaux (Musso, 1999) sont infléchis par le « mythe » de la stabilité politique, dans un contexte intérieur et extérieur souvent perturbé. Nous verrons également que ce constat ne se limite pas seulement à la France. Cette lecture s'inspire des relations que Claude Raffestin théorise entre réseau, territoire et pouvoir. Pour lui, *« tout réseau est une image du pouvoir du ou des acteurs dominants »* (Raffestin, 1980). Alors, le réseau de transport, par la matérialité qu'il donne au territoire, est un instrument parmi d'autres de la concrétisation d'un projet sur un territoire défini (Debrie, 2008).

Notons toutefois que l'étude de la portée politique des réseaux de transport est antérieure à la pensée formulée par Claude Raffestin. Plusieurs auteurs ont montré comment la détermination des frontières fait partie des composantes du développement d'un réseau national : des études empiriques ont analysé l'impact du réseau ferroviaire dans les Balkans (Beaver, 1941) et en Roumanie (Turnock, 1979). Plus récemment, une étude portant sur la Belgique a montré comment ce nouvel Etat de 1815 s'est structuré autour du seul chemin de fer (Beyer, 2011). Conscient de sa place centrale au cœur du commerce européen, la construction du réseau ferré belge est *« un investissement économique rentable, un gage de cohésion interne et une grande opération de prestige diplomatique »* (Van Der Herten, Van Meerten et Verbeurgt, 2001). Ses tensions permanentes avec le Royaume de Hollande lui font perdre sa suprématie sur le Rhin, la Meuse et l'Escaut. Ainsi, dès le milieu du XIXème siècle, le réseau ferroviaire belge est le plus dense d'Europe et vient concurrencer le Rhin réaménagé en reliant Anvers et son arrière-pays par une ligne surnommée le « Rhin d'acier ».

Pour autant, le réseau de transport n'embrasse pas toujours l'enveloppe territoriale. Le développement des réseaux de transports dans les colonies en est un exemple, dont les conséquences sont encore perceptibles et tenaces aujourd'hui. En étudiant la relation entre tracé et frontière, Jean Debrie montre que *« l'évolution des réseaux reste centrée sur des bases nationales »* (Debrie, 2008) mais néanmoins très limitées, puisque c'est son lien avec la métropole (Bavoux, 2000) qui structure leur présence dans les espaces littoraux, et leur absence dans les espaces continentaux. A travers l'exemple des réseaux en Afrique de l'Ouest francophone, il montre que l'intégration régionale (au-delà des frontières nationales) est largement absente, plusieurs décennies après l'indépendance de ces pays.

A l'inverse, les recompositions territoriales peuvent poser le problème des recompositions des réseaux de transport. Trevor Howkins examine l'héritage dont bénéficie le nouvel Etat polonais en 1920 et comment il affecte le réseau ferroviaire préexistant et les futures nouvelles constructions (Howkins, 1996). Il pointe notamment des différences de normes entre les pays et voit dans la démarcation russe, dont l'écartement des rails est de 1,52m au lieu de 1,44m, une volonté de se différencier des puissances occidentales. Bien que la géographie du réseau ferroviaire ait été appréhendée dans le dessin des nouvelles frontières, la Pologne ne possédait pas d'accès direct à la Baltique, devait négocier des corridors de transit avec ses voisins et héritait d'un réseau tantôt

orienté vers l'Allemagne, tantôt vers la Russie. La construction de nouvelles lignes au centre du pays pour connecter le réseau a eu des conséquences économiques désastreuses sous une pression extérieure croissante (Howkins, 1996).

Partant de ces quelques exemples, il ressort que le mythe « connexionniste » du réseau atteint ses limites quand il est confronté à des contextes historiques et politiques complexes. Des études régionales ont montré comment il a pu être le ciment de la construction nationale, mais aussi un frein à l'unité dans le cadre de recompositions territoriales, parce qu'il est porteur de fortes inerties. Qu'en est-il de l'exemple français, berceau de la pensée saint-simonienne ? Dans quelle mesure participe-t-il à l'émergence et à la stabilisation d'un Etat-nation ou au moins d'une unité nationale ? Le réseau est-il le garant d'une homogénéisation toujours salvatrice ?

Le contexte dans lequel s'est développée la doctrine saint-simonienne est celui d'une instabilité gouvernementale qui a conduit à la Monarchie de Juillet en 1830. Pour autant, l'instabilité politique va demeurer jusqu'à l'installation durable de la IIIème République, alors que se sont succédé auparavant la monarchie (1830-1848), la Seconde République (1848-1852) et le Second Empire (1852-1870). Pendant cette période, le rapprochement entre ingénieurs et politiques s'intensifie jusqu'à s'entrecroiser. La composition du corps des Ponts et Chaussées montre ainsi la proximité de la sphère politique et de la sphère de l'ingénierie civile, à l'origine des grandes infrastructures, qui ne se limitent pas au chemin de fer. Leur proximité vaudra la qualification parfois galvaudée de nombreux tracés de *« chemins de fer électoraux »* (Wolkowitsch, 2004). Des historiens verront dans le déploiement du rail l'extrême *« républicanisation de tout le territoire et le pousser à la conquête de ces cellules élémentaires de la vie publique que constituent les cantons. […] Programme cohérent issu du même fond d'optimisme humain, de solidarité égalitaire et de propagande partisane »* (Blanchard, 1942). Quelles sont les motivations des élus à obtenir l'accès au réseau ferroviaire du XIXème siècle ? C'est d'abord participer à la fièvre du rail, dans son acception humaniste, mais c'est aussi créer les conditions favorables à ce que le territoire ne soit pas marginalisé du fait de l'absence du réseau. De manière indirecte, on suscite le besoin du rail, comme condition d'adhésion à un projet national.

Ainsi, au-delà des visées humanistes et universalistes du rail, des voies navigables et des routes, c'est aussi un contrôle politique voire géopolitique que permet la construction d'un réseau : *« améliorer les communications d'une extrémité à l'autre de l'espace saisi afin de le maîtriser encore mieux est une constante de tout Etat »* (Grataloup, 1996). Ainsi, la mise en place des réseaux, comme maillage d'un territoire, apparaît rapidement comme une composante importante de l'organisation spatiale des Etats, et surtout du contrôle qu'ils y exercent. Jusqu'au XIXème siècle, si les routes ne sont pas forcément pensées comme composantes d'un réseau en tant que tel (Lepetit, 1986), elles sont construites pour assurer la défense du territoire et la perception de l'impôt (Ollivro, 2000). D'ailleurs, l'établissement de la première carte de Cassini, rendant compte pour la première fois de la totalité du territoire français, n'est-elle pas une volonté centralisatrice de contrôle du territoire par la connaissance de ces moindres recoins. Aussi, la croissance du réseau ferroviaire français s'inscrit dans une dynamique de croissance des nationalismes et d'installation durable des Etats-nations. Le renforcement de la cohésion interne passe alors par un objectif simple : aller au plus droit pour aller au plus vite. Pour cela, *« la forme étoilée intervient lorsqu'un pouvoir politique veut asseoir durablement sa domination politique »* (Ollivro, 2000) : c'est bien le cas du réseau français. On retrouve cette structure autour de Lyon à l'époque gallo-romaine. Avec le renforcement du pouvoir royal à l'époque moderne, *« la réalisation de tracés directs au départ de la capitale*

permet d'inscrire sur le territoire national la suprématie du pouvoir dirigeant » (Arbellot, 1973). On retrouve cette même forme dans le tracé des routes impériales du début du XIXème siècle (Reverdy, 1995).

Là où le mythe « unificateur » met en exergue la dimension connexionniste du réseau, le « mythe » de la stabilité politique, à travers le contrôle qu'il induit, met en valeur une autre dimension fondamentale des réseaux illustrée par les nœuds qu'il relie. En 1842, la *Loi relative à l'établissement des grandes lignes de chemin de fer en France*, initiée par Legrand, a été un tournant dans le développement du ferroviaire (Caron, 1997) en reproduisant la structure héritée des réseaux passés autour de la capitale. Dès lors, le réseau en étoile donne au point central une *« efficacité nodale hors pair »* (Ollivro, 2000). L'étoile a donc une dimension centripète qui oblige le passage par son centre : tout en permettant de créer des liens avec la province, Paris assure une mainmise sur l'espace. Dès lors, *« tout flux tend à être rattaché au centre, tend naturellement à être aspiré par lui »* (Ollivro, 2000). En revanche, de cette sorte, la structure stellaire fait éclater le réseau en deux grands ensembles distincts. Le premier regroupe ces fameuses branches qui participent à la nodosité du point central, Paris dans notre cas ; tandis que le second regroupe l'ensemble des autres branches. Partant de Paris, on va donc retrouver des lignes à longue distance, où la tendance pousse à relier les principales villes, comme édicté dans la loi Legrand de 1842. Le réseau voit donc l'apparition d'une seconde acception, où apparaissent les notions de *« relation, de maillage et de service »* (Offner, 1996).

Alors que l'ambition *« connexionniste »* défendue à de nombreuses reprises dans *Le Globe* ne semble pas souffrir d'opposition ferme, la doctrine politique saint-simonienne a suscité davantage d'interrogations et d'oppositions. Son principe *« à chacun selon ses capacités, à chacun selon ses œuvres »* place l'industriel dans une position autoritaire et aurait alors abouti à une hiérarchie positiviste dans laquelle l'intervention centralisée ouvrirait la voie à la *« fatalité de l'oppression »* (Marcel, 1975). Comme déjà évoqué, la question de la rapidité n'est pas la première dans la constitution d'un réseau ferroviaire national. Mais déjà, au XIXème siècle, les vitesses élevées ne sont effectives que sur quelques liaisons, alors que les temps de parcours sur le reste du réseau restent encore importants. Ainsi, la construction du réseau ferroviaire s'accompagne du *« processus qui conduit à la mise en place de territoires parcourus à des vitesses de plus en plus variables »* (Ollivro, 2000). Ce peut être lu comme un moyen indirect d'un contrôle accru d'un pouvoir centralisé mais ce peut être tout aussi lu comme un phénomène de sélection. Surtout, c'est la première fois que l'Homme est confronté à des vitesses de déplacement très différentes.

Le rapport à l'espace et au temps se trouve pour la première fois différencié, alors qu'un nombre de personnes de plus en plus important ont voyagé en train depuis 1840. Le « mythe » généré par le réseau et ses capacités de vitesse a alors occupé la large scène de ces deux derniers siècles. L'émergence de la hiérarchie du réseau étoilé sur Paris pour un meilleur contrôle aurait donc une double conséquence : une subordination à l'infrastructure et une subordination au temps lui-même. L'ensemble de ces excès est largement abordé dans les écrits de Paul Virilio, qui dénonce une course irrépressible à la vitesse, qui pousse alors à une différenciation de plus en plus forte (Virilio, 1995). Le rapport à la vitesse n'aurait-il alors été pas assez pensé dans la théorie saint-simonienne ? Rien n'est moins sûr, quand le philosophe affirme que *« dans le domaine de l'aménagement du territoire, le temps l'emporte désormais sur l'espace, mais il ne s'agit plus*

comme naguère d'un temps local et chronologique mais d'un temps mondial, universel, qui s'oppose non seulement à l'espace local de l'organisation foncière d'une région, mais à l'espace mondial d'une planète en voie d'homogénéisation » (Virilio, 1995). Ainsi, un réseau poussé à son extrême aboutit à la création d'une *« dromosphère »* (Virilio, 1977) dans laquelle l'accélération des vitesses de transport aboutit à un rétrécissement de l'horizon spatio-temporel, voire communicationnel quand on y ajoute l'intensification des moyens de télécommunication, initiés eux aussi au XIXème siècle. Ainsi, adoptant la métaphore de la *« dérive des continents »*, on peut lire l'expansion d'un réseau de transport comme celle qui permet la reconquête d'un territoire ou alors celle qui permet le renforcement des nœuds au profit de quelques points du territoire, faisant du reste un territoire secondaire … Alors, si le marxisme s'est largement inspiré de la pensée de Saint-Simon, parce qu'il le considère comme un père de la sociologie (Musso, 1999), il l'accuse aussi de *« socialisme impérial »* qui fait de l'industriel un dominant oligarchique, un régulateur de la société venu d'en haut, alors que la pensée marxiste pense au contraire que c'est le groupe le plus nombreux, le prolétariat, qui détiendrait les clés du progrès.

Faisant du fer un vecteur de progrès et de cohésion, les acceptions de réseau évoquées jusqu'à maintenant poussent à la différenciation progressive des territoires, dont nous aurons l'occasion de revenir à de nombreuses reprises dans ce manuscrit. Les objectifs ne sont plus seulement philosophiques ou politiques, ils penchent aussi rapidement vers la sphère économique : *« Richesse et vitesse, voilà ce que le monde admire et ce vers quoi chacun tend. Chemins de fer, postes rapides, bateaux à vapeur et toutes les facilités de la communication, voilà où s'en va le monde de la culture pour se surpasser, se sur-cultiver et par là, persévérer dans la médiocrité »* (Goethe, p.13). Ainsi, à l'épreuve du débat, la pensée saint-simonienne semble se heurter à de croissantes critiques, qui voient dans sa doctrine la justification d'un pouvoir central renforcé, dont la planification est l'un des instruments.

A la définition de Léon Lalanne de 1863[5], on y insère donc le point, comme lieu de croisement des lignes, et autour duquel se manifeste une force, tantôt centrifuge, tantôt centripète, selon la fonction qu'on lui assigne. En se différenciant sur le territoire sur lequel il s'implante, le réseau oblige à la construction d'une nouvelle métrique. D'après Jacques Lévy, le réseau se définit comme un *« espace à métrique topologique »* par opposition au territoire dont la métrique est *« topographique »* (Lévy et Lussault, 2003). Dans ce cadre, Maurice-François Rouge aurait donc raison, dès 1953, de faire du réseau un objet scientifique nécessitant la construction d'une nouvelle discipline. A défaut, de multiples chercheurs d'horizons divers se sont attachés à doter ce concept d'un contour notionnel et méthodologique, dont nous discutons les principaux points dans la section suivante.

1.1.3. **Du mythe à la théorie des réseaux**

[5] *« Voies de communication qui sillonnent les pays civilisés, s'infléchissent dans leurs directions et s'entrecoupent mutuellement de manière à recouvrir le territoire d'une sorte de tissu, de filet à mailles inégales, à contours irréguliers. L'expression de réseau caractérise donc d'une manière heureuse l'ensemble des voies de communication d'un Etat »*

Par l'introduction de la notion de métrique définie comme un *« mode de mesure et de traitement de la distance »* (Lévy et Lussault, 2003), nous faisons du réseau un objet que l'on peut différencier du territoire. Dès lors, il faut dépasser le simple essai de définition proposé par Léon Lalanne en 1863 pour le doter d'un ensemble conceptuel stable et de propriétés communes. L'objectif est ici de montrer comment la construction de la théorie des réseaux est intrinsèquement liée au « mythe » qu'ils génèrent depuis deux siècles. Si certains craignent la dérive anachronique, rappelons par exemple que l'accroissement des routes aux XVIIème et XVIIIème siècles fait que *« l'ensemble des points de départ et l'ensemble des points d'arrivée sont connexes. Le système se comporte comme un réseau »* (Guillerme, 1991). Aux prémices de la construction d'une théorie des réseaux, on part bien d'une multitude de points que le réseau solidarise par la construction de lignes. Dans cette section, nous ne nous limiterons pas aux seuls réseaux de transport, mais par analogie, évoquerons également d'autres réseaux parce qu'ils partagent un certain nombre de propriétés qui viennent éclairer l'effet structurant des réseaux. Parmi eux, on ne retiendra en revanche que les réseaux qualifiés de *« techniques »*, c'est-à-dire ceux qui contiennent un *« apport scientifique substantiel »* qui ont *« un impact vital sur la façon dont un réseau est organisé et exploité »* (Salsbury, 1995). Ainsi, la mesure d'un réseau doit être capable de rendre compte de son organisation d'une part, de son exploitation d'autre part : l'objectif n'est pas historique mais est celui de la compréhension des modalités d'évolution des réseaux, répondant de manière plus ou moins continue au mythe de l'effet structurant.

Le *Groupe de Recherche Réseaux*, créé en 1983, a grandement contribué à l'établissement d'une théorie des réseaux, même s'il fédère davantage d'économistes, ingénieurs, historiens et sociologues que de géographes dans un premier temps (Dupuy et Offner, 2005). L'objet *« réseau »* n'a souvent pas été évoqué dans les théories urbanistiques de la première moitié du XXème siècle au moins (Dupuy, 1991), alors qu'il n'était abordé qu'à une échelle plus large, pour l'essentiel interurbaine. De même, l'intérêt économique des réseaux s'est longtemps cantonné aux seuls réseaux de télécommunications, jusqu'à il y a une vingtaine d'années (Dupuy, 1993). Pour autant, la théorie des réseaux appréhende cet objet selon trois entrées (Dupuy et Offner, 2005), qui sont de notre point de vue trois types de réponses différentes de l'effet structurant :

- Une **structure** : *« morphologie d'agencement des lieux et de leurs connexions » ;*
- Une **dynamique** : *« temps forts pour des espaces forts » ;*
- Une **logique** : *« lois d'évolution »*.

Ce découpage rapproche de manière quasi inéluctable les réseaux techniques des réseaux territoriaux (Dupuy, 1987a). La **structure** d'un réseau est alors régit par trois axiomes, résumés dans la *Figure 1.* 2.

La première abstraction qu'il propose rapproche le mythe *« connexioniste »* des saint-simoniens de la théorie de la géographie du pouvoir (Raffestin, 1980) : les objectifs de mise en réseau partent du constat d'une *« diversité et d'une hétérogénéité spatio-temporelles »* (Dupuy, 1987a). Ce constat est la condition de la création d'un réseau : l'association universelle vise à lier ces points. En revanche, le « mythe » ne mentionne pas la différenciation déjà effective de ces points *(Figure 1. 2, 1a)*, caractérisés par des niveaux de population, des niveaux de pouvoir différents.

La seconde abstraction est celle des *« projets transactionnels »* où *« l'espace est transformé en territoire par la visée, l'intentionnalité de l'acteur »* (Raffestin, 1980). Elle dépend en partie de la différenciation des points issus du premier axiome. Le « mythe » des effets structurants renforce alors cette hétérogénéité par les visées politiques, géopolitiques qu'il défend. Ainsi, on distingue bien le projet transactionnel issu de l'Etoile de Legrand *(Figure 1. 2, 1b)*.

La dernière abstraction est surtout l'œuvre de l'opérateur qui transforme ces projets en un véritable *« réseau de projets transactionnels »*. Si l'on imagine des initiatives individuelles, alors elles font réseau dans cette troisième étape. Dans le cadre d'une construction largement centralisée, cette dernière étape est l'accomplissement du « mythe » des effets structurants, comme compilateur de l'imaginaire du premier axiome et de la maximalisation du second axiome, et de l'apport scientifique qui peut y contribuer *(Figure 1. 2, 1c)*. Ainsi, les théoriciens des réseaux défendent souvent *« l'idée d'une aptitude intrinsèque du réseau à croître sans soucis de quelconques obstacles »* (Dupuy, 1987a). Ce réseau se rapproche de la *« structure émergente »* évoquée plus haut *(cf. 1.1.1, p. 23)* mais alors il semblerait *« vain de rechercher des lois propres d'évolution des réseaux […] c'est d'ailleurs la notion même d'effet structurant […] qu'il convient de remettre en cause »* (Offner, 1993b). Pourtant, cette hypothèse endogène peut être réfutable quand elle s'inscrit sur un territoire largement hétérogène et différencié, où le développement des réseaux est aussi lié à la transformation des usages.

L'agencement des éléments primaires, associé au dynamisme des projets de liaison sont les préalables à la naissance de la **dynamique** du réseau. La métaphore des cycles de vie des produits est alors utile à la compréhension de la genèse d'un réseau (Garrison et Levinson, 2014). Les cycles des innovations de Kondratieff (Kondratieff, 1992) se rapprochent naturellement de la croissance des réseaux techniques, où l'on peut identifier la succession des différentes innovations, ciment du développement de nouveaux réseaux *(Figure 1. 3)*. Dans ce cadre, le point de départ historique du réseau correspond à l'innovation technique sous-jacente. Ce modèle semble être adapté pour les réseaux de télécommunications, nous prendrons ici l'exemple de la disparition progressive des Minitel et des fax. Cependant, cette lecture revêt plusieurs limites. La cyclicité proposée par Kondratieff s'inspire elle-même des cycles de Schumpeter, qui décrivent la disparition d'une innovation par l'avènement d'une autre. Or, à l'inverse de ce chantre de la destruction créatrice, *« l'arrivée d'un nouveau venu n'est que rarement synonyme d'éviction pour les réseaux antérieurs »* (Offner, 1993b). Ainsi, l'apparition du chemin de fer n'a pas signé la fin du déploiement des routes. D'autres théoriciens des cycles d'innovations, comme Mensch dans son modèle de métamorphose (Mensch, 1980), réfutent donc la cyclicité proposée par Kondratieff et montrent comment la stagnation d'une innovation, parce qu'elle est arrivée à son plein potentiel, appelle une période de transition vers la recherche d'une autre innovation *(Figure 1. 3)*. Nous verrons dans le *Chapitre 3* dans quelle mesure on peut appliquer la métaphore du cycle de vie au réseau ferroviaire français.

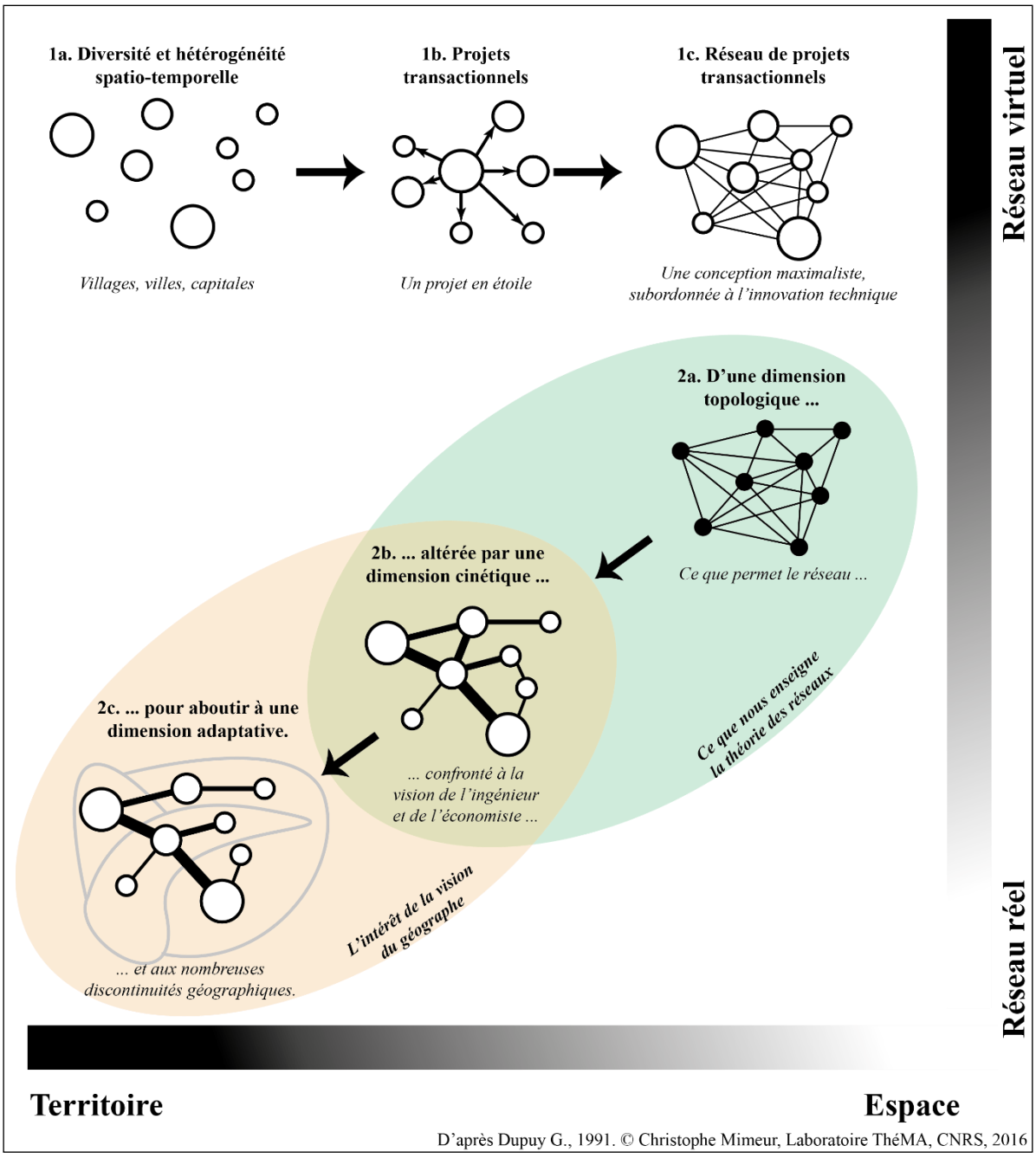


Figure 1. 2. Approche croisée du passage du réseau virtuel au réseau réel, du réseau technique au réseau territorial

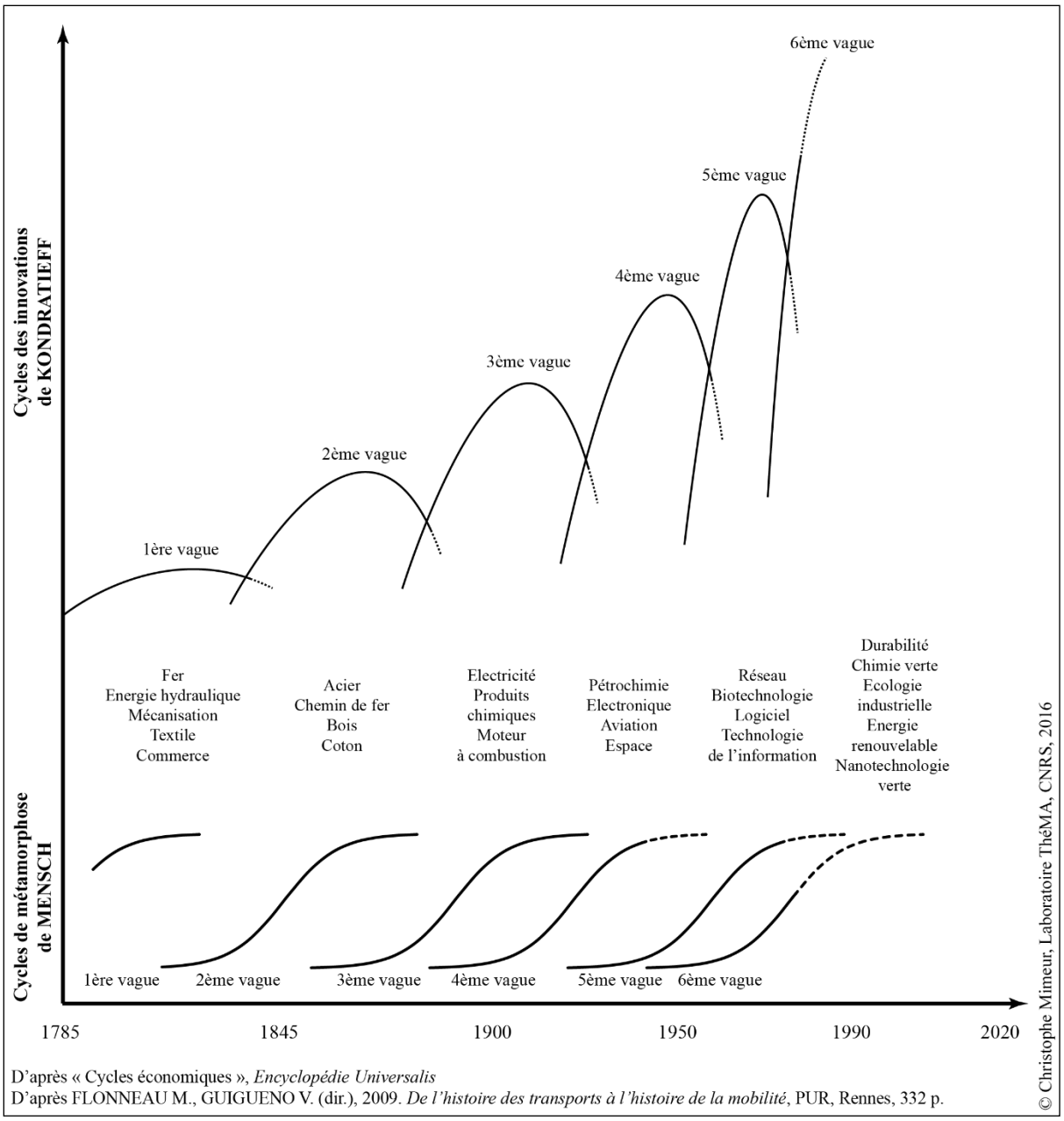


Figure 1. 3. Les cycles de réseaux : plus d'accumulation que de destruction …

Pour autant, la représentation du déploiement des principaux techniques depuis le XIXème siècle permet d'observer des effets de décalage entre l'innovation et la croissance des réseaux. Le schéma tendanciel de l'extension des réseaux techniques *(Figure 1. 4)* dessine bien des courbes logistiques, qui expliquent le transfert de la virtualité à la réalité du réseau (Dupuy, 1991), même si le déploiement de ces réseaux est inégal dans le temps, et on le suppose nécessairement dans l'espace. Si le téléphone est inventé à la fin du XIXème siècle, sa large diffusion ne sera effective qu'à partir des années 1960, alors que l'apparition de la télévision dans les années 1950 s'accompagne d'un succès immédiat. L'apparition d'un réseau serait donc la marche vers une fatale généralisation, voire une banalisation, à rapprocher de l'idéal connexionniste du réseau saint-simonien *(Figure 1. 2, 2a)*, toutefois soumis à l'aléa du bouleversement des usages.

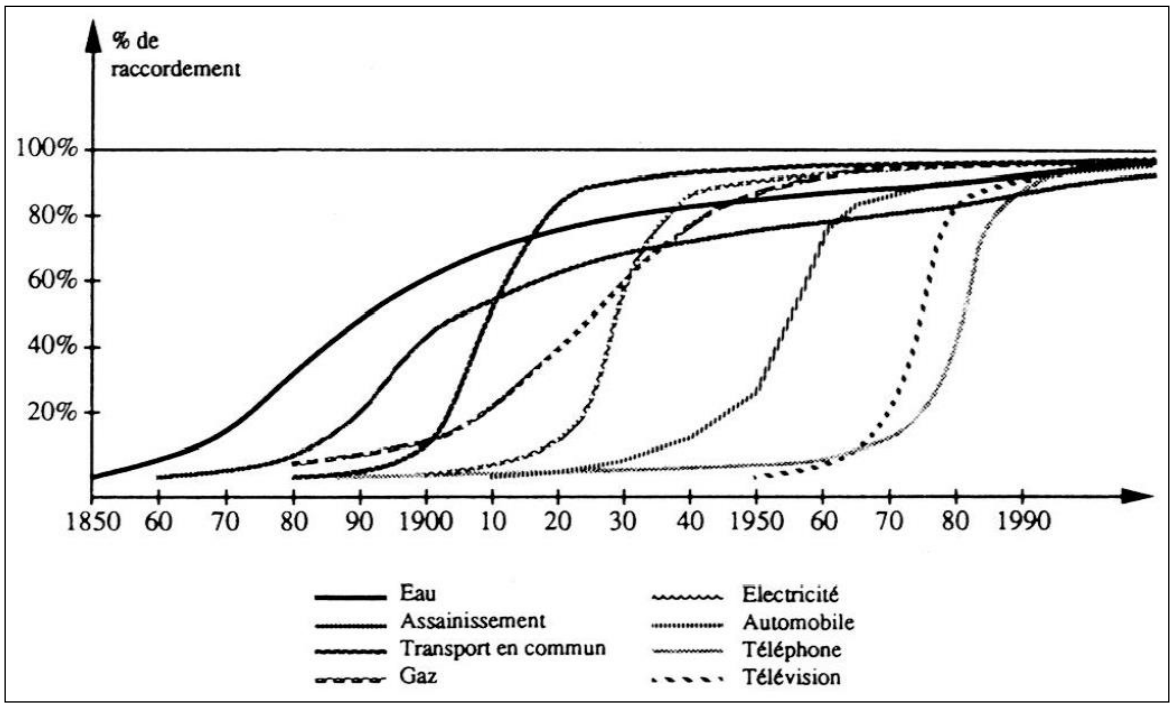


Figure 1. 4. Le schéma tendanciel des réseaux techniques (*Source :* (Dupuy, 1991))

A partir de l'approche axiomatique des réseaux de projet (Dupuy, 1987a ; Raffestin, 1980) et de l'approche des rythmes de vie des réseaux techniques (Dupuy, 1991), la réalité du réseau résulte d'une mise en cohérence sociotechnique (Curien, 2000). Ainsi la hiérarchie issue du projet transactionnel, comme traduction de la volonté organisatrice issue de la théorie saint-simonienne prend corps dans la dimension cinétique *(Figure 1. 2, 2b)*, qui est rendue possible par l'innovation technique, la hiérarchisation des vitesses (Ollivro, 2000 ; Virilio, 1995). Il en ressort deux visions différentes et complémentaires de la constitution d'une **logique** de réseau :

- La vision de l'ingénieur : on parle de *« réseau-interconnexion »*, qui pose les questions de *« l'organisation physique et l'inscription dans l'espace »* ;
- La vision de l'économiste : on parle de *« réseau-intermédiation »*, comme *« instrument d'échanges entre des producteurs et des consommateurs »*.

Ainsi, les propriétés d'un réseau sont le résultat de la confrontation entre sa structure, sa dynamique et sa logique. Nécessairement, son développement est soumis à *« différentes contraintes techniques, économiques, politiques »* et le réseau qui en résulte est *« un compromis entre réseau maximum et moyens à disposition et conditions réelles »* (Raffestin, 1980). Si la structure obéit de manière évidente au mythe de l'effet structurant, les processus dynamiques et logiques associés au développement du réseau introduisent de la différenciation, suivant que la dynamique implique un déploiement plus ou moins rapide du réseau, de manière plus ou moins différenciée selon les choix logiques qui sont actés. De notre point de vue, on pourrait toutefois déplorer que dans l'analyse de la logique d'un réseau, la vision du géographe ne soit encore pas aujourd'hui mentionnée naturellement, et pourtant, *« le réseau, à l'instar du champ agissant sur espace aréolaire, structure la « plaine » qui a trop souvent sous-tendu les modèles d'économie spatiale »* (Claval, 1981). Ainsi, parmi les frictions qui viennent contrarier l'imaginaire du projet de réseau, les composantes

géographiques *(Figure 1. 2, 2c)* sont au moins aussi importantes que celles qui sont économiques et politiques (Claval, 1981 ; Dupuy, 1993 ; Hughes, 1993). Nous verrons toutefois dans quelle mesure les géographes se sont peu à peu immergés dans cette thématique dans les prochaines sections.

Toujours est-il que l'*« instrument du bien collectif »* défendu par les saint-simoniens trouve sa traduction dans la théorie des réseaux. Il ne renvoie pas nécessairement à l'Etat mais exprime la présence de fortes externalités, dont les enjeux sont le dimensionnement et la taille (Curien, 2000). Dès lors, les lois d'évolution du réseau sont d'après nous loin de n'être que la confrontation des visions de l'ingénieur et de l'économiste. Elles font entrer ici en jeu l'ensemble des acteurs qui font système sur le territoire. La traduction du projet de réseau en réseau réel participe d'un choix à deux modalités : une logique équitable qui vise l'accès maximal de tous au réseau d'une part ; une logique efficace qui vise une construction minimale profitant au plus grand nombre d'autre part. Ces deux logiques sont encore une fois la plupart du temps imbriquées pour parvenir à un équilibre de taille.

Du point de vue de l'économiste, le bien *« réseau »* se détache de la théorie économique classique dans la mesure où il faut atteindre une taille critique à partir de laquelle les coûts fixes importants liés à construction de l'infrastructure sont compensés par l'utilisation intensive du réseau, c'est ce que les économistes appellent l' *« effet de club »* (Curien, 2000). Toujours selon eux, les importants investissements en capitaux fixes à la création du réseau peuvent n'être la plupart du temps consentis que par la puissance publique. Le bien collectif devient un *« service public »*, et dans le cadre de l'économie de réseaux, le marché s'organise le plus souvent en monopole naturel, dans lequel l'Etat joue souvent un rôle central. Mais dès lors que la dépense publique liée à l'investissement infrastructurel est croissante, le « mythe » de l'effet structurant ne suffit plus, et l'impact des infrastructures de transport sur le développement du territoire sur lequel elles s'implantent doit être sinon démontré, au moins justifié.

Cette première section a permis de montrer comment le *« mythe des effets structurants »* s'est constitué autour d'une théorie largement doctrinale, qui a su inspirer par la suite de nombreux courants de pensée sociologiques et politiques (Marcel, 1975). Par extension, ce « mythe » est selon nous largement traduisible dans les fondements de la théorie des réseaux, lorsqu'elle mobilise des concepts dans lesquels on peut aisément replacer les différentes composantes du saint-simonisme. En rapprochant réseaux techniques et réseaux territoriaux, la théorie des réseaux permet de mobiliser un arsenal de propriétés et d'indicateurs (Dupuy, 1987a) car elle permet de traduire une volonté collective, sinon politique, couplée à des enjeux techniques et scientifiques. Ainsi, en 2005 deux chercheurs de l'OCDE affirmaient que *« l'infrastructure de transport est un atout de vitalité sociale et économique, qui structure l'espace et détermine les mobilités »* (Short et Kopp, 2005), même si l'évaluation de cet effet reste un vaste champ scientifique dans le monde entier.

1.2. Qualifier les effets des infrastructures de transport

En 1944, Leland Jenks décrit le caractère décisif du chemin de fer dans le développement économique américain autour de trois conséquences distinctes (Jenks, 1944) :

- Il a un impact dans une profusion de créations d'entreprises ;
- Il stimule tous les secteurs économiques, et notamment le secteur financier ;
- Il contribue de manière directe au développement de l'économie nationale en offrant des services de transport.

S'il peut apparaître comme prophétique, alors il tempère son propos en affirmant qu'*« il apparaît qu'il n'y ait pas de techniques satisfaisantes pour donner une mesure précise de cette contribution dans l'expansion économique »* (Jenks, 1944). Alors que la théorie des réseaux la rapproche des cycles économiques, la question des liens qui les unissent se pose. Ainsi, on pose l'hypothèse qu'une large majorité d'études s'est focalisée sur une causalité linéaire de la croissance des réseaux sur le développement territorial. La définition de l'effet structurant de François Plassard s'applique ainsi quand l'infrastructure de transport engendre des effets *« qui se manifestent par des changements dans les structures économiques »* (Plassard, 1977). Ces études concernent tout autant la croissance du chemin de fer, que l'expansion des routes et autoroutes au XXème siècle, ainsi que le développement des transports urbains.

Une revue de la littérature montre comment les économistes, par la croissance de la science régionale, se sont accaparés la question des effets structurants des infrastructures de transport, alors même que la dimension spatiale tend à devenir de plus en plus présente et prégnante dans les démonstrations. Ainsi, nous montrerons comment la géographie anglo-saxonne appréhende de manière discrète la question, alors que la dimension temporelle tend à être quant à elle bien représentée. De plus, nous verrons comment les contributions scientifiques préfèrent traiter du développement économique que des mutations démographiques. En résonance avec les affirmations formulées par Leland Jenks, nous proposons de poser quelques questions simples, pour voir quels fondements théoriques sont à l'œuvre et quelles réponses sont apportées, sur le modèle des dix questions majeures de David Banister et Joseph Berechman (Banister et Berechman, 2001), dans un cadre territorial plus large. A chaque fois, nous nous efforcerons aussi de montrer les limites et critiques adressées à ces approches, et comment le géographe se situe par rapport à elles.

1.2.1. **Infrastructure et développement : un lien global et peu spatialisé**

La construction du « mythe » associée aux bienfaits du réseau et la théorie qui en a émergée montrent bien que le réseau n'obéit pas aux règles classiques de la théorie économique, où *« en termes d'utilité, un réseau dans sa globalité apporte davantage à la collectivité que la somme de ses parties »* (Curien, 2000). Il s'agit donc de *« mesurer l'utilité économique et sociale des investissements en mobilisant les méthodes et les outils du calcul économique »* (Bérion, Joignaux et Langumier, 2007). Dans ce contexte émerge la *« socio-économie des transports »* dans la seconde moitié du XXème siècle, que l'on nomme simplement *« économie des transports »* dans le monde

anglo-saxon (Commenges, 2013). Nous nous attacherons ici à nous focaliser sur les études poursuivant l'objectif de l'évaluation économique des infrastructures de transport, alors qu'elle englobe également des dispositions de quantification et de modélisation des mobilités. Le raisonnement par l'utilité nous permet de poser deux grandes questions de type macroéconomique.

a. *Une nouvelle infrastructure de transport a-t-elle un effet significatif dans le développement économique ?*

L'évaluation socio-économique des infrastructures s'est basée essentiellement sur la théorie néoclassique de la croissance selon laquelle les infrastructures de transports sont une composante à part entière du processus de production et de création de richesse (Boarnet, 1995 ; Chi, Voss et Deller, 2006). La puissance publique était dès lors liée intrinsèquement au développement des réseaux, du moins dans une acception théorique défendant la thèse des avantages comparatifs où *« les infrastructures de transport constituent un facteur de production public utilisé par les entreprises qui […] bénéficient d'externalités positives permettant d'améliorer la productivité des facteurs de production privés »* (Barro, 1990). Ainsi, lors d'une allocution hebdomadaire en 2008, Barack Obama affirme : *"we will create millions of jobs by making the largest new investment in our national infrastructure since the creation of the federal highway system in the 1950s. We'll invest your precious tax dollars in new and smarter ways, and we'll set a simple rule – use it or lose it. If a state doesn't act quickly to invest in roads and bridges in their communities they'll lose the money"*. Ici, le president américain s'appuie sur de précédentes conclusions issues du Federail Aid Highway de 1956, qui a permis la construction de 66 000 kilomètres de routes sur une décennie, ce qui constitue le plus grand plan d'investissement public consenti dans l'histoire des Etats-Unis (Garrison, 1960 ; Keeler et Ying, 1988). Analysant l'impact du Federal Aid Highway de 1956 sur le transport de marchandises, des chercheurs montrent que les bénéfices estimés viennent justifier entre un tiers et la moitié des investissements consentis par l'Etat Fédéral et les Etats (Keeler et Ying, 1988).

Aussi, un rapport anglais a montré l'impact du développement des chemins de fer en Angleterre au XIXème siècle (Eddington, 2006). Dans la seconde moitié de ce siècle, l'investissement annuel s'élève à 1,5 % du PIB anglais. Une étude de la capitalisation des investissements montre que cet effort portait à une valeur capitalisée de 30 % du PIB de l'époque. Ce même rapport montre que l'apport du chemin de fer sur l'industrie du coton et de l'acier en Angleterre a été déterminant de même que sur les gains de bien-être appréhendés par le taux de rendement social, qui s'élevait à 15 % en 1865.

Reprenant les hypothèses posées par la théorie de la croissance et notamment les hypothèses du cercle vertueux des infrastructures publiques proposé par Robert Barro (Barro, 1990), Aschauer

a suscité de nombreuses passions quand il a justifié au début des années 1990 le recul de la productivité américaine par la diminution des investissements publics d'infrastructures (Aschauer, 1989). De nombreux travaux empiriques (Gramlich, 1994) ont alors suivi pour montrer comment son appréciation du taux de rendement de l'investissement public est surestimé par de nombreux problèmes économétriques.

L'une des principales critiques adressée à cette approche est précisément le sens de la causalité. Son application globale supposerait que le capital public est endogène à la production, parce qu'il dépend du niveau de l'impôt et intrinsèquement du niveau de revenu (Lafourcade et Mayer, 2009). Le deuxième reproche adressé à la théorie de la croissance est la très possible omission d'éléments participant au développement économique d'une part, au choix des infrastructures d'autre part.

De notre point de vue de géographes, nous remarquons rapidement le caractère largement aspatial de cette approche. Aussi nous posons l'hypothèse que parmi les nombreux risques d'omission, l'absence des caractéristiques liées à l'espace géographique étudié compte parmi les plus importants, dans la mesure où les infrastructures de transport, par la métrique particulière qu'elles font émerger, *« déforment l'espace et renforcent son hétérogénéité »* (Bérion, Joignaux et Langumier, 2007). Un second type de questions prend naissance naturellement …

b. *Une nouvelle infrastructure de transport a-t-elle des effets homogènes sur un territoire ?*

Suite aux succès parfois controversés de la théorie de la croissance, la théorie des pôles de croissance semble être une réponse complexifiée aux effets structurants des infrastructures de transport (Chi, Voss et Deller, 2006). Elle a le mérite de poser l'hypothèse d'un développement inégalement réparti dans l'espace (Perroux, 1964), et participe à l'émergence croissante de la science régionale comme champ particulièrement actif dans l'étude des liens entre infrastructures et développement économique (Rietveld et Vickerman, 2004), parce que *« les réseaux de transport soulignent, entretiennent et amplifient les anisotropies du territoire, ses lignes de force* » (Lévy et Lussault, 2003). La théorie des pôles de croissance décrit deux phénomènes concomitants : des effets d'entraînement locaux ; des effets de diffusion dans un ensemble régional.

Les premiers sont bien décrits par Saskia Sassen, qui voit dans la mobilité un rôle pivot dans l'émergence de ce qu'elle qualifie de *« ville globale »* (Sassen, 1991) en montrant, que *« la mobilité ne se traduit pas seulement par un potentiel physique, de déplacement des personnes et des marchandises, ou virtuel, de diffusion d'informations. Elle suscite aussi un changement profond dans le système de redistribution des pouvoirs urbains »* (Roseau, 2012). Ainsi, certaines villes deviennent de fait hyperpuissantes, par l'accumulation de ces infrastructures, et sont l'objet d'une compétition féroce, renforçant par ailleurs les disparités entre ces espaces de grande richesse et leur

périphérie.

C'est en revanche dans sa seconde acception que la théorie des pôles de croissance participe à l'étude des infrastructures sur le développement différencié des territoires, à travers l'étude de la convergence et de la divergence du développement des territoires. Elle donne ainsi naissance à la théorie du développement régional potentiel (Nijkamp, de Graaff et Sigar, 1981 ; Rietveld, 1989 ; Vickerman, 1995) qui pose l'hypothèse que les disparités régionales puisent leurs origines dans les inégalités d'allocation de ressources, dans le contexte d'un système spatial ouvert. On retrouve alors d'un côté les facteurs de développement potentiel, les ressources dont disposent une région, et les facteurs de production. Ainsi, les infrastructures de transport, composante essentielle du potentiel de développement, sont perçues comme un instrument de politique publique visant à la réduction des inégalités entre les régions (Nijkamp, 1986 ; Ollivro, 2000). La diminution des coûts de transport, rendue possible par l'augmentation de l'accessibilité – notion discutée dans la section suivante – est au cœur du processus de réduction des inégalités.

Dans notre travail, cette approche théorique trouve une résonance particulière, parce qu'elle reprend elle aussi à son compte le cycle des produits, déjà à l'œuvre dans la théorie des réseaux, mais ici adapté au développement régional. Les propriétés des réseaux viennent participer à ce développement, dans la mesure où elles présentent un haut-degré de service public et d'indivisibilité (Nijkamp, 1986). Ainsi, une courbe logistique peut décrire une typologie de régions suivant le niveau de développement, lui-même fonction de la dotation en infrastructures. La *Figure 1. 5* permet de traduire cette typologie en isolant les effets potentiels de l'infrastructure de transport. On peut également lire cette courbe comme une trajectoire d'une région isolée sur le long-terme, dans laquelle l'investissement en infrastructure se traduit par le « décollage » du développement régional. Dans ce cas, l'approche issue de cette théorie est largement inspirée du « mythe » connexionniste, même si les catégories IV et V mettent en exergue de potentiels effets pervers, qui au-delà d'un certain niveau de développement, impliquent de potentiels problématiques effets d'étranglement. Alors que nous avons largement perçu le réseau dans ses infinies vertus, de trop nombreux réseaux feraient apparaître des désaménités (Boarnet, 1995 ; Chi, 2012), parce que les infrastructures sont des générateurs de bruit, de diverses pollutions et de potentielles congestions.

Dans son étude des régions néerlandaises dans les années 1970 et 1980, Peter Nijkamp (Nijkamp, 1986) aboutit à une typologie issue d'une analyse multidimensionnelle, décrivant un cluster important isolant les infrastructures de transport comme déterminant du développement régional.

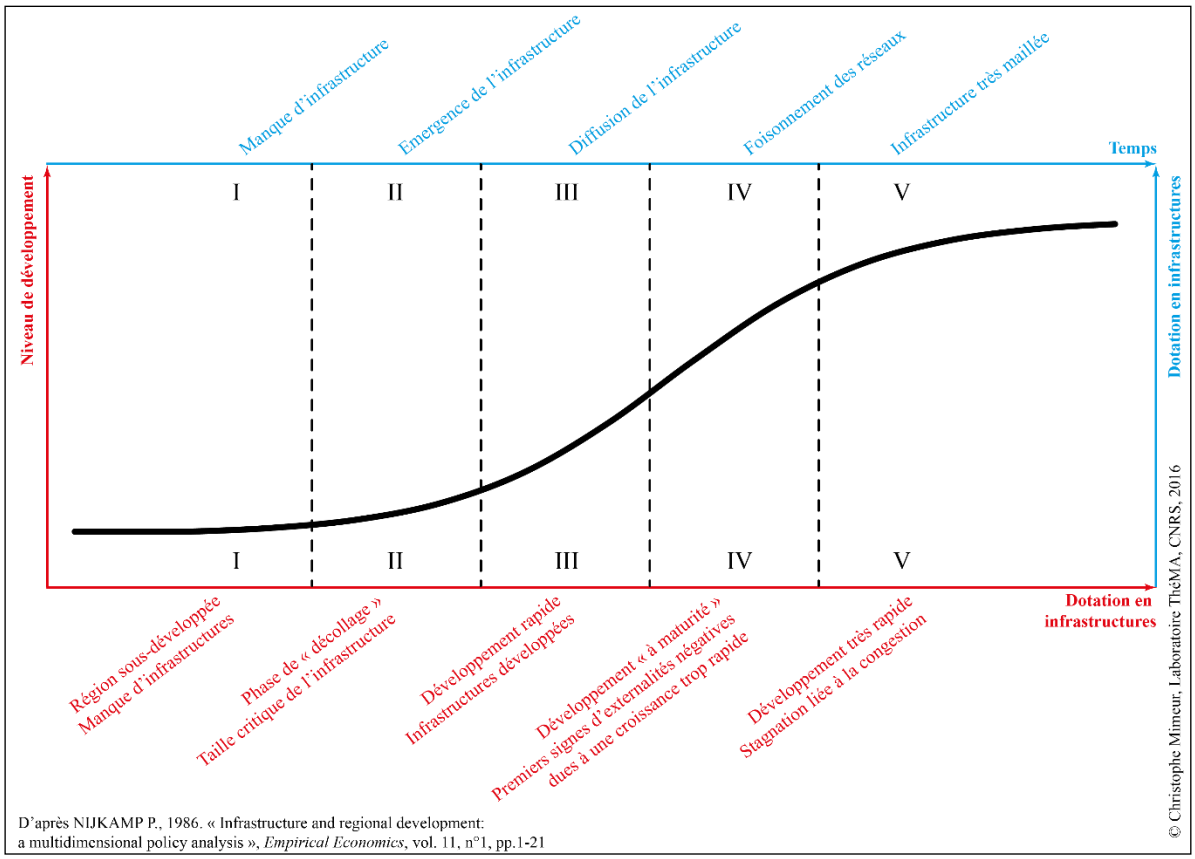


Figure 1. 5. Infrastructures de transport et développement régional
Approche typologique et temporelle

Cette approche est aussi largement présente dans les travaux traitant de la convergence européenne par les investissements en infrastructures, de transport particulièrement (Vickerman, 1995). Deux processus sont alors en jeu : compenser les régions défavorisées et contraindre les régions trop favorisées, en portant une attention particulière aux périphéries, à travers un montage financier à l'échelle du Marché Commun. Au cœur du processus de convergence européenne, les réseaux Trans-Européens ont alors une visée multiscalaire, régionale, nationale et supranationale. Pour autant, une étude portant sur les régions européennes entre 1990 et 2004 a montré que la croissance de la densité d'autoroutes, par unité de PIB, n'a pas de lien statistique évident avec la croissance économique (Crescenzi et Rodríguez-Pose, 2008), même si elle apparaît comme une condition qui doit être accompagnée d'autres politiques sur le capital humain et l'innovation, qui sont des facteurs plus marqués de l'expansion économique. Pire, l'absence d'un tel accompagnement pourrait engendrer des effets contraires. Ils viennent ainsi confirmer l'hypothèse dès 1986, selon laquelle l'effet de l'infrastructure de transport sera d'autant plus fort que les forces territoriales préexistantes le sont déjà (Nijkamp, 1986).

Les principales critiques adressées à cette approche s'attachent à montrer qu'en se concentrant sur l'aspect inter-régional, on perd de vue l'aspect intra-régional (Rietveld, 1989 ; Vickerman, 1995), qui est fortement lié à la structure régionale de la production et de sa portée par

ailleurs. Ainsi, l'investissement dans une grande infrastructure de transport à l'échelle d'une région européenne ne peut avoir d'effets tangibles que si les réseaux secondaires sont développés en complémentarité de la nouvelle infrastructure.

Ainsi, le lien entre infrastructures et développement paraît globalisé et largement aspatialisé. En effet, les théories de la croissance endogène et du développement régional potentiel font l'hypothèse d'un libre mouvement des ressources sur un espace isotrope et homogène, de plus largement agrégé (Banister et Berechman, 2004). Dès lors, des considérations économiques les plus orthodoxes et largement macroéconomiques, la socio-économie des transports se déplace vers la science régionale.

1.2.2. **Un rôle croissant de l'espace, mais sans géographes**

D'autres courants de la socio-économie des transports s'intéressent particulièrement à l'espace. La théorie des pôles de croissance fait référence aux effets d'entraînement locaux, qui sont davantage explicites dans la théorie des localisations, capable d'interpréter les distributions géographiques (Chi, Voss et Deller, 2006), dans laquelle *« les coûts de transports et les rendements d'échelles sont les deux facteurs de base »* (Quinet et Prager, 2013) des stratégies microéconomiques. Alors, ici, l'espace prend toute sa place, et le géographe semble pouvoir prendre la sienne aussi. La science régionale s'est emparée depuis ses débuts de la théorie de la localisation, en érigeant Von Thünen comme le père de cette théorie, parce qu'elle permet l'introduction de frictions associées à l'espace et une approche qui se démarque de la macroéconomie.

Pour étudier les liens entre développement économique et infrastructures de transports, *« la mesure la plus judicieuse serait que l'impact des infrastructures de transport ne réside peut-être pas dans le surcroît de richesse induit par l'accroissement du taux d'investissement public mais dans la réduction du coût des transactions »* (Charlot et Lafourcade, 2000). On peut résumer cette théorie par l'hypothèse que *« les entreprises se localisent en tenant compte de leurs coûts d'approvisionnement, de la taille des marchés qu'elles servent et de la localisation de leurs concurrents »* (Quinet et Prager, 2013). Ainsi se développe tout un nouveau pan de la théorie économique, autour du prix Nobel 2008 de l'économie, Paul Krugman, nommée la Nouvelle Economie Géographique (Krugman, 1991). La théorie de la localisation, qui paraît plus familière pour le géographe, nous autorise à poser deux grandes questions.

a. *Une nouvelle infrastructure de transport génère-t-elle des effets de proximité ?*

Rappelons ici simplement que dans son livre *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, Von Thünen propose un modèle théorique d'organisation de l'espace dans lequel le coût du transport est un déterminant de la spécialisation agricole dans un espace isotrope et homogène (Lévy et Lussault, 2003 ; von Thünen, 1826). A partir d'une ville centre entourée de sa campagne, l'objectif poursuivi par le modèle est de maximiser le profit de chaque surface agricole, fruit d'un arbitrage entre le coût foncier, le coût de transport d'une part, et les recettes de la vente d'autre part. La théorie de la spécialisation qui en découle propose un gradient centre-périphérie *(Figure 1. 6, 1.)*. Nous adaptons ici le modèle de Burgess, dont la visée est la représentation de la spécialisation résidentielle (Pumain et Saint-Julien, 2001). Les coûts de transport, simplement fonction de la **distance au centre**, tendent à spécialiser les espaces les plus proches du centre comme étant les plus rentables ; tandis que la localisation absolue tend à spécialiser les cultures les plus intensives au plus proche des villes, car leur consommation d'espace est moindre dans un contexte de pression foncière (Pumain et Saint-Julien, 2001).

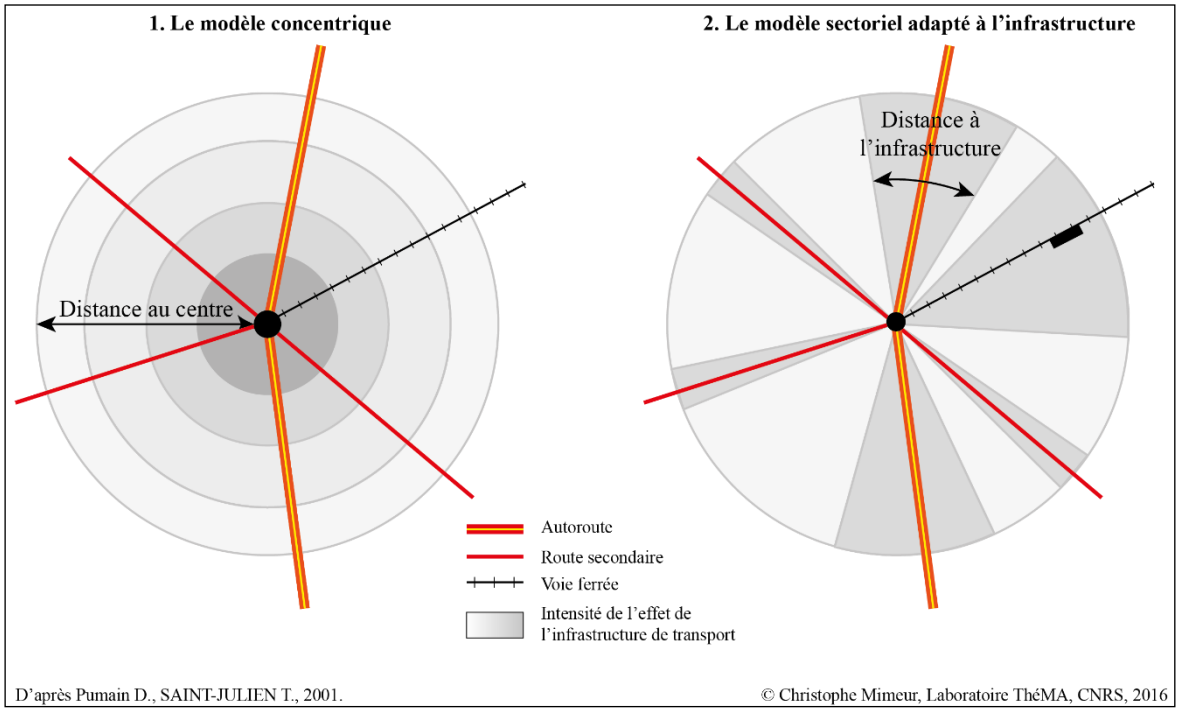


Figure 1. 6. Les modèles d'organisation de l'espace, base théorique de l'étude des effets structurants

Se dessinent ici les premiers contours du concept d'accessibilité, comme variable fondamentale des analyses coûts-bénéfices des infrastructures de transport. Si nous aurons à discuter plus loin dans ce manuscrit du concept d'accessibilité (Geurs et Van Wee, 2004 ; Hansen,

1959) dans ses acceptions mathématiques puis géographiques, la définition de l'accessibilité au sens économique du terme a été formulée dans les années 1970 et 1980 (Koenig, 1980). L'accessibilité décrit ici la *« facilité avec laquelle une activité peut être atteinte depuis une localisation donnée en utilisant un certain système de transport »* (Morris, Dumble et Wigan, 1979) et permet l'apparition de bénéfices tendant à la spécialisation et aux économies d'échelle (Gutiérrez, Condeço-Melhorado et Martín, 2010). Cette approche laisse donc une large place à la force centripète, où les gains d'accessibilité d'un lieu, c'est-à-dire l'effet du développement des infrastructures de transport, se traduisent par une augmentation de l'attractivité de ce lieu. Décrivant le développement du chemin de fer en France, l'historien et géographe Gérard Blier décrit cette force centripète en affirmant que *« grâce au chemin de fer et à son essor foudroyant, c'est depuis et vers Paris, plus que jamais, qu'à partir du milieu du XIXème siècle, les marchandises et les Hommes vont circuler »* (Blier, 1996). C'est ce que Paul Krugman démontre en affirmant que *« lorsque le coût de transport devient suffisamment faible, les forces d'agglomération sont renforcées et les activités industrielles se concentrent dans une seule région »* (Krugman, 1991).

Reprenant une simple approche concentrique, une étude empirique menée en Espagne a montré que la création d'entreprises a augmenté de 14 % entre 1980 et 1994 dans les communes situées à moins de 10 kilomètres d'une infrastructure (Holl, 2007). Cette étude pointe aussi les effets pervers de ces conclusions dans les espaces les plus enclavés. De leur côté, des chercheurs canadiens ont montré *un « impact positif clair, indépendant »* de l'accessibilité sur la croissance de l'emploi local, dans la mesure où elle concentre l'emploi dans les lieux les mieux desservis, en distinguant toutefois les secteurs d'activités (Apparicio, Séguin et Leloup, 2007).

De manière intrinsèque, cette approche se concentre aussi davantage sur l'échelle urbaine et sur la capacité de la concentration des infrastructures de transport à participer à la création d'effets d'agglomération (Combes et Lafourcade, 2012). Ces effets montrent comment la productivité d'une entreprise bénéficie, toutes choses égales par ailleurs, de la densité d'infrastructures de transport urbain mais aussi d'entreprises à proximité. Selon certains, cette force d'attraction permise par l'infrastructure de transport serait un déterminant de la localisation des investissements étrangers (Mayer et Trevien, 2012). Leur synthèse[6] cite une étude de 2009 portant sur la relocalisation des sièges de firmes américaines entre 1996 et 2001. Etudiant le potentiel impact d'une plate-forme aéroportuaire, ils montrent que le gain d'attractivité d'une aire urbaine s'accroît de 40 % en présence d'un petit aéroport, 90 % en présence d'un grand aéroport. En 2000, une autre étude a montré un impact positif et très significatif de la présence d'une grande autoroute dans la localisation de filiales étrangères aux Etats-Unis entre 1989 et 1994.

Les études portant sur l'impact des infrastructures sur la population sont plutôt rares (Adams, 1970 ; Chi, Voss et Deller, 2006 ; Lichter et Fuguitt, 1980). Pour autant, les travaux de Duranton et Turner ont montré que l'accroissement des infrastructures dans une agglomération participe à la croissance démographique, de manière en revanche indirecte, par l'évolution locale de l'emploi (Duranton et Turner, 2012). Ils montrent qu'en doublant le nombre de kilomètres d'autoroutes dans une ville, alors l'emploi croît de 15%.

[6] On trouvera les références précises dans (Mayer et Trevien, 2012). On y retrouve entre autres les auteurs Strauss-Kahn, Vives, Coughlin, Segev.

L'analyse des dynamiques spatiales urbaines constitue aujourd'hui un champ largement exploré par les économistes, mais dans lequel les géographes trouvent aussi leur place, dans l'étude des phénomènes de densification. Il s'agit d'évaluer le rôle des infrastructures de transport dans les processus décrits, même si les méthodes mises en œuvre se rapprochent souvent des dispositifs économiques et économétriques (Bérion, Joignaux et Langumier, 2007 ; Fristch, 2007).

Comme pour les autres approches, la théorie de la localisation basée sur les effets de concentration pointe quelques limites. Si l'infrastructure de transport augmente le flux de personnes et d'activités économiques par la diminution des coûts de transport et l'augmentation de l'accessibilité, l'approche des économies d'agglomérations surestime la mobilité des entreprises et des personnes (Quinet et Prager, 2013). Alors, les infrastructures de transport peuvent au contraire encourager à la dispersion et à la création de nouvelles inégalités.

b. *La nouvelle infrastructure génère-t-elle des effets de dispersion et de nouvelles inégalités ?*

Cette question vient en écho des arguments évoqués dans la première section de ce chapitre, qui postulent une mise en hiérarchie progressive des réseaux, participant à une différenciation progressive des territoires, voire à une opposition progressive.

En conservant l'échelle de l'agglomération à l'origine des effets centripètes, se pose alors rapidement la question de la participation des infrastructures de transport dans l'étalement urbain. Des chercheurs ont mené deux études différentes aux Etats-Unis et en Chine et concluent que l'étalement urbain a été favorisé autour de grandes villes par des investissements aussi bien dans de grandes radiales autoroutières que dans de larges anneaux. De même, ils imputent la baisse de la concentration spatiale des activités au maillage ferroviaire trop élevé (Baum-Snow et al., 2012). Aussi, d'autres études évaluent la modification des valeurs foncières face aux investissements en infrastructures de transport, et montrent comment la proximité d'une infrastructure de transport, même éloignée du centre, modifie le foncier (Swenson, Eathington et Otto, 1998).

Par extension, à l'inverse des phénomènes de densification que nous avons évoqués dans les lignes précédentes, les effets de dispersion associés au développement des infrastructures de transport participent de l'étude de l'évolution des formes et structures urbaines, notamment en termes d'étalement urbain (Baum-Snow, 2007). Son étude menée aux Etats-Unis montre que la présence d'une autoroute a fait diminuer sa population de 18 %, alors qu'elle aurait augmenté de 8 % si l'autoroute n'avait pas été construite.

Reprenant l'hypothèse simpliste de la théorie de la localisation adaptée du modèle de Von Thünen, le modèle de Hoyt semble ici une juste traduction du phénomène centrifuge des infrastructures de transport autour d'un centre, où la déconcentration des activités semble suivre les axes de transports. Alors que le modèle théorique semble bien établi pour décrire la ville de Chicago

des années 30, de nombreuses limites ont été posées quant à sa validité car il ne prend pas en compte l'automobile, qui obligerait aujourd'hui à sortir des limites urbaines (Pumain et Saint-Julien, 2001). Toutefois, la représentation que nous en faisons permet encore d'aujourd'hui d'interpréter la localisation des activités économiques par ce type de représentation *(Figure 1. 6, 2.)*, où la simple **proximité d'une infrastructure** déciderait d'un développement territorial.

Ici les effets structurants des infrastructures de transport ont pour principal effet la concentration des facteurs de production dans les zones les mieux desservies, mais qui ne sont pas forcément au centre. Aussi, quand les effets d'agglomérations liées aux infrastructures se conjuguent aux effets de proximité des autres entreprises, alors on aboutit à la séparation entre de *« grandes régions attractives (le centre) et des petites régions moins dynamiques (la périphérie) »* (Quinet et Prager, 2013), réifiant le gradient centre/périphérie bien connu des géographes, sans pour autant qu'ils participent intensément à sa démonstration dans le champ de la socio-économie des transports.

L'extension de ces hypothèses dans un cadre qui dépasse les limites urbaines permet d'identifier une différenciation progressive des territoires, abordée par Paul Krugman, en affirmant que *« lorsque les coûts de transport sont faibles et n'apparaissent plus comme une barrière à l'échange, la région périphérique offrant des salaires plus bas que la région spécialisée dans l'industrie, certaines entreprises sont incitées à se délocaliser pour réduire leurs coûts salariaux »* (Krugman et Venables, 1995). On peut y lire un double paradoxe. Le premier est l'effacement progressif de la variable des coûts de transport dans ce processus. En tout cas, cette vision vient atténuer l'effet de l'infrastructure de transport comme la variable d'ajustement par excellence de la localisation des activités (Bérion, Joignaux et Langumier, 2007). Dans les effets décrits par Krugman et Venables, les effets structurants de l'infrastructure sont aussi à lier à d'autres conditions. Alors que jusqu'à maintenant, l'évaluation de l'impact d'une infrastructure de transport paraissait linéaire sur le développement territorial, alors cet état de fait vient corroborer une complexification de cet effet. Alors on en arrive à nuancer l'impact de la diminution des coûts de transport dans les analyses coûts-bénéfices (Krugman, 1991).

Le second consiste à reprendre le « mythe » connexionniste abordé plus haut. Ce paradoxe pointe les limites de l'approche par la concentration des activités. Si le géographe a largement utilisé ces modélisations de l'organisation de l'espace, il pointe aussi rapidement ses limites. La principale est que l'accès au centre ne reflète qu'une partie de la réalité (Bailly, 1973) et ignore largement les interactions spatiales. On sort ici du cadre urbain pour le cadre interurbain. Ainsi, des chercheurs américains ont montré les différentiels d'impact des autoroutes dans les zones urbaines, péri-urbaines et rurales (Chi, Voss et Deller, 2006), où la baisse des coûts de transport participerait à un *« processus de diffusion des activités dans les régions périphériques »* (Krugman, 1991). Ainsi, cette différenciation est ici l'œuvre de spécialistes des statistiques spatiales et des SIG dans le cadre de l'étude d'interactions entre transport et population. Ils fondent leurs hypothèses sur la théorie des pôles de croissance, ici perçue d'un point de vue micro-économique et intra-régional, et qui semble prendre en compte une mobilité relative des facteurs de production. Les bienfaits accordés aux infrastructures de transport, ici particulièrement aux autoroutes, sont alors de trois ordres (Chi, Voss et Deller, 2006) :

- Elles favorisent l'implantation d'entreprises et la création d'emplois ;
- Elles n'obligent plus ainsi à se déplacer vers le centre ;
- Elles permettent la création de nouveaux services liés à l'infrastructure.

Les conséquences sont le possible développement de nouveaux centres, le plus souvent spécialisés. Le questionnement du polycentrisme semble périphérique des théories économiques souvent utilisées dans l'évaluation socio-économique des infrastructures de transport, alors qu'il est plus familier du géographe. Il a d'ailleurs été un instrument d'aménagement du territoire en France, comme nous le verrons plus tard dans le manuscrit. On pourrait même avancer l'idée que ces théories suscitent de nombreux doutes chez les économistes, tant ils ont participé à en montrer les limites (Boarnet et Haughwout, 2000 ; Haskins, 2002).

Ainsi, on peut déplorer que malgré une prise en compte de l'espace, appréhendé par la distance et l'accessibilité, la géographie ne soit que peu représentée dans les dispositifs exposés jusqu'à présent. La science régionale fait une part belle à l'économie, et a longtemps peiné à s'ouvrir à d'autres disciplines (Rietveld et Vickerman, 2004). Ainsi, la dimension spatiale que nous avons analysée jusqu'à présent peine à traduire des discontinuités, des rugosités et des frictions géographiques. De manière raccourcie, on peut déplorer la réduction de l'espace à un facteur de production parmi d'autres dans les modélisations économiques et économétriques.

1.2.3. **Les enseignements de la profondeur temporelle**

Si l'espace semble provoquer des réticences, alors le temps en revanche a été largement traité dans la littérature traitant des effets structurants du transport. Dans une étude de court terme, un chercheur met en garde ses lecteurs, conscient que les *« les périodes d'étude des impacts des infrastructures (1970-1975 et 1976-1980) étaient essentiellement très courtes et doivent pourtant être confrontées à des effets de long terme, particulièrement parce que le développement du capital social a généralement besoin d'une longue période de gestation. »* (Nijkamp, de Graaff et Sigar, 1981). Cette section permet de voir comment cette question, à l'aune d'une profondeur temporelle, mobilise non seulement l'économie, mais s'ouvre aussi progressivement à la géographie et à l'histoire.

a. ***Quelle place dans l'économie historique ?***

Le XIXème siècle est le terrain sur lequel l'économie historique cherche à évaluer l'impact du chemin de fer, comme *« la »* technologie à l'initiative de la Révolution Industrielle (Rostow,

1962), celle qui a permis le décollage d'une croissance auto-suffisante. Dès lors, une « évidence historique » s'est construite pour de longues décennies dont certaines réminiscences sont encore évoquées dans des rapports et contributions récentes (Eddington, 2006 ; Short et Kopp, 2005).

Et pourtant, deux contributions sont restées célèbres et sont encore presque nécessairement citées dans les travaux traitant des effets structurants des transports, en réponse à l'automatisme évoqué par Rostow. La contribution la plus célèbre, aujourd'hui encore citée allègrement[7], est celle d'Albert Fishlow en 1965 intitulée *: American Railroads and the Transformation of the Ante-bellum Economy* (Fishlow, 1965). La seconde contribution est celle de Robert Fogel : *Railroads and American Economic Growth* (Fogel, 1962). Ces deux articles auraient marqué l'économie historique, la faisant plonger dans une nouvelle vague (David, 1969).

Ils sont une réponse à un axiome d'*« indispensabilité »* décrit par Fogel. De notre point de vue, leurs recherches s'attachent davantage à démêler le mythe de l'effet structurant plutôt que de mesurer l'effet de l'infrastructure. Ainsi, ils posent le problème dans un autre sens, en appliquant la théorie de l'économie sociale : que serait-il advenu sans le chemin de fer ? Ils mesurent donc les dépenses qu'il aurait fallu consentir si les déplacements devaient se faire par la route ou sur les voies navigables. On fait donc l'hypothèse que ces alternatives participeraient à l'augmentation des coûts de transport, entraînant l'abandon de terres agricoles et une baisse de la production industrielle. Leurs conclusions sont éloquentes et de nombreuses critiques viendront montrer d'ailleurs leurs discours rhétoriques, en prise avec les débats épistémologiques qui agitaient l'économie historique (McCloskey, 1985). Robert Fogel est donc presque provocateur quand il affirme :

> *« despite its dramatically rapid and massive growth over a period of a half century, despite its eventual ubiquity in inland transportation, despite its devouring appetite for capital … the railroad did not make an overwhelming contribution to the production potential of the economy »* (Fogel, 1962).

Pour autant, sa conception de l'économie sociale a été largement critiquée parce qu'elle ne prend pas en compte la mobilité possible des facteurs de production et la demande amoindrie par l'absence de chemin de fer. Conscient de ces limites, Albert Fishlow arrive aux mêmes ordres de conclusions que Robert Fogel, mais avec des chiffres et un discours plus nuancés : *« they played a more passive role »* (Fishlow, 1965). Ainsi, Fogel évalue les pertes subies par l'absence de chemin de fer de moins de 5 %, alors que Fishlow les évalue à moins de 15 %.

Ces premiers travaux en ont initié de nombreux autres, dont l'objectif est de montrer comment la trajectoire de développement des lieux desservis s'est plus ou moins démarquée de celle des lieux ne bénéficiant pas d'une infrastructure de transport. Prenant en compte davantage de facteurs dans le calcul du bien-être (*social saving*), Donaldson arrive donc à une toute autre conclusion concernant le développement du réseau de chemins de fer en Inde entre 1850 et 1930, sous administration britannique. En confrontant un groupe bénéficiant du réseau d'un autre qui n'en a pas bénéficié bien que c'était prévu, il montre que le revenu par tête du premier groupe a augmenté

[7] Une simple recherche sur Google Scholar fait état de 577 citations. Pour comparaison, l'article de Fogel (1962) n'est cité que 114 fois, et l'article de Leland Jenks (Jenks, 1944), publié dans des périodes comparables, évoqué plus haut dans le manuscrit n'est cité que 109 fois.

de 16 %, toutes choses égales par ailleurs (Donaldson, 2010).

De son côté, reprenant l'aire d'étude de Fishlow dans le Midwest américain, Jeremy Atack aboutit à la conclusion que la desserte par le chemin de fer a contribué à l'attractivité des territoires pour l'industrie (Atack, 2013). En revanche, il confirme les conclusions de Fishlow sur l'impact modéré du chemin de fer sur la densité de population mais montre en revanche comment la présence du chemin de fer constitue un élément majeur de l'urbanisation dans le Midwest des Etats-Unis entre 1850 et 1870 (Atack et al., 2010 ; Fishlow, 1965). De son côté, Hornung montre que la croissance démographique des villes prussiennes bénéficiant d'une desserte entre 1840 et 1871 était de 1 à 2 points supérieures à la croissance démographique des autres villes. Il postule pour cela que les villes ont accueilli de plus grandes entreprises, favorisant l'emploi et les migrations vers la ville (Hornung, 2012a).

On peut en revanche déplorer que de telles études sur le cas français soient encore très limitées, surtout sur cette période (Quinet et Prager, 2013).

b. *Quelle place dans la géographie et l'histoire ?*

On peut toutefois dénombrer quelques travaux en géographie, dont la portée reste mineure, dans la mesure où il s'agit de monographies. Il s'agit là d'une vaste tradition géographique, dont elle a longtemps abusé au XXème siècle (Bavoux, 2009). Même au début des années 2000, François Plassard montre encore l'apport de telles approches pour *« comprendre comment s'articulent les changements des conditions de transport avec les transformations d'un territoire de taille relativement réduite »* (Plassard, 2003).

Claude Harmelle analyse donc l'impact du chemin de fer dans la région de Saint-Antonin (Harmelle et Elias, 1982), en concluant donc ainsi naturellement que le train a permis des déplacements qui ne l'étaient pas auparavant, mais qu'en revanche, la réduction des coûts de transport a favorisé l'importation, pénalisant dans le même temps les productions locales. Si les conclusions nous paraissent localisées, elles n'en sont pas moins très géographiques : *« quand préexistent des motifs d'échange […] des opportunités de contrôle et de capitalisation des flux […], alors des routes sont frayées […] des moyens de transport plus ou moins adaptés voient le jour »*.

Quand ils se heurtent à la question démographique, les géographes peinent à dégager de véritables conclusions quant au lien entre le chemin de fer et la démographie, comme cela a été le cas en Limagne et dans le Massif Central (Caralp-Landon, 1959 ; Chatelain, 1947). Leur approche du réseau pourrait alors sembler minimaliste, exposé surtout comme un élément de transformation du paysage (Dupuy et Offner, 2005). En 1975, Michel Chesnais propose une approche régionale et pour cela est le premier géographe français à utiliser le formalisme des graphes (Chesnais, 1975 ; Dancoisne, 1984). Pour lui, associé aux concepts de modèle de gravitation, cette nouvelle approche permet de *« situer l'étude des transports comme un élément fondamental de l'analyse régionale »* (Chesnais, 1975). Pour autant, les conclusions de ces travaux sont confrontées à la difficulté de la généralisation, tant l'échelle spatiale d'étude semble très réduite. En revanche, on peut souligner la profondeur temporelle de ces travaux, qui montrent que le *« temps est une donnée importante pour*

percevoir les recompositions » (Bérion, Joignaux et Langumier, 2007). Ainsi, les travaux d'Etienne Auphan compare les situations ferroviaires en France, en Allemagne de l'Ouest et en Grande Bretagne, en montrant comment les réseaux actuels participent de l'obsolescence des chemins de fer (Auphan, 1989), dans une perspective temporelle.

Par ailleurs, quand on se penche sur les premiers travaux historiques qui traitent de l'effet structurant des infrastructures de transport, la frontière entre l'histoire et la géographie semble poreuse. Notre but n'est pas ici de faire une revue de la littérature historique qui traite des réseaux de transport. On notera juste les tendances principales. A notre connaissance, peu d'historiens français se sont penchés sur l'impact des infrastructures de transport sur le territoire. Georges Ribeill est souvent considéré comme l'historien des chemins de fer, se consacrant surtout à l'histoire de l'entreprise, l'histoire de ses travailleurs et le poids de ses ingénieurs (Ribeill, 1984, 1993). François Caron a lui aussi retracé l'ensemble de l'histoire des chemins de fer en deux tomes, que beaucoup de ses confrères qualifient de remarquables (Caron, 1997). Il a toutefois participé activement au Groupe Réseaux, en s'employant à déceler des phases successives de développement du réseau, appuyées par des décisions politiques ou relevant de l'ingénierie (Ribeill, 1985a), sans la conscience de *« faire réseau »*. De son côté, Bernard Lepetit parle du réseau routier comme d'un *« impensable réseau »* (Lepetit, 1986), alors même qu'il est déjà largement hiérarchisé. Les aspects politiques sont donc souvent les plus prégnants des études historiques françaises. Il faut aussi noter que ces auteurs sont particulièrement ouverts à compléter leur approche historique par des explorations spatiales de leurs questions.

C'est surtout sur les scènes européennes et anglo-saxonnes que l'attrait pour les réseaux de transport va être le plus fort chez les historiens. Et les conclusions sont parfois surprenantes. Ainsi, retraçant l'historiographie portugaise, Luis Silveira souligne le regard pessimiste qu'elle a longtemps eu sur les effets du chemin de fer, justifié par les nombreux intérêts privés et la corruption qui en découle, jusqu'à l'impossibilité de l'économie portugaise de supporter la construction (Silveira et al., 2013). En opposition, la littérature récente met davantage en valeur le développement industriel et l'urbanisation, comme si un rapprochement vers l'économie historique s'opérait progressivement. A l'échelle européenne, l'étude historique des réseaux de transport paraît indispensable pour comprendre et appréhender les logiques de la construction d'un réseau de transport européen. Ils mettent ainsi en exergue la permanence des stratégies nationales dans un contexte pourtant plus large, comme s'il s'agissait d'une permanence depuis le XIXème siècle (Martí-Henneberg, 2013).

De l'autre côté de l'Atlantique, on notera l'initiative de W.G. Thomas qui analyse le chemin de fer comme un déterminant des dynamiques sociales, économiques et politiques de la nation américaine (Healey, Thomas et Lahman, 2013). Son projet *« Railroads and the making of Modern America »* vise à montrer comment le développement des réseaux a participé à la mobilité de la main d'œuvre, champ longtemps ignoré par le manque de données. L'enrichissement des archives participe à de nouveaux éclairages sur des phénomènes de grèves ou migratoires. Il est à noter qu'il est aussi à l'initiative de la reconstitution du réseau ferroviaire américain dans un système d'information géographique (SIG).

En conclusion de cette section portant sur les leçons de la profondeur temporelle,

l'évaluation des effets structurants du transport a fait la part belle aux économistes, alors même qu'ils ont nécessairement à manier les échelles et les temporalités, sans que les géographes ni même les historiens y trouvent une place importante. Les principaux enseignements viennent remettre en cause le « mythe » suprême des infrastructures de transport, et visent particulièrement à le nuancer et l'évaluer de manière objective et scientifique. Parmi les principales conclusions, on retiendra que :

- L'effet résulte de deux forces contradictoires : l'une pousse à rejoindre le centre, l'autre à la convergence entre les territoires ;
- L'effet *« résulte pour l'essentiel des premiers investissements qui établissent le réseau, et non des extensions et/ou des aménagements effectués à un stade plus avancé »* (Lafourcade et Mayer, 2009).

Alors Albert Fishlow croise ces deux affirmations en 1965 en affirmant :

"The magnitude of these changes in the transportation sector was fully matched by changes in other parts of the western economy […] a theoretical structure invoking lower transport costs can explain [the changes in the economy], […] such a logical relationship does not establish unequivocal causality. If population migrations had already extended the agricultural frontier, if urban agglomerations had already become prominent, transportation improvments may have justified and reinforced such development rather than initiated them" (Fishlow, 1965)

1.3. De la remise en cause à la controverse

L'analyse que nous venons de conduire a montré à quel point il est complexe d'isoler l'effet de l'infrastructure de transport sur le développement économique et territorial. Les principales conclusions auxquelles nous avons été confrontés tendent à affirmer que l'investissement en infrastructures de transport ne participe pas tellement à une création nette de richesse mais plutôt à une redistribution des conditions de la croissance (Boarnet, 1995), faisant des échelles un outil indispensable à son étude. L'évaluation économique doit donc s'enrichir d'autres approches, au sein de la socio-économie des transports.

Pour autant, ces constats n'ont pas freiné l'axiome d'*« indispensabilité »* (Fogel, 1962). Déjà en 1953, le constat était plutôt provocateur, quand Maurice-François Rouge déplorait que *« malgré l'importance extraordinaire qu'ils ont prise dans notre vie, l'attention n'a jamais été attirée jusqu'à présent sur les réseaux, leur croissance, leur multiplication, et sur la nature même de cette structure, qui emprisonne le globe dans des trames de plus en plus serrées et solides : réseaux de chemin de fer, de voies navigables, de routes, d'électricité, de téléphones, de distribution … Si les citadins, considérés à titre individuel, ne peuvent vivre sans les services fournis par les réseaux, on peut se demander ce que serait la vie économique, en général, sans leur existence … »* (Rouge, 1953).

Cette section vise à montrer comment les effets structurants des infrastructures de transport, théâtre de *« porosités entre univers académique et société civile »* (Offner, 2014), fait l'objet de récupération rhétorique, alors que le champ de recherche qui l'entoure tend à se complexifier.

1.3.1. **Une mainmise de l'aménagement du territoire**

L'effet structurant semble durablement *« ancré dans les milieux politiques »* (Bavoux et al., 2005). Il convient ici de rappeler quelques éléments, selon nous propice à la profusion de rapports, expertises et contributions académiques (Bazin et al., 2011). Parallèlement, dans un article devenu célèbre, en 1993, Jean-Marc Offner, économiste, dénonce la *« mystification politique »* (Offner, 1993a) de l'effet structurant, qui s'accompagne d'une rhétorique bien rodée, alors que les effets automatiques avaient déjà été pointés du doigt (Plassard, 1977). Cet état de fait est à relier au rôle important des pouvoirs publics dans l'organisation, la gestion des réseaux, où *« l'expertise sur les transports est presque totalement internalisée par l'administration »* (Commenges, 2013).

Le discours qui accompagne les investissements en infrastructures de transport a pour objectif de légitimer les décisions politiques, fortes empreintes de l'aménagement du territoire et de justifier la forte dépense publique, en vantant les évidents effets positifs. Le transport est donc une partie prenante des plans d'aménagement du territoire et d'urbanisme, succédant aux volontés géopolitiques et politiques développées dans la première section. A toutes les échelles, les transports sont de toutes les politiques publiques, porteurs d'impacts territoriaux, que les géographes envisagent par la répartition des populations et des activités sur le territoire (Bavoux et al., 2005 ; Merlin, 2007).

L'instrumentalisation des transports dans les politiques publiques est de tous les modes, dans un contexte de diversification des modes de transport après la seconde guerre mondiale, accompagnée d'une large démocratisation de l'automobile (Dupuy, 1999) et une croissance du secteur aérien. Par ailleurs, ils seront tout aussi utiles aux différentes orientations des politiques de transport, quand la crise pétrolière et la rupture du dynamisme de la croissance économique viennent contrarier le système de déplacements en place (Flonneau et Guigueno, 2009). Une lecture politique est donc nécessaire dans l'étude des choix d'aménagement, que les anglo-saxons traduisent par le terme de *« foi »* (Banister et Berechman, 2004) dans les politiques publiques de transport. La *Table 1. 1* montre comment les priorités des politiques de transport ont évolué depuis les années 1950, sur les scènes française et européenne. Ainsi, on voit comment les politiques se sont orientées entre 1950 et 1970 vers l'équité territoriale, dans le but de désenclaver les régions les plus reculées. Les orientations données entre 1970 et 1990 poussent aux renforcements du développement territorial par une meilleure dotation en infrastructures, tandis qu'on lit depuis les années 1990 une priorité donnée à la grande vitesse ferroviaire, afin de renforcer la compétitivité des territoires et la création d'un réseau de transport à l'échelle européenne.

	Europe		France	
	Priorités	***Actions***	**Priorités**	***Actions***
1950-1970	Priorité aux zones rurales	*Extensions vers les régions périphériques*	Priorité à la couverture nationale	*Prémices du réseau autoroutier*
1970-1990	Priorité au développement régional et à la cohésion régionale	*Dotation régionale en infrastructures de transport*	Priorité à la grande vitesse et au développement régional	*Développement de la grande vitesse ferroviaire*
		Renforcement du noyau européen		*Développement des systèmes ferroviaires régionaux*
				Extension du réseau autoroutier Contournements locaux
1990-	Priorité au réseau européen et à la grande vitesse	*Nouveau système de financement*	Priorité au développement européen et au renforcement de la grande vitesse	*Réseau transatlantique*
		Dotation régionale en infrastructures de transport		*Développement de l'intermodalité régionale*
		Création du marché unique		*Grande vitesse européenne*

Table 1. 1. Les priorités de la politique publique de transport
D'après BANISTER D., BERECHMAN J., 2004. *Transport investment and economic development,* Routledge, 383 p.

La permanence des discours politiques liés aux effets structurants s'est aussi trouvée renforcée par deux autres états de fait. Le premier est à imputer aux lois de décentralisation de 1982, renforcées par le Code général des collectivités territoriales (1996), qui étend les compétences des collectivités territoriales (Merlin, 2007). Le volet *« Transport »* est tout particulièrement concerné, puisque la région est chargée de l'organisation des services ferroviaires régionaux de transport, le département du transport interurbain et du transport scolaire, les communautés d'agglomération des transports urbains. Le second renforcement de la place des effets structurants des transports dans les politiques d'aménagement est la croissance des problématiques environnementales. Considéré comme un secteur contribuant de manière significative aux dérèglements, les réglementations et orientations des politiques publiques ont été nombreuses depuis la première crise pétrolière (Bavoux et al., 2005).

Ces deux aspects participent au squelette de la Loi d'Orientation des Transports Intérieurs (LOTI), du 30 décembre 1982. Dans un contexte de crise économique structurelle et de conclusions nuancées quant aux effets structurants des infrastructures de transport (Plassard, 1977), l'article 14 prévoit que *« les grandes opérations d'infrastructures réalisées avec le concours de financements publics, doivent faire l'objet d'un bilan des résultats économiques et sociaux rendu public »*[8]. De notre point de vue, les dispositions législatives sont un levier important dans l'explication de la

[8] www.legifrance.gouv.fr

permanence de l'abondante littérature traitant des effets structurants, parce qu'elles supposent des expertises, académiques ou non, avant et après l'investissement, mobilisant les théories exposées plus haut dans notre propos.

Mais, dépassant la seule sphère du politique, Jean-Marc Offner réitère ses critiques quand il dénonce les *« mythologies politico-journalistiques »* (Offner, 2014). Dès lors, *« avoir une infrastructure de transport »* semble une modalité indispensable, si ce n'est au développement économique, mais plutôt à l'image d'un territoire. Ainsi, l'infrastructure participerait à la *« revitalisation »* d'un territoire (Banister et Berechman, 2004). Les territoires qui en profitent n'hésitent pas à leur faire savoir, à grands coups d'articles et dossiers de presse. En août 2014, l'hebdomadaire *Challenges* titre *« Comment Bordeaux profite à plein de l'effet TGV »*[9] et affirme que l'effet TGV n'a jamais été démenti, en montrant la hausse des prix de l'immobilier. Ainsi, dérivé sans doute du mythe de l'association universelle, le concept de mobilité qu'on y associe est devenu un *« attribut de la modernité »* (Bavoux et al., 2005). Faisant un état de l'art des effets de la grande vitesse ferroviaire, des auteurs montrent comment la littérature aime à traiter l'amélioration du bien-être social, les changements d'image, l'amélioration de la qualité de vie et les effets psychologiques (Bazin et al., 2011). *« Vous voulez visiter Le Mans le temps d'un week-end ? […] Appelez-nous, on s'occupe de tout »*[10] : voici le slogan de l'office du tourisme du Mans, profitant de l'infrastructure à grande vitesse pour asseoir son succès touristique.

On voit comment la puissance publique s'est largement emparée, de manière permanente, du concept des effets structurants, en en faisant un moyen de légitimer son action, un moyen de faire la promotion de son territoire. Aussi, ces dernières lignes montrent comment la grande vitesse ferroviaire a largement contribué depuis trente ans à la pérennité de cette rhétorique, alors même qu'elle a contribué aussi sur la scène française à la complexification de la notion de territoire dans l'appréhension des effets des infrastructures.

1.3.2. **Acteurs et échelles dans la complexification du territoire**

Le contexte institutionnel de la décentralisation et de la loi LOTI, d'une part, et la remise en cause des effets structurants automatiques dans les années 1970 issue de recherches de plusieurs disciplines (Fogel, 1962 ; Plassard, 1977) sont selon nous deux éléments qui ont contribué à la complexification de l'approche du territoire. De plus, alors que jusqu'à maintenant, l'étude de l'effet structurant s'est surtout consacrée à l'offre de transport, la demande locale de transport semble prendre une place croissante dans les relations entre réseau et territoire (Banister et Berechman, 2004), faisant entrer d'autres disciplines dans la socio-économie des transports. Se

[9] *Challenges*, publié en ligne le 28 août 2014
[10] « L'effet TGV au Mans », *L'Expansion, L'Express*, 27 novembre 2013

dessine alors un nouveau cadre *« de plus en plus fondé sur la négociation territoriale associant les différents groupes d'acteurs intéressés par l'infrastructure et riche de questionnements »* (Bérion, Joignaux et Langumier, 2007).

Parallèlement à la valeur sociale accordée à la mobilité, des travaux se sont attachés à démontrer les effets pervers du développement des réseaux, concluant parfois jusqu'à l' *« enfer des réseaux »* (Dupuy, 2011). Loin du bien-être social, les termes de fracture et de dépendance décrivent de nouvelles inégalités, nées de l'extension des réseaux, et de la hiérarchisation déjà évoquée aux échelles urbaines et interurbaines. D'un point de vue individuel, le réseau fait apparaître des *« premium spaces »*, desservis par des réseaux performants et onéreux, mais participant à la fragmentation urbaine, qui interpellent les politiques publiques actuelles, contraintes par les objectifs de dérégulation. L'action publique est alors appelée à participer à la correction de ces inégalités, d'où une demande croissante de sa part d'analyse d'impact dans laquelle elle tient son propre rôle.

Les réflexions scientifiques se sont alors penchées sur des approches territorialisées, profitant des nombreux projets d'ouverture de nouvelles autoroutes depuis les années 1970 et de nouvelles lignes à grande vitesse ferroviaire en France depuis les années 1980. Elles se conjuguent a une *« inflexion actorielle »* (Lévy et Lussault, 2003), où les géographes tendent à davantage prendre en compte l'action individuelle et collective dans les processus de décision. Ils dotent alors le territoire d'une logique interne, qui fait système, introduisant de nouvelles terminologies, telles le potentiel, l'interaction (Frémont, 2014).

L'existence de telles études est également à rapprocher d'une demande sociale, qui s'est concrétisée parfois par la création d'observatoires des effets territoriaux[11] (Bérion, Joignaux et Langumier, 2007). Ces observatoires adoptent alors une échelle réduite, celle de l'infrastructure et une temporalité tout aussi réduite, traditionnellement *ex-ante* et *ex-post*. Les premières conclusions de tels observatoires sont que *« l'impact est fortement conditionné par les potentialités locales préexistantes, le dynamisme des acteurs locaux et les initiatives prises par les collectivités territoriales en tenant compte de la spécificité du contexte et des atouts à valoriser »* (Cornilly, Danzanvilliers et Furgaut, 1988).

De nouveaux vocables ressortent de ces approches, qui montrent une causalité directe très faible, comme nous l'avons déjà vu, et qui montrent l'importance *« d'un contexte territorial spécifique et des jeux d'acteurs particuliers »* (Bazin et al., 2011). Les termes de potentialité et d'appropriation montrent comment les acteurs locaux ont souvent un rôle déterminant, mais aussi comment les usages peuvent provoquer des transformations économiques ou géographiques. Ce n'est pas l'infrastructure en elle-même, mais son appropriation par les acteurs locaux qui peut provoquer un effet de développement. Ainsi, dès les premières lignes à grande vitesse, les gares TGV ont été perçues comme des potentiels de développement économique et des stratégies

[11] Les observatoires territoriaux ne se bornent pas aux seules études des effets d'infrastructures de transport. La DATAR en a recensé alors plus de 500 en France, sous des formes juridiques diverses et des champs d'activités nombreux.

d'anticipation ont souvent eu pour but la valorisation des espaces autour de la nouvelle gare (Facchinetti-Mannone et Bavoux, 2010) : ces stratégies prennent la forme de revitalisation de quartiers de gare dans les grandes villes et de la création de zones d'activités à proximité immédiate des gares périphériques.

L'analyse conduite par les observatoires territoriaux apporte un paradoxe entre des conclusions bien établies venant fortement nuancer l'effet direct des infrastructures, et *« l'âpreté du jeu politique, la persistance d'une croyance ou du moins d'une référence aux effets structurants »* (Bérion, Joignaux et Langumier, 2007). Pour autant, plusieurs critiques à l'égard des observatoires territoriaux ont été formalisées :

- Les approches empiriques et territorialisées doivent être complétées dans des approches formalisées capables d'introduire les ressources socio-économiques locales dans une modélisation économétrique (Boiteux, 2001) ;
- Les approches ne prennent pas assez en compte l'évolution des usages, et les bouleversements des mobilités induites par la nouvelle infrastructure de transport ;
- Les approches conditionnelles sont directement liées aux contextes spatiaux et temporels de chaque étude, rendant impossible toutes les tentatives de généralisation (Delaplace, 2014).

Plusieurs types de réponses viennent enrichir ces aspects. La première est sans aucun doute abordée par les travaux économiques et économétriques évoqués plus haut. La seconde trouve écho dans les enquêtes de déplacement qui suivent la mise en place des infrastructures et par la modélisation des trafics. Il s'agit d'ailleurs d'un champ que la géographie explore largement grâce aux avancées de la modélisation à l'échelle urbaine des interactions entre infrastructure et occupation des sols.

En revanche, la troisième ouvre des perspectives qui trouvent un écho particulier dans notre travail. Elle vient aussi corroborer notre propos sur les nombreuses monographies du début du XXème siècle. Nous nous contenterons ici d'évoquer les principales conclusions, qui viennent enrichir les discours des effets des infrastructures de transport sur le territoire.

Ces géographes quantitativistes explorent une *« vision plus abstraite de l'espace »* (Dupuy et Offner, 2005) par l'appréhension des réseaux. L'étude des transports s'ouvre à la plus large étude des réseaux, entourés des concepts qui les accompagnent. A partir d'une approche beaucoup plus formalisée et à une échelle plus large, Denise Pumain conclut en 1982 que *« la présence ou l'absence de desserte par le réseau de chemin de fer n'a pas été un facteur déterminant de l'évolution démographique »* (Pumain, 1982) au XIXème siècle, alors qu'en revanche, le chemin de fer se dessinerait *« en fonction d'une hiérarchie de tailles et de dynamismes urbains qui préexistait à son installation »*. Ainsi on peut voir apparaître le concept d'interaction spatiale, dont on a pointé le manque dans le modèle théorique de l'organisation de l'espace, quand elle conclut que *« la nouvelle infrastructure a favorisé une augmentation de la taille et de l'espacement des centres correspondant à un niveau donné de fonction »*. Les effets structurants sont alors ici à penser non pas de manière causale et linéaire mais dans les deux sens de la causalité, se rapprochant ici plutôt des arguments avancés par l'économie historique des années 1960.

Plus récemment encore, Anne Bretagnolle reprend ce concept de système hiérarchisé. Elle s'est intéressée à la mise en espacement progressive des villes, réifiant ici directement le réseau avec ses concepts de structure, dynamique et logique, évoqués dans la première section (Bretagnolle, 2003, 2009). Elle rapproche ainsi le niveau de centralité des villes des gains ou des pertes d'accessibilité dans le réseau de chemin de fer entre 1830 et 1954, en édictant le mécanisme de la coévolution. La différenciation des territoires s'accompagne ici d'une hiérarchisation : si l'infrastructure de transport s'améliore significativement entre deux agglomérations, alors elle peut impacter de manière négative d'autres agglomérations, en les laissant de côté. L'impact des réseaux de transport est illustré par l'espacement progressif des villes. Alors, on pourrait conclure *« qu'une augmentation de l'accessibilité interurbaine […] augmenterait les inégalités à l'intérieur des régions »* (Quinet et Prager, 2013), à partir d'un phénomène de *« sélection géographique »* (Janelle, 1969 ; Mackinder, 1902). Les économistes parlent alors de la construction d'une *« économie d'archipel »* (Veltz, 1996).

Les travaux évoqués ci-dessus impliquent l'introduction de concepts largement familiers du géographe. Ainsi, ils visent à montrer l'impact des infrastructures de transport dans le processus de centralité, qui ne s'inscrivent pas dans des dynamiques de court terme. Ici, le caractère cumulatif peut être qualifié de *« mode d'acquisition progressif de la complexité »* (Pumain et Saint-Julien, 2001), qui fait référence selon nous à la *« structure émergente »* évoqué dans la théorie des réseaux, et qui traduit la complexité par laquelle *« plus les relations internes d'un corps sont nombreuses, subtiles et complexes, mieux ce corps est organisé et plus il agit sur son environnement »* (Musso, 1987). Sans qualifier la manière dont ils agissent, on peut ici affirmer l'existence des effets structurants des infrastructures de transport.

Dès lors, la localisation géographique prend ici une place centrale, sur un espace pourvu de discontinuités et de frictions, dont on peut en imputer certaines au réseau lui-même. Aussi l'éclairage historique semble montrer la participation du réseau de transport à la différenciation des territoires sur le long terme. De plus, on voit que l'apport du géographe permet ici de réinsérer l'objet *« réseau »* à proprement parler, par l'étude de son évolution, posant là une pierre importante pour la suite de notre travail.

L'approche complexifiée que nous venons d'aborder vient démontrer comment la causalité entre infrastructure de transport et développement territorial n'est plus envisagée de manière unidirectionnelle et nécessaire. Les publications des quarante dernières années témoignent d'un élargissement des questionnements, alors même que l'*« effet structurant »* reste un solide argument des discours et décisions politiques. Cette persistance alimente le débat scientifique de manière permanente, dans lequel l'ensemble des protagonistes impriment sa marque, par ses méthodes et par son vocabulaire.

1.3.3. **Une controverse de permanente actualité scientifique**

En France, dans la littérature traitant des interactions entre réseau et territoire, un article est resté célèbre, puisqu'il est presque toujours cité depuis. En 1993, dans *L'Espace géographique*, Jean-Marc Offner publie *« Les effets structurants du transport : mythe politique, mystification*

scientifique » (Offner, 1993a). Structuré autour de deux grandes parties, cet article revient d'abord sur la difficulté de la définition des effets avant même leur démonstration, souvent qualifiée de douteuse, déterministe et simpliste. Il remet alors en question la causalité linéaire, toutefois déjà largement abordée dans la littérature (Boarnet, 1995 ; Nijkamp, 1986 ; Plassard, 1977). Plus incisif, il dénonce aussi le manque de rigueur méthodologique, dénonçant les procédés orthodoxes de l'économie. D'après lui, la question des relations entre transport et territoire doit être celle d'un renouveau conceptuel. Il propose alors d'emprunter au langage mathématique la notion de *« congruence »* : elle permet d'étudier deux structures qui évoluent de manière parallèle et qui s'adaptent de manière réciproque. L'objectif est alors l'identification clarifiée de ces congruences, où la piste temporelle semble la meilleure à explorer. L'analogie vers les systèmes est possible, quand ils se définissent par le *« primat de la relation sur l'objet »* (Lévy et Lussault, 2003), où il pourrait être vain de chercher des relations de cause à effet à l'intérieur de ce système.

Mais, vingt ans après la publication de cet *« article qui a fait date »* (Plassard, 2003), *L'Espace géographique* a proposé un dossier en 2014, qui témoigne des débats souvent passionnés que suscite cette question. Ce dossier, nommé *« Les controverses de l'Espace géographique »*, permet la contribution du même auteur qui réitère ses affirmations ainsi que celle d'autres chercheurs, géographes et aménageurs, venant répondre aux invectives parfois provocatrices du géographe et urbaniste. Et la façon de nommer ce dossier, de la part de la revue, n'est pas anodine. La définition donnée par le dictionnaire du Petit Robert parle d'une *« discussion argumentée et suivie d'une question »*. Quand elle est scientifique, la controverse est un concept central dans notre appréhension de l'effet structurant : elle *« se caractérise par la division persistante et publique de plusieurs membres d'une communauté scientifique, […], qui soutiennent des arguments contradictoires dans l'interprétation d'un phénomène donné »* (Raynaud, 2003). Etant donné la dimension du mythe qui s'est construit, on peut aussi légitimement se rapprocher de la définition de l'*« éristique »*, art du débat, quand on s'intéresse tout aussi au discours sur la réalité que sur la réalité elle-même (Schopenhauer, 1998). Dans cette section, nous évoquerons donc la controverse du point de vue de l'histoire des sciences, alors que notre positionnement dans le chapitre suivant s'en inspirera d'un point de vue épistémologique. Le philosophe Jean-Claude Baudet montre comment la constitution de savoirs issus d'une controverse est souvent l'imbrication entre science et technique (Baudet, 2015). Notre lecture de la construction du mythe et d'une théorie des réseaux y trouve donc une résonance particulière. Il n'est pas étonnant que ce philosophe ait publié son ouvrage aux éditions *La Boîte à Pandore*.

Ainsi reconnaît-on le mérite de Jean-Marc Offner avec son titre provocateur d'avoir ouvert une *« boîte noire »* (Topalov, 2015). Et dans cette boîte noire, on remarque rapidement comment le terme de *« réseau »* dépasse le cadre scientifique, ajoutant de la complexité à la controverse. Nous proposons ici de rendre compte de la controverse, et surtout de la boîte noire qui la renferme, par la représentation de la diversité sémantique des termes traitant des effets structurants des infrastructures de transport *(Figure 1. 7)*. L'ensemble des vocables utilisés l'ont déjà été au fil du chapitre, c'est pourquoi nous avons fait le choix de ne pas surcharger la figure par le rappel des sources, que l'on retrouvera plus haut au fil du texte.

Notre interprétation de la littérature des effets structurants propose une lecture du micro/local jusqu'au macro/global. L'axe horizontal oppose quant à lui les aspects positifs du mythe

aux aspects les plus négatifs. Une première analyse rend compte de la diversité de la nature des mots. On notera donc que les vocables se situant aux opposés de l'axe horizontal sont des qualificatifs dont l'impact est fort, témoignant de l'ancrage du mythe des effets structurants. En revanche, plus on se rapproche du centre de notre représentation, plus les vocables désignent des processus et des phénomènes.

On peut identifier le quart en haut à droite comme le *« berceau »* du mythe des effets structurants, dans lequel on va retrouver l'**association universelle**, permise par le nouveau **bien collectif**, lui-même gage de **gains de productivité**, du moins selon la montée des **discours politiques**, qui restent encore d'actualité.

Le quart en haut à gauche se trouve à l'opposé. Assez rapidement, les premiers détracteurs de cette vision mystique dénoncent cet **axiome d'indispensabilité**, parce que le **maillage** du réseau exerce un contrôle. Des opinions souvent chargées d'idéologie l'accusent de visées **impérialistes**, tandis que l'opposition entre ceux qui ont accès au réseau et ceux qui n'y ont pas accès crée une **fracture**.

Alors, les plus grands défenseurs du mythe ont tendance à opposer des arguments plus individuels, contenus dans le quart en bas à droite. Le réseau permet pour chacun l'**ouverture de l'horizon spatio-temporel**, dans lequel l'**échange** a toutes les vertus. Cette ouverture aux autres permet de véhiculer une **image du territoire**, rendue possible par le réseau, apparaissant alors comme un **révélateur** du développement territorial.

Au niveau local, les effets pervers associés aux effets structurants des infrastructures n'apparaissent qu'assez tard dans la littérature, à un moment où les réseaux se sont imposés dans de nombreuses sphères tant de la vie publique que dans la vie privée, concourant à l'**enfer des réseaux**. Ils sont perçus comme créateur de nouvelles **inégalités**, parce que certaines catégories de personnes seraient **dépendantes** de ces réseaux. Le réseau est alors porteur d'**exclusion**.

Parallèlement, l'autre groupe de vocables utilisés identifie des processus et des phénomènes, à toutes les échelles. Nous défendons ici l'idée que la controverse scientifique vise d'une part à se désolidariser des théories ou discours les plus clivants quant aux effets structurants, d'autre part à se doter de concepts et méthodes propices à l'identification et la qualification d'effets, aussi bien à l'échelle globale qu'à l'échelle locale. Ainsi, sur la scène de cette controverse scientifique, les vocables empruntent à différentes disciplines, où les mêmes phénomènes ont parfois des mots différents.

Premièrement, par l'identification de **grands choix d'aménagement**, se rapprochant de l'analyse géographique et/ou politique, certains étudient les phénomènes de **convergence**. Alors, on pourrait conclure à un **effet marginal** de l'extension des réseaux, qui est lui-même conditionné par la **complémentarité** avec d'autres réseaux.

Favorable au **développement régional** pour les économistes, l'extension des réseaux participerait plutôt pour les géographes à une **différenciation** du territoire, où la population et les activités participent à des mouvements de **concentration** ou de **diffusion**.

Bénéficier de l'accès à un réseau, pour les économistes, revient à bénéficier d'un **effet de**

« club », et ainsi, se trouver dans la **congruence** qui lie réseau et territoire. Chez les géographes, on étudie par exemple comment cet effet participe d'une **coévolution** avec la **hiérarchisation** existante, créatrice d'un **effet « tunnel »**, traduction d'une mise en espacement.

A une échelle plus locale, le réseau est le support de nombreuses **interactions**, fonction du **potentiel** ancré sur le territoire. Ces trente dernières années, les géographes ont pris acte du rôle indispensable de l'**action publique** dans la prise en compte des **opportunités** procurées par le réseau.

La diversité sémantique que nous venons de constater montre la passion avec laquelle la question est souvent traitée, sans doute parce que *« l'emprise de l'Homme semble surtout se réaliser au moyen des réseaux »* (Rouge et Dupuy, 1989). D'autres dénoncent le caractère totalitaire que prend peu à peu le réseau dans la société (Castells et Delamare, 1998). Cette emprise participe ainsi directement à l'entretien de la controverse.

Une fois que l'on a constaté la controverse, se pose la question de sa lecture. Sociologue et épistémologue français, Jean-Michel Berthelot défend la thèse que la controverse *« engendre une diversité de lectures possibles, également légitimes dès lors qu'elles sont pertinentes »* (Berthelot, 2002). Si l'on cherche à opposer les différents points de vue, il en ressort une étrange impression qu'il n'y a pas d'arguments supérieurs qui permettent de trancher la question de manière rationnelle. L'intérêt est en revanche bien plus grand lorsque l'on se penche sur la diversité et la différenciation des arguments contenus dans la boîte noire. Dès lors, il semble qu'il n'y ait plus *« de lecture plus vraie ou plus fausse, pas davantage de lecture neutre »* (Berthelot, 2002). Nous défendons ici l'idée qu'au-delà du constat d'un couple réseau/territoire qui fait système, nous devons effectuer des choix permettant de mettre l'accent sur certains phénomènes contenus dans la controverse. Parmi les termes communs, à la fois aux intensités du lien et aux sphères thématiques explorées, le questionnement sur les échelles semble faire un certain consensus et être une clé d'entrée dans cette controverse.

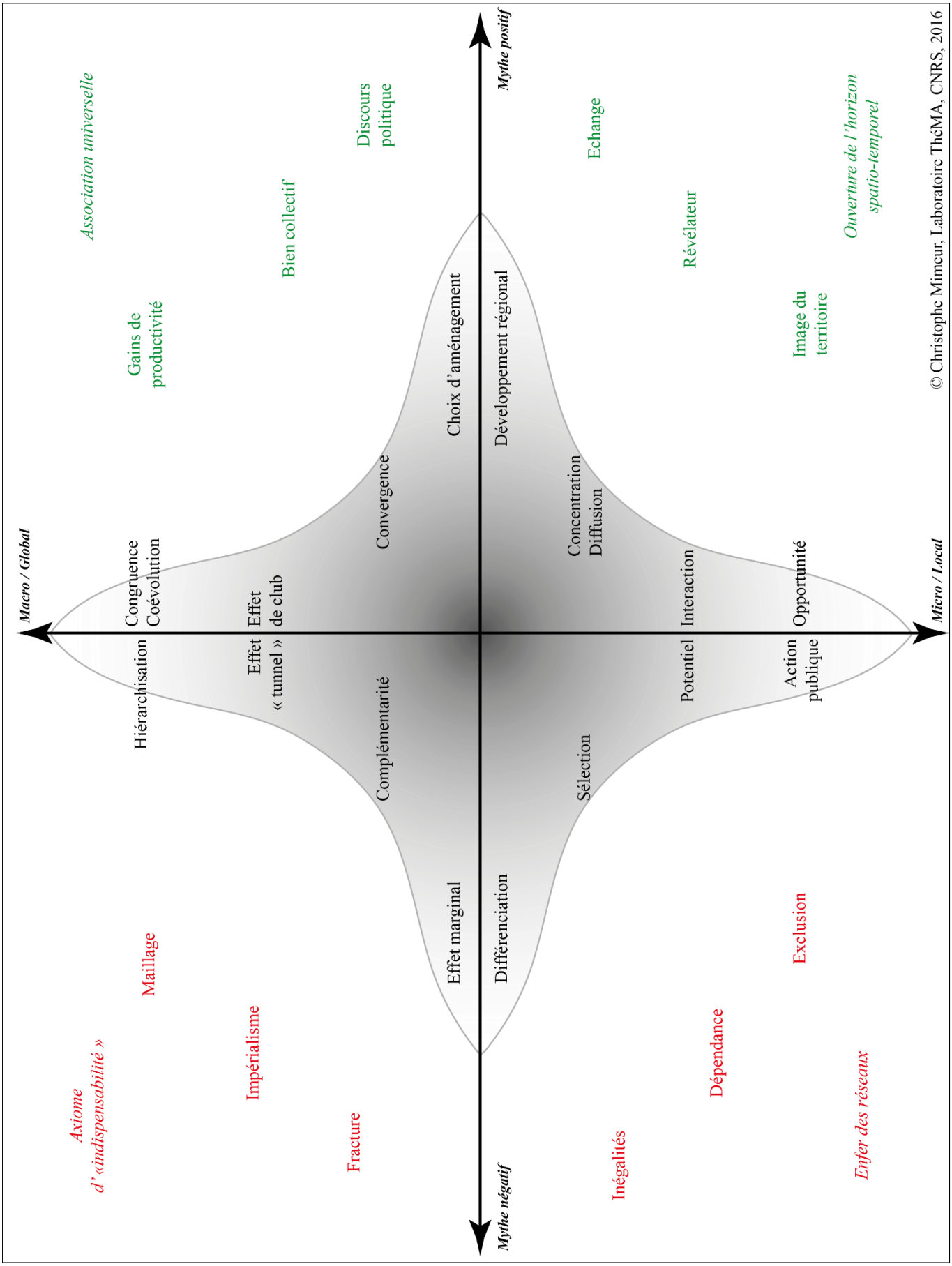


Figure 1. 7. Les effets des infrastructures de transport : du mythe à la controverse

Conclusion

Ce premier chapitre a permis de montrer les rapprochements que l'on peut faire entre d'une part la construction du mythe de l'effet structurant des infrastructures des transports et d'autre part la construction d'une théorie des réseaux, définissant des concepts et des propriétés qui leur sont propres et qui seront fondamentaux dans la construction de notre propre démarche de recherche. On a vu comment la doctrine saint-simonienne, économique, politique, spirituelle et sociale s'est vue propulsée dans les sphères centralisées de l'Etat, elles-mêmes étroitement liées au milieu industriel, dans un contexte d'innovations techniques.

Malgré une conceptualisation tardive du réseau, en tant qu'objet scientifique, les études portant sur cet effet structurant sont le marqueur d'un questionnement permanent depuis deux siècles, qui ont accompagné les innovations techniques et les avancées scientifiques. Les travaux sont multidisciplinaires même si l'approche économique est bien marquée. Leur diversité est assurée par des contributions théoriques et empiriques, macro-économiques et micro-économiques, aux échelles globale et locale, aux temporalités courte et longue. De cet état de l'art, nous retenons l'abandon d'un recours systématique aux effets structurants automatiques. Il en ressort d'une part que les causalités directes sont plutôt faibles, contrariées par des forces centrifuges et centripètes et d'autre part que le temps est un aspect important dont il faut tenir compte. La remise en cause des effets à partir des années 1970 a participé à la permanence de l'actualité scientifique de cette problématique, largement associée à l'action publique. Par-là, elle a participé aussi à l'entretien d'une controverse scientifique, dans laquelle nous avons identifié une diversité sémantique, fonction de l'intensité du lien et de la thématique étudiée.

De plus et malgré l'implication croissante des géographes dans la littérature, la géographie et l'aménagement *« ont tiré jusqu'à présent moins de profit des réflexions sur les réseaux »* (Dupuy et Offner, 2005), même si la géographie des transports semble tirer son épingle du jeu. Ces deux mêmes chercheurs identifient ce manquement par la difficulté de la géographie à faire évoluer sa vision de l'aménagement, dans laquelle le réseau semble *« s'émanciper des circonscriptions pour dessiner selon leurs propres dynamiques et leurs logiques des espaces autres ».*

Dès lors, on ne peut plus s'affranchir des jeux d'échelles. Alors que les économistes, orthodoxes ou non, reconnaissent dans l'échelle temporelle une clé dans la compréhension des effets structurants, les géographes montrent à quel point l'échelle spatiale d'analyse est substantielle aux conclusions tirées. Il nous semble qu'un positionnement scientifique appuyé par une réflexion sur les échelles spatiales et temporelles est la clé d'entrée dans cette controverse d'une éminente et permanente actualité scientifique, conscient qu'un large spectre ouvre de larges explorations en géographie.

Chapitre 2. LA CONSTRUCTION DE NOUVELLES ECHELLES DANS UNE PERSPECTIVE GEOHISTORIQUE

« Le temps met tout en lumière. »
THALES

Introduction

Réfléchir aux choix des échelles spatiales et temporelles, c'est un exercice que le professeur d'histoire-géographie propose à tous ces élèves dans la première phase de l'analyse d'un sujet. Même la composition de géographie au baccalauréat se doit d'être bornée dans un cadre spatio-temporel. Pourtant, si l'historien s'attache toujours à justifier les bornes temporelles de son étude, le géographe est plus discret, voire *« muet »* (Reynaud, 1992). Le rapprochement de l'histoire et de la géographie est ancien. Celui de l'espace et du temps l'est beaucoup moins. La définition de l'espace et du rapport à l'espace est un questionnement épistémologique chez les géographes. La diversité des échelles spatiales semble évidente, tant le géographe cherche à combiner les échelles, à les traverser. Celle des échelles temporelles le semble moins, peut-être déjà parce qu'au singulier comme au pluriel, le temps prend un *« s »*, nous abstenant de nous justifier ou même d'y réfléchir. La question semble en revanche se poser quand on cherche à identifier une ou des temporalité(s). Quand il cherche à évaluer les effets structurants d'une infrastructure de transport, le chercheur a donc du mal à *« comparer la situation avant et après l'implantation de l'infrastructure »* (Bretagnolle, 2014), parce qu'il existerait des *« effets de long terme »* (Nijkamp, 1986), nous obligeant désormais à la précision scalaire.

L'objectif de ce chapitre est la définition de notre positionnement dans le champ de la géographie des transports, avec un raisonnement géohistorique sur la longue durée, dans une approche quantifiée de l'information géographique. Pour cela, il s'agit de définir de quelle manière nous avons envisagé le temps dans notre discipline (2.1.), et plus spécifiquement à l'intérieur du champ de la géographie des transports. Nous nous insérons dans un champ émergent, issu du *« spatial turn »* des sciences sociales dans les années 1980, qui donne une place centrale à l'information géographique dans les *humanités numériques*. Elles deviennent *GeoHumanities* ou *Spatial Humanities* dans notre discipline (2.2.). Enfin, nous défendons une approche quantifiée du lien entre réseau et territoire, nous obligeant ici à la définition des échelles spatiales, au croisement de la géohistoire et de la géomatique, nous permettant de définir des cadres d'exploration de l'effet structurant (2.3.).

2.1. Une géohistoire pour des réflexions à de larges échelles

La France est un cas particulier des relations entre histoire et géographie. Elle est l'un des rares pays dans lesquels ces deux disciplines ne font qu'une dans l'enseignement secondaire. Pour autant, leurs relations seraient l'expression d'un *« couple orageux »* (Sierra, 2012). En 1694, le Dictionnaire de l'Académie Française affirmait que *« la géographie est nécessaire pour bien savoir l'histoire »*. La métaphore orageuse montre comment la géographie et l'histoire sont deux disciplines qui se sont entremêlées depuis 150 ans, jusqu'à ce qu'un champ géohistorique semble émerger en France, s'appuyant sur des concepts qu'elle s'efforce de définir. Nous nous attacherons à replacer ce particularisme dans un contexte plus global. Nous verrons enfin comment la démarche géohistorique peut trouver sa place dans la géographie des transports, où le temps prend une résonance toute particulière.

2.1.1. **Les vicissitudes des liens entre histoire et géographie, et pourtant …**

Si nous avons déjà montré la diversité sémantique qui englobe la problématique des effets structurants des infrastructures de transport, celle qui entoure les relations entre géographie et histoire n'est pas moins riche. Dans l'*Encyclopédie de géographie*, Maric-Vic Ozouf-Marignier a recensé les expressions traitant du rapport entre géographie et histoire (1992). Si les plus optimistes parlent de la géographie comme la *« moitié de l'histoire »*, voire son *« œil »*, d'autres la qualifient de *« servante de l'histoire »*, de *« science auxiliaire »*, de *« discipline nourricière »*, pire, de *« rejeton »*. Plus tard, des géographes taxeront la géographie historique comme une *« géographie du passé »* (Volvey et al., 2005), qui a tout de même perduré jusqu'au milieu du XXème siècle. Histoire géographique, géographie historique puis aujourd'hui davantage géohistoire : voilà trois vocables qui se sont succédé voire superposés, témoins d'une difficulté intrinsèque de pratiquer, sinon de nommer, des recherches qui lient histoire et géographie.

C'est la notion de frontière qui a fait de la géographie une discipline largement diffusée au XIXème siècle. A l'époque de l'installation de la IIIème République, le pouvoir y voyait un moyen de stabiliser le régime. A partir des localisations de batailles antiques, des centres ecclésiastiques, l'enseignement de la géographie est permis par Auguste Longnon, auteur d'un atlas historique depuis la période romaine. La période coloniale est toute aussi propice à la stabilisation de la discipline géographique, dans laquelle la géographie militaire trouve sa place dès le milieu du XVIIIème siècle (Boulanger et Trochet, 2005). L'étude du terrain à des fins stratégiques passe par le croissement de critères physiques et humains. Cet élan se transforme en une géographie de guerre au début du XXème siècle, où les géographes sont mobilisés pour établir de nouvelles cartes, enseigner l'aménagement de l'espace militaire, exploiter les ressources locales. Ainsi, les géographes, en se faisant connaître des services de l'Etat, s'émancipent de l'histoire, d'autant plus que Paul Vidal de la Blache a initié cette séparation à la fin du XIXème siècle.

L'école vidalienne est née en réaction à l'histoire, et donc aussi du rejet de la place donnée à la géographie par Auguste Longnon. Elle cherche à se doter d'un vocabulaire qui lui est singulier,

se rapprochant du paysage et du naturalisme. Ainsi, les monographies qui suivront décriront de manière détaillée une portion de territoire, où la description des structures ne laisse que peu de place au temps, qui n'apparaît que de manière secondaire. Si la géographie s'émancipe de l'histoire, la géographie historique tend presque fatalement à être marginalisée, et donc l'histoire tend aussi à oublier la géographie, parce que pour exister, Pierre Bourdieu montre comment un champ doit se distinguer de ceux déjà en place (Bourdieu, 1976). Seul Roger Dion persiste dans la géographie historique, qui se sert de la dimension temporelle comme d'un rempart au déterminisme naturel (Dion, 1936), où la géographie historique ne se limiterait alors qu'à une analyse des *« espaces révolus, situés dans le passé »* (Claval, 2005), dans laquelle il identifie une stabilité des paysages depuis le Moyen-Âge. La géographie tend plutôt à cette époque à une recherche de structures, sans intervention du temps, avec une démarche modélisatrice.

La situation anglo-saxonne est moins floue, peut-être parce que l'opposition entre les disciplines n'est pas aussi tenace. La revue *Journal of Historical Geography* existe depuis 1975, alors que des groupes de spécialité sont ancrés dans les rassemblements de l'Association des Géographes Américains (AAG) et l'Association des Sciences Sociales pour l'Histoire (SSHA). Ainsi, on arrive à distinguer plusieurs pratiques de la géographie historique outre-Manche et outre-Atlantique. Clifford Darby, quand il publie en 1976 *A New Historical Geography of England* (Wynne, 2003), décrit la manière dont se fait la géographie historique, nécessairement assujettie à l'analyse historique (Claval, 2005) :

- L'historien se charge d'identifier des périodes et des tranches de temps ;
- Le géographe fait en sorte de dresser un tableau de l'aire d'étude en fonction des données disponibles dans chaque tranche de temps, et s'affaire à rendre compte des transformations entre chaque tableau de périodes.

Toujours au milieu du XXème siècle, Andrew Hill Clark, canadien, s'intéresse à l'impact des colons sur la faune et la flore de l'Île du Sud en Nouvelle-Zélande. Son travail porte surtout sur les étapes d'introduction, de mise en valeur et de dynamique des espèces dans un contexte d'émergence d'associations écologiques (Clark, 1949). On peut aussi citer ici les travaux d'Alan Baker, qui s'intéresse à la vie associative dans les campagnes du Centre-Ouest de la France, en s'abstenant des découpages des historiens à partir de l'étude des comités des fêtes, des pompiers bénévoles, qu'il met en parallèle du développement ferroviaire par exemple. On voit ainsi comment, dans la géographie historique anglo-saxonne, les travaux ont évolué de l'assujettissement aux périodisations imposées par les historiens jusqu'à la mise en évidence de temporalités propres aux phénomènes qu'étudient les géographes.

En France, l'essor de temporalités propres est resté discret, tant la géographie historique française était marginalisée. On peut voir dans les monographies régionales comment l'évolution des genres de vie est envisagée de manière indépendante des dates de règnes et des régimes politiques (Claval, 2005). Bien que secondaire, le temps montre une forte inertie, faisant émerger la *longue durée*. Dans une perspective nationale, la *Géographie historique de la France*, dressée par Xavier de Planhol en 1988 distingue deux très grandes périodes : une période qui court de la Préhistoire au XIXème siècle, et une seconde, faite de différenciations et soubresauts. Si ces

géographes ont identifié des *« durées beaucoup plus amples, dans lesquelles se coulent les évolutions des techniques et des modes de vie des populations »* (Claval, 2005), ils n'ont pas reçu une forte audience dans leur discipline, alors même que des historiens comme Lucien Febvre, Marc Bloch et Fernand Braudel ont été reconnaissants de ce caractère novateur.

Naturellement, les mouvements scientifiques radicaux ont vu dans l'identification de temporalités propres un moyen de réinterpréter l'œuvre marxiste. Certains voient dans cette nouvelle interprétation un moyen d'analyser la domination des Hommes sur leur espace, dans lequel les infrastructures tendent à jouer un rôle croissant. Milton Santos, géographe brésilien, représente un courant original qui trouve une résonance dans notre travail. Ainsi, il s'attache à analyser l'évolution des formations sociales par l'environnement, largement dessiné et modifié par les techniques qui façonnent les moyens de communication (Santos, 1972). Une dernière tendance est à repérer dans la croissance des temporalités dans les approches culturelles (Ogborn, 1999). Elle est née des approches paysagères issues des monographies et des études rétrospectives. Elle s'est trouvée une autre place dans l'émergence de trajectoires individuelles (Hagerstrand, 1967), qui visent à identifier des groupes sociaux, porteurs d'une temporalité et d'une spatialité qui leur sont propres, dans des temporalités beaucoup plus courtes.

En France, le structuralisme que permet la *« Nouvelle Géographie »* depuis les années 1970 a laissé la rétrospective de côté, en renouvelant l'introduction des temporalités dans la géographie, même s'il demeure une *« inégale attention portée aux concepts d'échelle et de temporalité »* (Volvey et al., 2005). Trois grandes tendances sont identifiées pour illustrer l'introduction de la temporalité en géographie :

- La géographie culturelle, en introduisant des dynamiques possibilistes ;
- La géopolitique, parce que le recours à la longue durée est nécessaire à l'analyse des déséquilibres politiques des territoires ;
- L'analyse spatiale, issue des rapprochements avec l'archéologie, puis la géographie, notamment urbaine (Lepetit, 1984 ; Pumain et Lepetit, 1993).

Les anglo-saxons, qui ne semblent pas avoir vécu des rapports conflictuels, s'inquiètent en revanche du déclin de la géographie historique. Ils voient dans le développement de l'histoire environnementale une mainmise des études où l'espace tend à disparaître (Guelke, 1997). Pour un certain nombre d'entre eux, la « survie » de cette discipline passe par son ouverture (Ogborn, 1999). Et dans ce cadre, ils observent davantage de réticences de la part des historiens que des géographes, dans leur propension à s'emparer d'objets visant à rassembler le discours et leur matérialité, dans des temporalités définies.

Alors que la géographie anglo-saxonne n'a pas connu les mêmes vicissitudes qu'en France, ses membres identifient quelques risques d'effritement si elle ne fait pas l'effort de s'ouvrir à d'autres disciplines. La situation française est singulièrement différente : elle a connu une longue période, de la dépendance à l'histoire à la marginalisation du reste de la discipline géographique. Pour autant, le renouveau de la *« Nouvelle Géographie »* a participé à la prise en compte croissante de temporalités, parfois courtes, mais parfois longues. Dans ce dernier cas, la discipline

géohistorique s'appuie sur des fondements théoriques anciens qu'elle s'efforce de formaliser aujourd'hui, bien que la démarche reste encore timide dans l'Hexagone. En 1990, le colloque Géopoint consacré à *Histoire, temps et espace* montre à quel point la question était encore neuve. Quinze ans plus tard, Philippe Boulanger et Jean-René Trochet posent la question : *Où en est la géographie historique ?* Ce colloque montre la montée des préoccupations géohistoriques.

2.1.2. **La construction de systèmes spatiaux dans la longue durée**

La *« géohistoire »* ne semble pas avoir de traduction. Si le dictionnaire de Jacques Lévy et Michel Lussault propose les vocables *Geohistory* et *Geohistorik, l'Historical Geography* est le terme le plus souvent employé dans le monde anglo-saxon. Fernand Braudel est celui qui utilise cette notion pour la première fois dans *La Méditerranée et le monde méditerranéen à l'époque de Philippe II* (Braudel, 1949). Il divise alors le temps en trois parties *(Figure 2. 1)*. Le *« temps géographique »* est celui qui a trait aux rapports de l'homme et du milieu, avec des fluctuations dotées d'une forte inertie. A chaque échelle temporelle est alors associée une échelle géographique, proportionnée. Plus tard, dans *Civilisation matérielle, économie et capitalisme* (Braudel, 1980), il s'intéresse davantage aux impacts spatiaux dans le temps social, celui de la dynamique des relations sociales. Il utilise alors des notions largement spatiales, et raisonne en termes de localisation relative. Ici, les échelles spatiales ne sont pas nécessairement proportionnelles aux échelles temporelles. C'est à notre sens dans l'articulation de ces temporalités qu'il fait émerger la géohistoire contemporaine, quand celle de 1949 était d'inspiration vidalienne. Christian Grataloup propose alors la définition suivante :

> *« étude géographique des processus historiques, la géohistoire consiste à mobiliser les outils du géographe pour composer une explication des évènements et des périodicités, partant de l'hypothèse que la localisation des phénomènes de société est une dimension fondamentale de leur logique même »* (Lévy et Lussault, 2003).

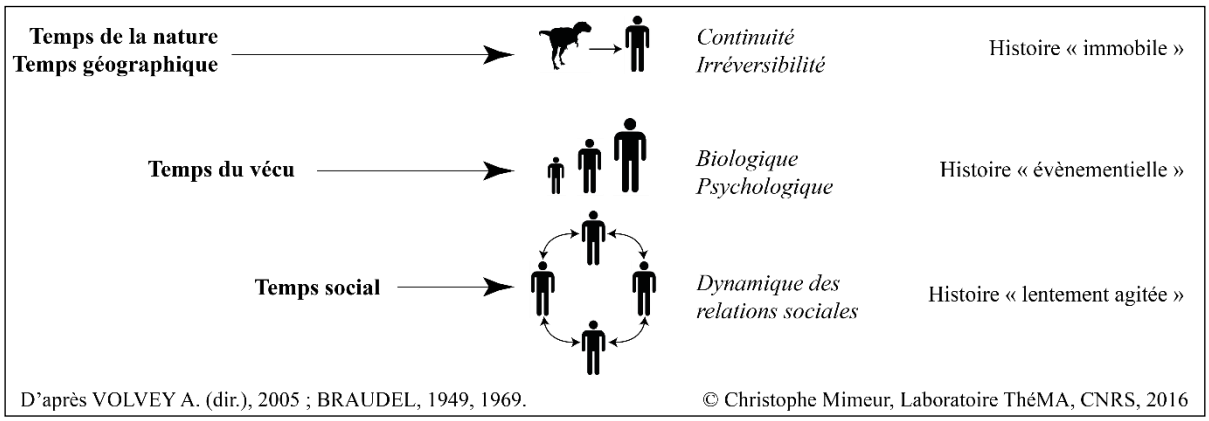


Figure 2. 1. Echelles de temporalités

La singularité de cette approche réside déjà dans la prééminence géographique par rapport à l'histoire. Elle reconnaît aussi la pluralité des temporalités, et de fait, des échelles spatiales. L'originalité contemporaine réside quant à elle dans l'utilisation des outils et dans son impératif d'explication. Cette acception l'éloigne définitivement de la géographie historique classique *(Table 2. 1)*. Elle s'est longtemps cantonnée à rendre compte de faits anciens quand la géohistoire identifie des processus, aux causalités multiples et diachroniques, sans pour autant que la période étudiée soit révolue. Ici, nous saisissons dans quelle mesure notre approche du lien réseau/territoire est constitutif d'un raisonnement géohistorique.

	Géographie historique	Géohistoire	Géohistoire du lien réseau/territoire	
Cadre temporel	Société révolue	Lien avec le présent	Permanence des réseaux	
Cadre spatial	Espace révolu	Articulation scalaire	Processus de diffusion des réseaux	
Cadre spatio-temporel	Cadre imposé par l'historien	Production et analyse des découpages	Contraction spatio-temporelle	
Localisation	Permanence de la localisation	Inertie spatiale	Relativité des localisations	
Temporalités	Causalité linéaire et synchronique	Causalité multiple et diachronique	Phénomènes de permanence et de décalage	
		Espace de liberté		**d'un positionnement géohistorique**

D'après Grataloup, 2005, 2015 ; Jacob-Rousseau, 2010 © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 2. 1. De la géographie historique à la géohistoire contemporaine … appliquée au lien réseau/territoire

Selon nous, la géohistoire amène à interroger la posture épistémologique du géographe quand il manipule les échelles à la fois spatiales et temporelles. Cette exigence paraît d'autant plus fondée que la géohistoire nous dote d'un espace de liberté, autorisant de nombreux découpages spatiaux et temporels. L'objectif ici est de cerner la démarche géohistorique en la dotant de concepts clairement définis et de méthodes directement issues de la science géographique.

Christian Grataloup propose de rapprocher le couple Espace/Temps du couple Nature/Culture, plus commun des géographes. Cette opposition est aussi classique chez les sociologues et les philosophes qui opposent *« science de la nature »* et *« science de l'esprit »* (Colliot-Thélène, 2004). Nous proposons ici de croiser les deux approches pour identifier la zone dans laquelle nous plaçons la démarche géohistorique. En combinant territoire, milieu et espace (Grataloup, 2015), nous identifions un domaine d'application de la géohistoire qui procède d'une double lecture :

- Ce que l'on regroupe dans la Nature explique la manière avec laquelle l'Homme s'accommode du milieu physique, par l'établissement de lois ;
- Ce que l'on regroupe dans la Culture relève de la *« signification culturelle des phénomènes de la vie »* (Weber, 1965), par l'identification d'*« individualités historiques »*.

Alors nous positionnons la démarche géohistorique à la charnière d'une lecture du milieu et de la territorialisation *(Figure 2. 2. a.)*. Cette démarche procède d'une identification des éléments sociaux nécessaires à la compréhension d'un phénomène, sans toutefois devoir s'efforcer à adopter une démarche heuristique dans laquelle on devrait systématiquement séparer ce qui relève du milieu de ce qui relève de la dimension humaine de l'espace. Ainsi on voit comment cette démarche tend aujourd'hui à prendre en compte un *« temps social »* braudélien dans lequel l'étude des positions relatives permet de donner des clés quant à la structuration d'un système socio-spatial.

Selon nous, rendre compte d'un système socio-spatial, au sens géohistorique, permet de dépasser la séparation Espace/Temps. En revanche, les combiner peut relever d'une démarche verticale ou d'une démarche horizontale. C'est par cette clé de lecture que nous positionnons l'objet géohistorique au croisement de ces deux approches et de la traditionnelle opposition Nature/Culture. Cette dernière devient complémentaire par l'ouverture d'un prisme qui décrit le processus de construction de connaissances géohistoriques. Il se trouve au centre d'un triptyque associant la position, le voisinage et la trajectoire individuelle. L'introduction de la dimension temporelle participe à la transformation d'un simple espace en un lieu (Cresswell, 2013), par la quête d'une identité géohistorique *(Figure 2. 2. b.)*. L'auteur parle ainsi du sens du lieu, *« sense of place »*, qui ne prend pas ici la signification de la direction, mais plutôt celle de la perception.

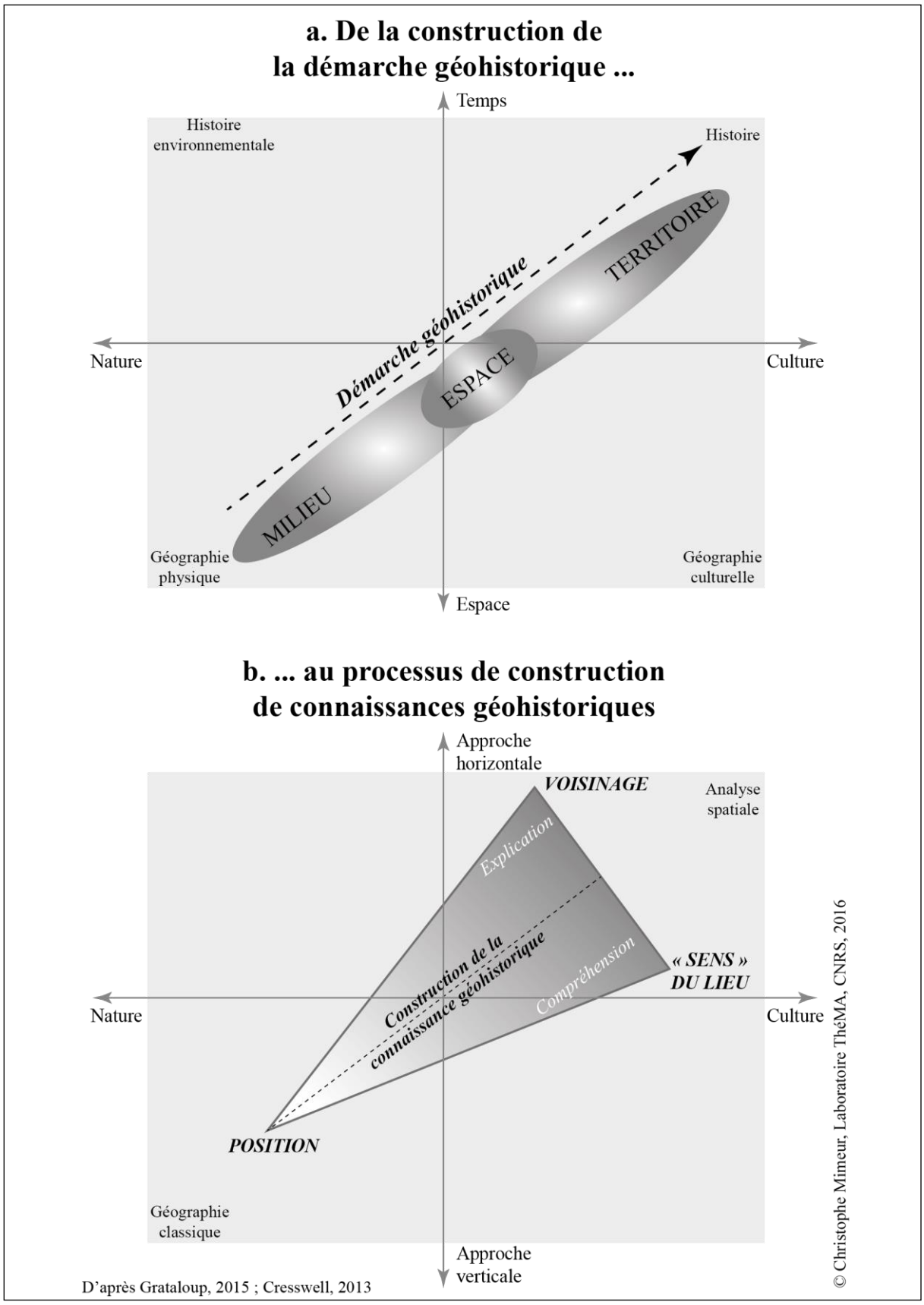


Figure 2. 2. De la construction de la démarche géohistorique au processus de construction de connaissances géohistoriques

Lisons ici notre volonté de conserver la prééminence géographique dans le processus. Nous l'expliquons aussi par la charnière à laquelle notre discipline se trouve lorsqu'il s'agit de *« comprendre »* et d'*« expliquer »* (Colliot-Thélène, 2004). Dans sa définition, Christian Grataloup propose une solution large dans laquelle les outils résident principalement dans les modèles spatiaux et la modélisation graphique (Grataloup, 1994 ; Volvey et al., 2005). L'objectif est alors de rapprocher le duo compréhension/explication dans le processus de modélisation. Nous retenons pour l'instant une définition large, comme *« énoncé théorique fondé sur la simulation schématique et formalisé, dans un langage donné, d'un objet de connaissance et d'une réalité »* (Lévy et Lussault, 2003). La dernière section de ce chapitre réutilisera ces contours dans la définition de notre posture épistémologique.

L'analyse spatiale dote ainsi les démarches géohistoriques de nouvelles ressources. Dans les années 1980, la croissance des modèles géographiques et leurs succès amènent les géographes à les utiliser pour rendre compte de phénomènes situés dans le passé. Christian Grataloup est un précurseur dans notre discipline quand il analyse la mondialisation (Grataloup, 1996) : la chorématique lui sert à rendre compte d'une très longue durée, dans laquelle il identifie des types d'évolution répondant à des principes spatio-temporels. En édictant le *« principe de Bagdad »*, il cherche à démontrer comment un lieu-carrefour procède de forces centrifuges et centripètes, tantôt productrices de nouveaux échanges, tantôt productrices de risques d'invasions. La très longue durée de la mondialisation montre aussi comment les découpages des continents participent aux interrogations actuelles de la mondialisation, comme la question de la place de la Turquie en Europe.

La démarche n'en reste pas moins timide en France. De même, les travaux mobilisant ces concepts ne se réclament pas nécessairement du champ géohistorique (Jacob-Rousseau, 2010). Et pourtant, certains travaux entrent en résonance naturelle avec le processus que nous avons mis en évidence dans la *Figure 2. 2*. Une attention particulière est portée aux systèmes de peuplement et à l'objet *« ville »*, dans lesquels les géographes font intervenir les temporalités propres à chaque ville, leur position relative et leur place dans l'environnement. L'objectif de montrer le processus dynamique qui sous-tend la formation des villes et leurs évolutions passe par des modélisations d'auto-organisation ou de construction urbaine (Frankhauser, 1997 ; Pumain, Bretagnolle et Degorge-Lavagne, 1999 ; Pumain et Lepetit, 1993 ; Sanders et al., 1997). L'appartenance à ce courant n'en est pas moins implicite dans l'appréhension du temps long par les archéologues. L'archéologie spatiale apparaît dès le milieu des années 1970 en Angleterre. Leur rapprochement est progressif vers les autres sciences sociales, dont la géographie montre une convergence des concepts, où le temps long cherche à identifier des changements, des bifurcations, des récurrences (Favory, Nuninger et Sanders, 2012). On citera aussi le dynamisme de deux groupes de recherches, Modys – *« Modélisation des dynamiques spatiales »* – et TransMonDyn – *« Modéliser les grandes transitions de l'évolution du peuplement dans l'Ancien et le Nouveau Monde »*. Aussi, il faut souligner la prise en compte des temporalités à l'intérieur de très vastes projets de recherche : l'exemple du LabEx DynamiTe montre comment la longue durée et les enseignements de type géohistorique constituent une dimension d'axes portant sur les dynamiques environnementales et sociales des risques, la construction des espaces, les dynamiques patrimoniales et les liens entre changements environnementaux et sociétés dans le passé.

De cette identification des concepts dans la géohistoire, nous retenons particulièrement la capacité de la géohistoire à une *« valorisation géoréférencée d'informations datées »* (Jacob-Rousseau, 2010), alors que nous aurons à défendre dans les sections suivantes notre positionnement dans une géographie théorique et quantitative. Nous retiendrons aussi de notre inscription dans le champ géohistorique l'espace de liberté dans lequel nous avons à définir des échelles spatiales et temporelles, en fonction des objets de notre recherche. En France, les enseignements de la longue durée sur les risques, de même que les postures naturalistes, laissent une large place à la mobilisation d'archives, telles la presse, les comptes-rendus de débats. Mais notre lecture de la pratique géohistorique en France pose la question de la place des temporalités dans le champ de la géographie des transports ainsi que celle de la mobilisation et de la mise en place d'un instrument de recherche.

2.1.3. **La longue durée, le parent pauvre de la géographie des transports ?**

Le champ de la géographie des transports semble avoir largement suivi l'évolution des approches en géographie, depuis la géographie vidalienne jusqu'à la Nouvelle géographie et l'analyse spatiale (Pumain et Saint-Julien, 2010 ; Sanders, 2001). Si la temporalité a fait aussi immersion dans ce champ de la géographie, elle n'est pas forcément de celles que Braudel a fait émerger de ses écrits. Les débuts français de la géographie des transports sont à corréler avec les études régionales que nous avons mises en valeur dans le *Chapitre 1*, d'une forte tradition vidalienne, se rattachant ainsi à une géographie rétrospective de notre point de vue. La *Géographie de la circulation* écrite par R. Capot-Rey ne fait qu'égrener par chapitre le découpage appliqué des genres de vie (Capot-Rey, 1946). En 1963, R. Clozier, dans un manuel au même titre, s'intéresse encore aux interactions entre transports et milieux naturels, tout en incorporant une approche par les modes de transport (Clozier, 1963).

En France, et même outre-Atlantique, on évalue pourtant le décollage de cette branche de la géographie à la révolution quantitative (Chang, 2004). A cette époque, les influences sont pourtant économiques : William Garrison écrit sa thèse sous la direction de William Isard, très impliqué dans la science régionale. C'est donc surtout à l'Université de Washington que la géographie des transports se développe, en lien avec les changements méthodologiques. Les géographes utilisent alors le formalisme des graphes pour modéliser les réseaux de transport, à l'intérieur desquels les modèles théoriques d'organisation de l'espace trouvent une nouvelle résonance, en parallèle des travaux de Peter Haggett (Haggett et Fréchou, 1973). L'école américaine de la géographie des transports tend alors se pencher vers une nouvelle discipline, *Transport Planning*, dans laquelle la planification des réseaux de transport passe par une approche socio-économique, utilisant les propriétés des réseaux, telles la connectivité, l'accessibilité et la centralité.

On lit ici une première incursion de la temporalité dans la géographie des transports. Torsten Hägerstrand a largement contribué à cette introduction temporelle, à double titre. Ses premiers travaux portent sur les migrations rurales en Suède au XIXème siècle : de cartes en anamorphoses à l'identification de régularités temporelles, il participe largement à une meilleure connaissance des processus de diffusion des innovations. Mais ses travaux ultérieurs occulteront durablement la longue durée des préoccupations des géographes des transports. Quand il arrive aux Etats-Unis, il développe les fondements théoriques de sa *Time Geography*, appliqués à la localisation de

ressources, à l'accessibilité des individus. Cette approche tend à prendre en compte des trajectoires individuelles dans les enchaînements et motivations de déplacements (Thévenin, 2010). De notre point de vue, ces approches ont largement influencé la dynamique de la géographie des transports. Elle laisse ainsi largement de côté les temporalités longues pour se préoccuper du présent, voire du futur.

Entre 2007 et 2009, David J. Keeling propose de revisiter les enjeux actuels posés en géographie des transports à partir d'un découplage en trois échelles *(Table 2. 2)*, du global au local (Keeling, 2007, 2009 ; Keeling et Blomley, 2007). Dans le tableau, nous étudions de quelle manière les géographes abordent ces questions par la longue durée, bien que ces démarches restent encore largement confidentielles. Les préoccupations globales tendent à lier la mondialisation et le développement de villes-mondes. De ce point de vue, l'auteur montre comment une approche géohistorique est nécessaire pour appréhender les relations entre les villes au niveau mondial, depuis les traditions centralisées de la colonisation jusqu'à la période actuelle. On notera alors les travaux précurseurs de César Ducruet, sur les flux maritimes à l'échelle mondiale (Ducruet et Notteboom, 2012) qui lie trajectoires de villes et flux maritimes. La croissance du trafic aérien pose de son côté de nouveaux enjeux environnementaux et implique le réagencement des aéroports et leur statut au sein des métropoles. L'échelle régionale, en plus d'être la passerelle entre les deux autres, est celle dans laquelle a émergé le nouveau paradigme de la mobilité (Keeling et Blomley, 2007) : la littérature traite alors des inégalités liées à la mobilité, lorsqu'elle permet de relier, de manière plus ou moins efficiente et équitable un lieu de travail et un lieu de résidence. Ces travaux trouvent une résonance particulière dans les politiques publiques de transport. L'auteur souligne alors des immersions de la géographie culturelle dans la recherche de l'héritage touristique sur la longue durée, comme élément façonnant les infrastructures de transport vers les parcs nationaux, les points d'eau. Au niveau local, les nouveaux enjeux sont liés à la place de la voiture en ville et de celle des transports urbains. Alors la géographie des transports s'attache aux comportements individuels de déplacement dans la ville, de manière à identifier les déterminants des choix modaux, privilégiant des approches désagrégées (Fox, 1995). Les problématiques associées à la congestion et au trafic de transit trouvent un écho particulier à ces échelles. Les innovations technologiques, la croissance des démarches modélisatrices ont conduit à l'élaboration de modèles d'interaction entre occupation du sol et système de transport.

La littérature francophone de la géographie des transports semble en partie influencée par ces nouveaux challenges, mais en partie seulement. Les travaux de la Commission française de géographie des transports semblent largement orientés par l'article phare de Jean-Marc Offner de 1993 (Dobruszkes et Lanneaux, 2010). Les questions relatives à la grande vitesse ferroviaire semblent une ressource inépuisable depuis les années 1980 : *« les débats sur les tracés des lignes, la localisation des gares nouvelles et l'articulation avec le réseau existant offrent constamment aux géographes des sujets de recherche, […] viennent ensuite les questions d'organisation régionale, dont l'intérêt pour les chercheurs semble pour partie […] re-dynamisé par la politique française de décentralisation »* (Dobruszkes et Lanneaux, 2010). On lit ici l'influence institutionnelle dans les orientations de cette sous-discipline, comme on l'avait souligné dans le traitement de la thématique spécifique des effets structurants des infrastructures de transport dans le *Chapitre 1*. De plus, il conclut à deux autres absences : celle d'une meilleure représentation des approches quantitatives et une autre des approches fondées sur les rapports sociaux. On ne peut toutefois pas

adhérer à l'ensemble des éléments de ces conclusions : d'abord parce que nous avons vu comment les nouveaux enjeux de la géographie des transports prennent en compte ces « absences », d'autre part parce que des chercheurs appartenant à notre laboratoire de recherche sont largement impliqués dans ces problématiques. Alors, sans conclure à une marginalisation de la géographie des transports française, on peut en revanche déplorer le déficit de dialogue au sein d'une instance regroupant les spécialistes de géographie des transports, par ailleurs également absente à l'échelon européen.

	Enjeux pour le champ de la géographie des transports	Quelle prise en compte de la longue durée ?
Echelle globale	Ville-monde Réseau-monde	Flux et centralités
Echelle régionale	Grande vitesse ferroviaire Inégalités liées à la mobilité Logistique interurbaine	Enjeux patrimoniaux Marque des réseaux
Echelle locale	Congestion Enjeux environnementaux Répartition modale Logistique urbaine	Risques environnementaux Tourisme rural

D'après Keeling, 2007, 2009 ; Keeling et Blombey, 2007

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 2. 2. Les nouveaux enjeux de la géographie des transports, peu confrontés à la longue durée

De plus, Emile Merenne propose six classements souvent opérés en géographie des transports (Mérenne, 2014) :

- Par mode : route, rail, voie d'eau, aérien, maritime, multimodaux ;
- Par nature : personnes, marchandises ;
- Par propriétaire : public, privé ;
- Par utilisateur : individuel, collectif ;
- Par type de site utilisé : site propre, site banal ;
- Par milieu : urbain, interurbain.

Si on peut faire entrer en résonance notre travail de recherche dans quelques-unes de ces catégories, la dimension temporelle est en revanche largement absente, qu'elle soit d'ailleurs courte ou plus longue. Pour autant, les transports ne sont pas absents de réflexion à de longues temporalités : on retiendra ici le dynamisme de T²M – *Association for the History of Transport, Traffic and Mobility* – dont le colloque annuel de 2016 traitera des espaces de flux et des frictions. En France, on portera une attention particulière à l'association pour l'histoire des chemins de fer français – Rails & Histoire – dont le comité scientifique accueille certes des historiens, mais aussi des géographes des transports comme Francis-Beaucire, Xavier Desjardins, Pierre Zembri.

A ce stade, nous retenons comment les temporalités longues sont timidement engagées dans la géographie des transports, plus souvent concernée par les innovations technologiques, au rang desquelles le Système d'Information Géographique (SIG) tient bonne place, au service de

problématiques immédiates. Parallèlement, conscient de l'atout du géographe à une *« valorisation géoréférencée d'informations datées »* (Jacob-Rousseau, 2010), nous nous engageons dans une démarche croisant l'information géographique, dans une temporalité, qui la fait se transformer en une information géohistorique.

2.2. Le *« spatial turn »*, entre questionnement et instrument

Si les effets structurants des infrastructures de transport peuvent être assimilés à la problématique de la poule et de l'œuf, alors le *« tournant spatial »* peut se prêter à la même analogie. Certains reconnaissent dans l'évolution de la géographie humaine depuis vingt ans la renaissance conceptuelle et méthodologique qui a permis de faire de la question spatiale une des plus dynamiques des sciences sociales (Warf et Arias, 2008). Pour autant, on pourrait tout aussi bien lire l'évolution des sciences sociales comme celle qui a créé le besoin de spatialité. Cette section vise à montrer comment le *« spatial turn »*, initié dans les années 1980, a fait la part belle à l'espace, et de plus en plus spécifiquement à l'information géographique. Les *Digital Humanities* sont une branche, une discipline, une scène de dialogue qui peinent encore à se définir, alors que la science géographique s'y est engouffrée en adaptant les Systèmes d'Information Géographique (SIG) à des problématiques temporelles. Cet outil entre alors en adéquation avec le questionnement que nous souhaitons mener dans ce travail de recherche : si les humanités peuvent encore éprouver des difficultés à dialoguer avec les sciences informatiques, alors la géographie est de celles qui se sont confrontées très tôt aux exigences d'adaptation de langage, à l'heure du renouveau quantitatif de la géographie.

2.2.1. **Les humanités à l'heure du *« digital turn »* et du *« spatial turn »***

Le *« digital turn »* des sciences humaines et sociales est difficile à saisir (Berry, 2011), tant le terme même d'humanités numériques est un oxymore (Dacos et Mounier, 2014) : livre contre écran, interprétation contre calcul, etc. Les essais de définition sont multiples (Gold, 2012 ; Schreibman, Siemens et Unsworth, 2008) : il faut lire dans l'émergence de ce mouvement les évolutions qu'entretiennent les sciences humaines et sociales et les sciences informatiques (Le Deuff, 2014). Les prémices de ce mouvement sont à trouver dans l'initiative de Roberto Busa à partir de 1946. Pendant vingt ans, il a participé à l'analyse lexicale complète des œuvres de Thomas d'Aquin, dans son *Index Thomisticus*. L'ordinateur fut alors un puissant allié dans le repérage de récurrences linguistiques, instaurant les premiers principes de la linguistique computationnelle. Traditionnellement, les rapprochements entre littérature et informatique sont les plus dynamiques dans le champ des humanités numériques (Dacos et Mounier, 2014). Dans les années 1980, le développement du mouvement *Humanities Computing* permet de sortir de perspectives purement disciplinaires pour envisager des protocoles transposables à plusieurs disciplines. Enfin, l'explosion

du web depuis le milieu des années 1990 participe de l'émergence des *Digital Humanities*, dans la mesure où le souci de la diffusion et de la communication des travaux devient un enjeu majeur.

De notre point de vue, les humanités numériques se sont considérablement développées depuis vingt ans, bien qu'elles soient peu représentées en France, s'engouffrant dans la nouvelle scène permise par Internet, sans toutefois se structurer sur le terrain scientifique. Nous identifions cette difficulté à celle même de se concevoir comme (i) un instrument de recherche, (ii) un outil de communication ou (iii) un objet de recherche (Dacos et Mounier, 2014). Ainsi, en France, ceux qui se réclament de cette branche, de ce champ, se sont fédérés autour d'une non-conférence sur les humanités numériques, *THATCamp Paris*[12]. En 2010, ils sont à l'origine du *Manifeste des Digital Humanities*, qui se fonde sur une approche pragmatique du tournant numérique, autour de principes forts, assimilables aux trois façons de concevoir les humanités numériques :

- une large diffusion des méthodes et du code (i) ;
- une pratique très ouverte, proposant le libre accès aux publications (ii) ;
- l'intégration aux dispositifs de formation (iii) ;
- la construction active de cyberstructures (i, ii).

A l'échelle mondiale, on recense de multiples essais de fédération de chercheurs, qui ont parfois échoués. L'Alliance des organisations pour les humanités numériques (ADHO)[13] recense la production scientifique à l'intérieur de ce champ. Elle fait état de cinq revues à comité de lecture. Pour autant, il faut souligner que la croissance des revues atteint ses limites quand certaines s'essoufflent au bout de quelques années[14]. Une autre organisation fédératrice rassemble les centres dédiés aux humanités numériques à travers le monde. Les membres fondateurs du dhCenterNet[15] sont composés à 50 % de centres canadiens, américains et britanniques. Les autres sont allemands, taïwanais, australiens, néo-zélandais et français. Il s'agit en France du Centre pour l'édition électronique verte[16], plate-forme de ressources électroniques, dont les activités ne se limitent toutefois qu'à la publication de livres, articles, notes et évènements. Pour autant, le rapport publié en 2014 par l'Institut Français encourage davantage au développement de centres des humanités numériques dans les universités, ainsi que le rapprochement des initiatives françaises du dynamisme international autour des humanités numériques. S'il existe une association francophone des humanités numériques – *Humanistica* –, on remarque que ses membres institutionnels, dont le nombre est limité, sont pour la plupart des laboratoires d'histoires, de linguistique, de communication. On compte parmi eux de nombreuses Maisons des Sciences de l'Homme. Enfin, il faut mentionner la cyberinfrastructure française *Huma-Num* : l'objectif est de mettre à disposition des chercheurs en sciences humaines et sociales un dispositif technologique, à toutes les étapes nécessaires à l'utilisation de données numériques. Il s'agit d'imaginer des dispositifs génériques à

[12] The Humanities and Technology Camp –THATCamp : http://tcp.hypotheses.org. Le manifeste est le résultat des débats du premier rassemblement THATCamp Paris les 18 et 19 mai 2010.
[13] http://adho.org
[14] La revue *Humanist, Discussion Group*, créée en 1987, a subi cette croissance des revues, et a disparu en 2008. La revue *Computers in the Humanities Working Papers* a disparu en 2009. Le dernier numéro de *Text Technology* remonte à 2007. Certaines autres revues ont fusionné, notamment celles qui proposaient la publication des actes des rassemblements.
[15] http://www.dhcenternet.org
[16] http://cleo.openedition.org

l'ensemble des sciences humaines et sociales. Cette infrastructure fait partie d'un réseau européen plus large DARIAH-EU, *Digital Research Infrastructure for the Arts and Humanities*.

Nous retenons de l'émergence des *Digital Humanities* une nouvelle scène de dialogues entre disciplines, semblant toutefois souffrir de structures ouvertes à l'ensemble des sciences humaines et sociales, en France en tout cas. On peut alors se demander s'il s'agit réellement d'un champ ou d'un *« accompagnement technologique, informatique et documentaire »* (Le Deuff, 2014). Dans notre travail de recherche, l'inscription dans le champ des humanités numériques se justifie par notre démarche géohistorique, qui propose de revisiter les effets structurants des infrastructures de transport à la lumière de la longue durée, en mobilisant des outils, à la fois objet et instrument de recherche. Notre recours à l'archive entre en résonance avec le *« digital turn »* des sciences humaines et sociales, qui ouvre un panorama croissant de données nécessaires à notre travail. Dans cette optique, nous nous positionnons en faveur de l'utilisation de logiciels libres et ouverts dans la conception de notre instrument de recherche et la construction de nos outils de recherche. Aussi, la dimension de la diffusion des données, dans une perspective géohistorique et patrimoniale, suscite dans nos travaux un grand intérêt, qui nous rapproche sensiblement des orientations pragmatiques des scientifiques se réclamant des humanités numériques. Dès lors, notre positionnement géohistorique trouve ici encore une place légitime quand l'information géographique se situe au cœur de notre démarche.

L'ADHO est une structure séparée en cinq groupes d'intérêt. Parmi eux, figurent en bonne place les *GeoHumanities.* Leurs préoccupations visent à adapter, développer et favoriser le développement d'outils nécessaires à la prise en compte de perspectives spatiales, spatio-temporelles et *« placiales » (en anglais)*[17], dans le cadre élargi des sciences humaines et sociales. Ainsi, nous identifions, à côté du *« digital turn »* des sciences humaines et sociales, un *« spatial turn »* (Finnegan, 2007 ; Warf et Arias, 2008), qui lui est quasiment synchrone. Il se concrétise par la présence de sessions consacrées aux études spatialisées dans de nombreuses conférences. Constatant la relative absence de l'espace dans les théories historicistes et structuralistes du milieu du XXème siècle, les milieux marxistes voient dans l'espace une clé de lecture de la société. Le regain d'intérêt pour cette dimension, que Kant élevait déjà à celle du temps, réside aussi selon Manuel Castells dans le passage à une société de l'information (Castells, 1998), dans laquelle il place le postmodernisme dans un contexte d'*« espaces de flux »*. Le réseau participe alors à l'homogénéisation de l'espace à partir d'une convergence de l'espace-temps. Certains y lisent la fin de la géographie, alors que d'autres voient au contraire un renouveau, dans lequel de nouveaux questionnements d'ordre local émergent, autour du sens des lieux, de leur identité. Tim Cresswell décrit alors le processus du tournant spatial par la transformation de *Space* à *Place* (Cresswell, 2013). Ainsi, les historiens identifient cette prise de conscience dans le contexte de réduction des distances permises par les innovations techniques (Janelle, 1969) : *« l'expérience de la simultanéité a redonné un sens à la distance »* (Bodenhamer, Corrigan et Harris, 2015). L'émergence de la vitesse participe de la *« dromosphère »* (Virilio, 1995) dans laquelle l'ubiquité est peu à peu remise en cause, et où la quête d'une nouvelle identité associée au lieu est croissante, au cœur des courants radicaux des sciences humaines et sociales. On comprendra comment cette vision fait écho dans

[17] http://www.geohumanities.org

notre question de recherche.

De manière générale, le *« spatial turn »* a participé au renouveau de la géographie aux Etats-Unis. Ainsi, si la géographie avait disparu de certaines universités, telles Harvard ou Stanford, le tournant spatial a participé de sa renaissance dans ces deux universités, prises pour exemple (Martí-Henneberg, 2011). Dans le premier cas, le Centre pour l'analyse géographique a signé le retour de la géographie dans la prestigieuse université de Harvard en 2006, qui s'est considérablement ouverte aux projets interdisciplinaires. A l'initiative de Peter K. Bol, éminent spécialiste des civilisations de l'est asiatique, la géographie à Harvard se spécialise dans l'information géographique (*GIScience*) au service des sciences de la terre, de la santé et de l'histoire[18]. A Stanford, la géographie est partie prenante du Centre des humanités numériques depuis la fin des années 1980, dans le Centre pour l'analyse spatiale et textuelle[19].

Bien que le terme ait moins de dix ans, on peut toutefois repérer quatre sous-branches dans les *GeoHumanities*, au croisement des tournants digital et spatial, largement semblable au développement des *Digital Humanities (Figure 2. 3)*. Ils sont explicités dans le préambule de la nouvelle revue *GeoHumanities*, éditée par l'Association des géographes américains (Dear, 2015). On trouve ainsi dans le premier numéro, de même que dans les sessions du colloque annuel de l'AAG à San Francisco en 2016, des contributions appartenant au trois premiers champs.

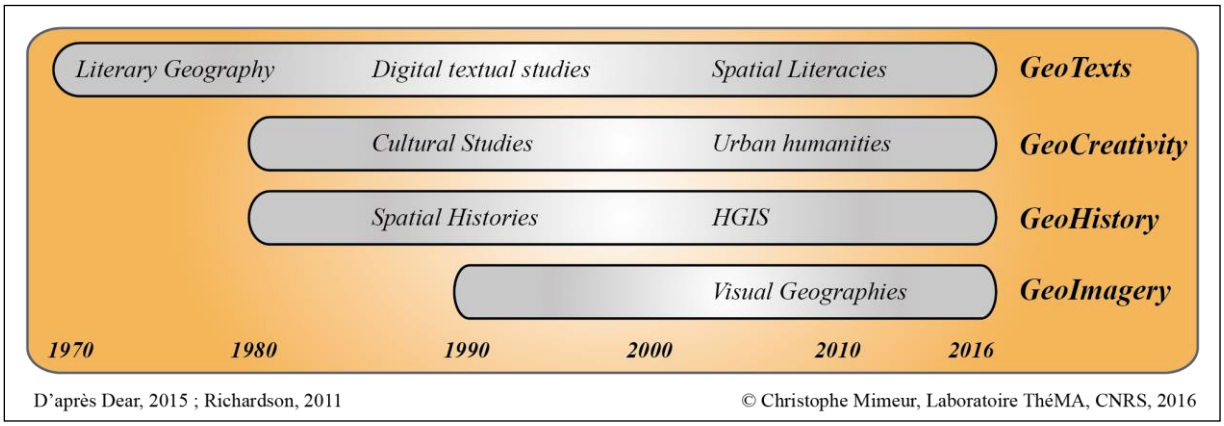


Figure 2. 3. GeoHumanities

Ce travail se situe en revanche dans les *Spatial Histories*. Considérant notre inscription dans le champ géohistorique français, la mesure de la temporalité n'est pas imposée par l'historien. Un ancrage davantage anglo-saxon place l'information géographique au cœur de notre travail. Ce sont les historiens qui sont les premiers enthousiastes au potentiel des sciences de l'information géographique (Dear, 2015), qui ne constitue pourtant qu'un pilier des *GeoHumanities*. Soucieux de souligner la prééminence géographique de notre travail de recherche, nous plaçons cette thèse dans le champ des *Spatial Humanities (Figure 2. 4)*, qui se définit comme fédérateur de chercheurs cherchant à introduire les concepts strictement géographiques au cœur des problématiques des

[18] *Report to the Provost on Spatial Analysis at Harvard University*, 2003
[19] http://digitalhumanities.stanford.edu

sciences humaines et sociales (Bodenhamer, Corrigan et Harris, 2010). Ainsi, dans leur ouvrage *Toward Spatial Humanities* (Gregory et Geddes, 2014), les deux auteurs Ian N. Gregory et Alistair Geddes associent le titre aux deux expressions *Historical GIS* et *Spatial History*.

Alors que nous retenons les possibilités de *« valorisation géoréférencée d'informations datées »* (Jacob-Rousseau, 2010), nous nous emparons de ce champ au croisement des humanités numériques, du tournant spatial dans les sciences humaines et sociales et dans la valorisation des lieux dans la science de l'information géographique. Si la large utilisation des SIG n'est plus à démontrer dans notre discipline, et qu'ils ne sont plus à définir, le passage au Système d'Information GéoHistorique (SIG-H) suggère l'adaptation des principes et des défis des classiques SIG, renfermant de nouveaux enjeux tant pour la géographie que pour la science de l'information géographique.

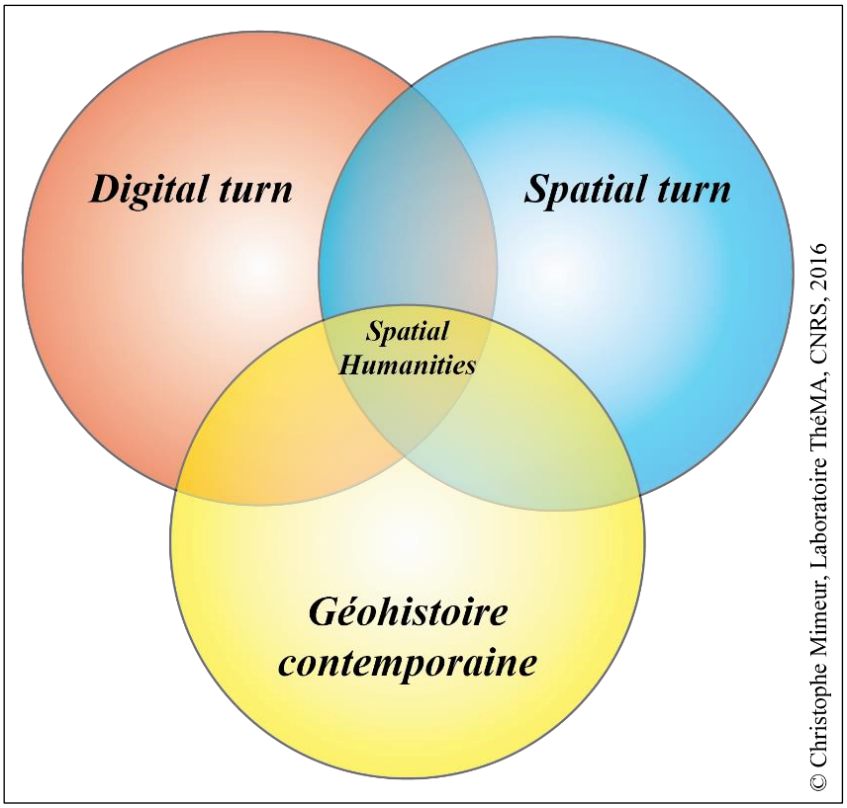


Figure 2. 4. Les *Spatial Humanities* dans la géohistoire contemporaine

2.2.2. **SIG-H : de l'information géographique à l'information spatio-temporelle**

En 1991, David J. Maguire dresse une liste de quatorze domaines d'application différents des SIG (Maguire, 1991), parmi lesquels il évoque l'occupation des sols, l'aide à la décision, l'inventaire urbain. Pourtant parmi eux, à aucun moment la dimension temporelle n'apparaît. On peut en revanche facilement identifier un rapprochement entre l'utilisation des SIG et l'archéologie (Barge et al., 2004) : ils sont une réponse au besoin de constituer des inventaires à l'échelle d'un chantier d'une part ; et à la spatialisation de bases de données structurées déjà existantes d'autre part. La culture d'inventaire est alors largement dominante en France au début des années 2000, suivant les directives du Ministère de la Culture. Les auteurs déplorent que *« le SIG n'est pas utilisé pour produire des résultats mais pour manipuler les données et constituer des documents »* (Barge et al., 2004). Aujourd'hui leurs analyses bénéficient des derniers développements en télédétection et des sciences de la terre.

C'est du côté de l'Association des Sciences Sociales pour l'Histoire (SSHA) qu'il faut trouver le dynamisme des SIG-H. La revue *Social Science History* est à l'origine d'un numéro spécial en 2000, dirigé par Anne Knowles. Dans son colloque annuel, la SSHA consacre des sessions spéciales grâce au groupe d'intérêt qui porte dans son titre les relations entre SIG et géographie historique. En 2005, un numéro spécial consacré au SIG-H est publié dans *Historical Geography*, dans laquelle la même coordinatrice souligne que la majorité des auteurs publie pour la première fois sur ce sujet, témoignant d'un outil nouvellement utilisé dans des perspectives historiques. En 2008, un colloque organisé par l'Université d'Essex rassemble 125 chercheurs en 21 sessions (Gregory et Geddes, 2014). Parallèlement à ces entrées thématiques, témoignant de l'émergence croissante des SIG-H, de premiers ouvrages à caractère général et méthodologique émergent sous l'influence d'Ian Gregory (Gregory et Ell, 2006, 2007 ; Gregory et Geddes, 2014 ; Gregory et Healey, 2007a). Si les premières années sont caractérisées par des démarches exploratoires adaptées à quelques cas d'études, alors leur approfondissement a requis le développement des techniques et des approches méthodologiques. Il s'agit ici de montrer comment les SIG-H ont obligé à repenser la manière de représenter l'information géographique, en y ajoutant la dimension temporelle.

Nous reprenons dans un premier temps quatre fonctions allouées au SIG classique par John Wilson et Stewart Fotheringham (Wilson et Fotheringham, 2008). En commentaire, nous confronterons deux exemples d'application : l'utilisation d'un SIG classique pour un diagnostic territorial, et l'utilisation d'un SIG-H pour une analyse des localisations patrimoniales *(Table 2. 3)*.

	SIG «classique»	SIG-H
Acquisition	Bases de données institutionnelles Données satellites / Télédétection	Archives statistiques / textuelles Archives photographiques
Archivage	Mise à jour des données automatiques / semi-automatiques	Retour à l'archive Contrainte du modèle de données
Abstraction	Cartographie statique	Cartographie dynamique
Analyse	Requêtes spatiales	Requêtes spatiales et temporelles

D'après Wilson, Fotheringham, 2008 ; Gregory, Ell, 2007

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 2. 3. Du SIG au SIG-H

Classiquement, l'information géographique est dotée de deux composants : un composant spatial qui permet de localiser un objet, et un composant attributaire, qui permet de le qualifier. Le SIG-H requiert l'introduction d'un troisième composant, temporel, tendant à complexifier les processus inhérents aux fonctionnalités d'un SIG *(Table 2. 3)*.

Les sources nécessaires à la collecte de l'information géohistorique constituent le premier enjeu. Si un recueil de données INSEE/IGN et propres aux collectivités est facilement intelligible, alors le recueil de chaque monument inscrit au patrimoine nécessite le recours aux référencements nationaux voire européens ou même d'études archéologiques dans le cas de monuments disparus. Le deuxième enjeu réside dans l'enrichissement des données attributaires. Si notre premier exemple impose un aperçu, le plus récent possible, du territoire, alors le second exemple suppose la prise en compte de la profondeur temporelle d'une même variable attributaire, comme sa fréquentation, son équipe de direction, nécessitant un recours à l'archive, pour chaque entité spatiale. Les requêtes spatiales sont un incontournable des outils SIG, pour définir par exemple l'aire de chalandise d'un commerce ou d'un service. En revanche, interroger la trajectoire d'une entité par la profondeur temporelle complexifie les requêtes spatiales dans un SIG-H. Enfin, si la cartographie dans le cadre d'un diagnostic territorial est une photographie du territoire, alors on peut se demander comment représenter les mutations des localisations des monuments dans un SIG-H, que la profondeur temporelle soit constituée d'aperçus à certaines dates, ou de manière continue dans le temps.

En 1964, B.J. Berry propose la construction d'une *« matrice d'information géographique »* (Berry, 1964) pour traduire la prise en compte conjointe de l'espace et du temps *(Figure 2. 5)*. D'abord, se référant à nos deux objets de recherche, les entités spatiales sont dans notre travail : le réseau ferroviaire d'une part ; le territoire, par ses unités administratives d'autre part. Chacun de nos deux objets sont ensuite qualifiés par leurs attributs respectifs, souvent quantitatifs.

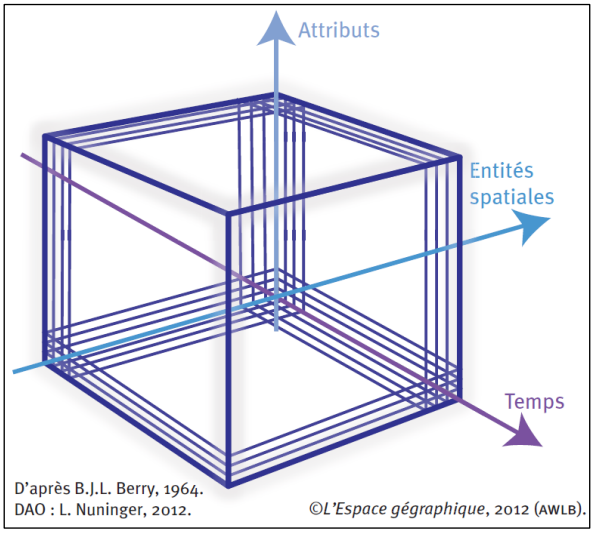


Figure 2. 5. La matrice d'information géographique.
D'après Berry H.J., 1964 ; FAVORY F. et *al.*, 2012

La prise en compte de la dimension temporelle est en revanche plus complexe. Ici, le recours à l'ontologie permet de représenter la complexité des objets spatio-temporels mobilisés dans un SIG-H. Issue d'un questionnement philosophique, l'ontologie est une science qui étudie les propriétés générales du monde, les objets qui le composent, qui s'organisent (Smith, 2001). Son recours se justifie dans le champ des *GeoHumanities* parce qu'elle participe au dialogue entre les sciences humaines et sociales et les sciences informatiques ; et parce qu'elle nous permet d'interpréter les principes géohistoriques explorés plus haut, adaptés au SIG-H. La représentation de l'information spatio-temporelle à l'intérieur d'un SIG-H obéit ainsi à deux conceptualisations différentes (Favory, Nuninger et Sanders, 2012 ; Grenon et Smith, 2004) :

- Les entités « SNAP » qualifiées d'endurantes : elles se caractérisent par une étendue spatiale permanente, dans tous les instants. Ce sont ces entités spatiales qui caractérisent notre objet « territoire » ;
- Les entités « SPAN » qualifiées de perdurantes : elles se caractérisent par une étendue spatiale, qui existe dans une étendue temporelle qui lui est propre. Dans notre travail, chaque entité de notre objet « réseau » n'a pas toujours existé.

Si ce découpage entre nos deux objets paraît intelligible, les entités SNAP et SPAN sont inséparables : une unité administrative est une entité SNAP, mais son processus dynamique, tel que l'évolution de sa population ou sa place dans la hiérarchie administrative relève des caractéristiques SPAN. En revanche, le réseau relève davantage d'une entité SPAN, dans la mesure où il n'est pas présent à tous les instants de la période étudiée, et possède des caractéristiques qui sont nécessairement dynamiques.

Dès lors, cette approche ontologique permet de replacer un outil de recherche, le SIG-H, au cœur de notre positionnement géohistorique. La conceptualisation de l'information spatio-temporelle autorise alors le chercheur à explorer l'espace de liberté que nous avons défini plus haut.

D'un point de vue méthodologique, l'appropriation d'un tel outil place notre travail de recherche dans la géomatique, participant au dialogue entre les sciences de l'informatique et les sciences humaines et sociales.

Ainsi, la complexification de l'information spatio-temporelle participe à de nouveaux défis incombés aux SIG-H. Les auteurs soulignent aujourd'hui les perspectives appliquées, participant à la construction de nouvelles connaissances historiques, géohistoriques.

2.2.3. **Les projets de SIG-H : expériences et enjeux**

Après avoir abordé les différents enjeux associés au passage de l'information géographique à l'information géohistorique, nous proposons ici une classification de SIG-H, dans laquelle nous accordons une place particulière à ceux traitant de réseaux de transport. Afin d'éviter un effet *« catalogue »*, nous proposons de donner un aperçu des SIG-H existants à travers une double lecture : celle du projet, et celle des échelles. Cette approche nous permettra de cerner les principaux enjeux et défis des SIG-H, d'un point de vue thématique et méthodologique.

Le panorama des SIG-H montre comment le terme de *« projet »* est largement associé à ces entreprises. Défini comme un *« ensemble finalisé d'activités et d'actions […] dans le but de répondre à un besoin défini dans des délais fixés et dans la limite d'une enveloppe budgétaire allouée »* (Boutinet, 1990), ce vocable semble tout à fait adapté à la démarche de construction d'un SIG-H. On peut ainsi rapprocher les différentes étapes du projet des fonctions principales allouées à un SIG *(Table 2. 3)*.

Premièrement, la construction d'un projet de SIG-H procède d'une commande. Elle est le plus souvent sociale. Nous avons interprété la renaissance de la géographie par une réponse à la contraction de l'espace-temps, où l'ubiquité et notre capacité à rejoindre tous les lieux en très peu de temps a fait renaître le besoin d'un attachement au lieu. Nous proposons la même lecture quant à l'émergence des SIG-H. Ainsi, dans son introduction à une rencontre entre professionnels de l'information géographique, Daniel Uny propose d'interpréter le *« petit regard dans le rétroviseur »* (Uny, 2014) par un besoin sociétal, patrimonial et mémoriel.

Deuxièmement, la construction d'un projet de SIG-H procède d'une facette unique et inédite, qui confronte le chercheur à un environnement incertain. Parmi les plus audacieux, on peut citer la construction de neuf SIG-H à l'échelle nationale. Les projets les plus précurseurs sont chinois (Bol et Ge, 2005) et britanniques (Gregory, 2005). D'autres ont été toutefois développés aux Etats-Unis (McMaster, Noble et Master, 2005), en Belgique (Vanhaute, 2005), en Russie ou en Corée du Sud. Parmi les enjeux les plus prégnants, on peut citer la délimitation des limites administratives. Ces éléments se rapprochent ainsi des problématiques géomaticiennes, dans lesquelles le *Modifiable Areal Unit Problem* (MAUP) tient une place importante (Wilson et Fotheringham, 2008). Etant donné les couvertures spatiales et temporelles, il apparaît que de telles entreprises représentent des projets de grande envergure, associés à des enjeux financiers et humains importants. Ils sont ceux qui mobilisent des fonds importants.

Troisièmement, la construction d'un projet de SIG-H procède d'une phase décisive de fédération de chercheurs, nécessaire à la prise de décisions et aux choix quant au contenu du projet. Les principaux résident alors dans la détermination des informations à incorporer avec les difficultés inhérentes à la dimension temporelle. Le chercheur est alors confronté à la rareté de la donnée, d'autant plus qu'il cherche une information quantifiée. Largement voisins des questionnements archéologiques, les questionnements relatifs à l'incomplétude, l'imprécision et l'incertitude des données sont des composantes inévitables d'un projet SIG-H, qu'il faut réinvestir à toutes les étapes du projet. L'exemple du développement des SIG-H urbains est éclairant à ce propos. En reconstituant l'histoire urbaine de certaines agglomérations, ce type de SIG-H propose souvent des entrées thématiques, issues du choix et de la disponibilité des données. Ainsi le SIG-H de New-York propose l'évolution de la population et les mouvements de concentration et déconcentration ethniques, à travers une série de plans. D'autres exemples sont plus ponctuels, où le SIG-H de Tokyo propose le recensement des infrastructures de transport et le SIG-H de Sydney propose une répartition des objets d'arts.

Les SIG-H consacrés aux infrastructures attirent une particulière attention dans notre travail de recherche et constituent par ailleurs une bonne illustration des trois écueils inhérents à la construction d'un projet de SIG-H. Ainsi, la première initiative américaine, déjà mentionnée dans le *Chapitre 1*, affirme avoir privilégié la quantité de l'information à la précision des données, faisant référence au deuxième défi posé par les SIG-H. Plus détaillé et dérivé du programme national de Grande-Bretagne, la numérisation des chemins de fer britanniques bénéficient de la précision de l'Atlas du Colonel Cobb (2006), faisant aussi écho au second défi posé par les SIG-H. L'initiative la plus complète est européenne : elle a bénéficié du financement de la Fondation Européenne pour la Science (ESF) et recense plusieurs dizaines de milliers de kilomètres dans 22 pays différents, en devant toutefois s'accommoder de la diversité des sources des différents pays (Morillas-Torné, 2012), rejoignant les deuxième et troisième défis d'un SIG-H, mais répondant aussi au premier. Si la rareté et l'hétérogénéité des données constituent un défi, alors le SIG-H est en revanche un formidable outil qui permet de combiner des types de données très différents.

Les trois défis des projets de SIG-H sont donc communs aux initiatives européennes, nationales et urbaines, même appliquées aux SIG-H dédiées au transport et sont l'expression des différentes étapes du projet. Ils sont d'ailleurs de la même façon confrontés à trois challenges. Nous proposons de rendre compte de la chronologie d'un projet de SIG-H, au croisement des préoccupations thématiques d'une part et méthodologiques d'autre part.

Le premier challenge réside dans le coût de tels projets. En effet, le défi d'incertitude et de rareté de la donnée, couplé aux difficultés d'ordre technique, ne permet souvent pas une collecte automatisée de données. Le SIG-H national britannique a coûté 3,5 millions de dollars, somme difficilement mobilisable par les sciences sociales (Southall, 2014). L'auteur montre comment le projet a dû mobiliser des chercheurs dans le domaine de la santé, des agences environnementales gouvernementales et des mécènes privés pour construire ce vaste chantier. Ainsi, la mise en place de projets nécessite de manière croissante de spécifier les enjeux méthodologiques mais aussi thématiques, où l'objectif est de démontrer la valeur économique et sociale d'un tel projet, dans la mesure où il bénéficie à d'autres domaines d'application.

Le second challenge procède des questionnements méthodologiques, que nous qualifierons de **« traditionnel »**. De ce point de vue, nous observons un « culte de la collecte », qui est chronophage et coûteux. Les phases de choix des sources, d'évaluation et de digitalisation délaissent peu à peu les préoccupations thématiques pour se concentrer sur l'architecture méthodologique nécessaire à la constitution du SIG-H. Nous posons ici l'hypothèse d'une maturité des chercheurs dans leur capacité à initier un projet de SIG-H, leur capacité à mobiliser des sources et à l'alimenter. Nous citerons toutefois ici le glissement vers une conception de SIG-H avec davantage de données qualitatives, ouvrant de nouveaux champs d'investigations (Gregory et Geddes, 2014).

Le troisième challenge procède au retour aux questionnements thématiques, que nous qualifierons ici d'**« émergent »**. Cet enjeu relève d'une capacité à trouver une audience élargie et plus intéressée aux résultats du projet plutôt qu'à la méthodologie employée pour construire le SIG-H (Gregory et Geddes, 2014 ; Knowles, 2005). Nous posons toutefois l'hypothèse d'un nécessaire retour à des préoccupations méthodologiques, cette fois-ci en lien constant avec les réponses à apporter à notre question.

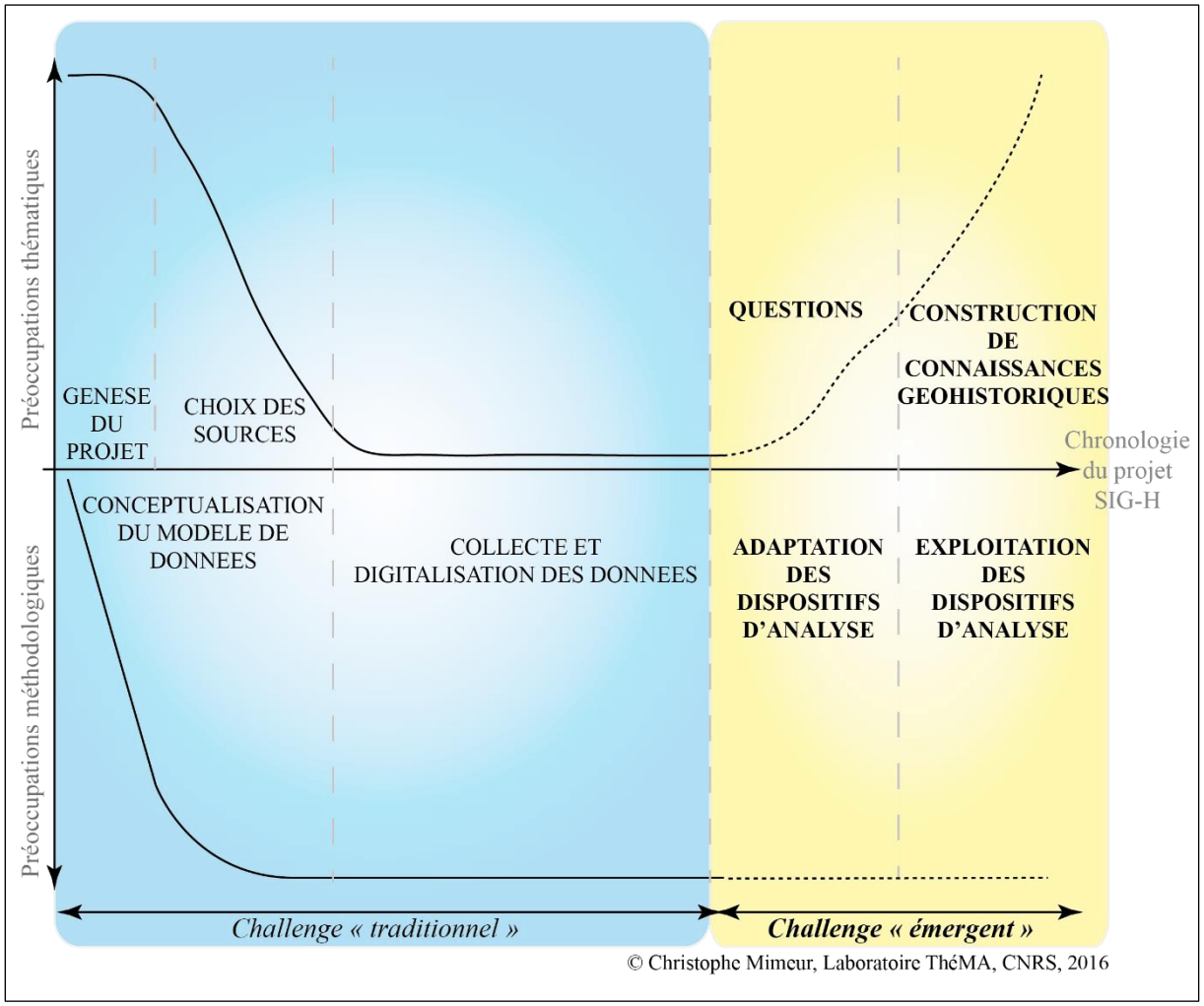


Figure 2. 6. Le projet du SIG-H : dépasser la collecte pour l'analyse

Nous positionnons notre travail de recherche dans ce challenge émergent. Le parallèle entre les préoccupations méthodologiques et thématiques montre comment un projet de SIG-H possède une dimension adaptative très forte, en lien avec l'environnement incertain dans lequel sa genèse opère. Il démontre par ailleurs que ce projet n'a pas de démarche intégrée, venant corroborer l'espace de liberté dans lequel notre positionnement géohistorique nous insère. La première phase du challenge émergent mélange questionnements scientifiques et construction méthodologique *(Figure 2. 6)*. Dès lors, l'espace de liberté défini plus haut nous oblige d'autant plus à définir les véritables questions de notre recherche, et nécessite de montrer comment nous nous insérons dans une approche quantifiée du lien réseau/territoire, dans des cadres de validité qu'il reste à construire.

2.3. Vers la définition de cadres d'exploration du lien réseau/territoire

« Il n'y a plus vraiment de géographes, d'économistes ni de spécialistes des transports pour croire aux effets structurants des transports » (Offner, 2014). Ainsi parle Jean-Marc Offner, vingt ans après son article provocateur. Pourtant, en se replaçant à l'échelle de la longue durée, nous posons l'hypothèse que cette affirmation n'est pas aussi définitive (Bretagnolle, 2014 ; Pumain, 2014). Pour cela, nous avons vu que nous nous dotons de nouveaux cadres spatio-temporels par la géohistoire, d'outils permettant de prendre en compte l'information géohistorique par le SIG-H.

Nourrir le débat et participer à la construction de nouvelles connaissances du couple réseau/territoire nous poussent ici à entrer dans les débats épistémologiques du statut des *« sciences historiques »* (Passeron, 1991) en général, de la production de connaissances du lien réseau/territoire en particulier. Denise Pumain pose l'hypothèse que le processus d'accumulation de connaissances procède de contextes épistémologiques et de conditions techniques (Pumain, 2005), distinction que nous conservons dans l'application de la construction de connaissances du lien réseau/territoire. Cette distinction trouve aussi sa place dans le challenge émergent dans lequel nous nous inscrivons dans les projets de SIG-H, qui suppose la construction des conditions de production de connaissances, sur les plans thématiques et méthodologiques.

2.3.1. **Du mythe à la science, du positivisme au constructivisme**

Nous proposons ici une entrée successive des grands positionnements épistémologiques quant aux connaissances de notre thématique de recherche, dont l'intérêt réside ici dans le passage du « mythe » à la production scientifique de savoirs. Les débats oscillant entre le déterminisme et le possibilisme ont été nombreux dans la géographie tout au long du XXème siècle. Peu à peu, les réflexions se sont enrichies des contextes culturels, des mécanismes de décisions (Claval, 1985). Alors que le déterminisme faisait de la causalité une primitive de l'explication géographique, le

possibilisme *« décourageait toute recherche de lois et toute explication systématique »* (Claval, 1985). Il s'agit plus largement ici de l'opposition entre des démarches inductives et déductives, au croisement de la compréhension et de l'explication.

Ici, nous estimons que l'explication de notre démarche épistémologique est encore plus justifiée car nos objets de recherche sont le moteur d'un mythe solide, perpétué depuis deux siècles (Offner, 1993a). En sciences sociales, la dimension du « mythe » a été largement abordée (Lévy et Lussault, 2003) : elle est souvent évolutionniste et largement positiviste. Les prémices du positivisme scientifique sont à chercher dans la philosophie d'Auguste Comte, dont le secrétaire n'est autre que Saint-Simon, à l'origine du mythe des effets structurants. Pour eux, l'étude des faits permettent davantage de décrire le monde que l'expliquer, parce que la recherche des causes d'un phénomène est à rejeter. Le mythe ainsi construit selon des considérations socio-techniques a dérivé vers un mythe scientifique. Dès lors, on parle de réalisme, dans la mesure où le positivisme fait l'hypothèse que les objets observés sont effectivement dans la nature. La compréhension du monde est donc ici inductive, dans laquelle les observations aboutissent à l'énoncé d'une loi. Elle est aussi rationnelle dans la mesure où le positivisme présuppose une méthode stricte, souvent mathématisée. Aujourd'hui, on a plutôt tendance à éviter de rentrer dans des considérations positivistes, parce que, selon leurs détracteurs, elles ne prennent pas en compte la complexité de l'humain et érigent naïvement par-dessus tout le progrès (Lévy et Lussault, 2003), en opposant les sciences de la nature et les sciences de l'homme, à l'instar de Karl Popper.

Face à cette démarcation, on pose la question du statut des *« sciences historiques »* (Passeron, 1991) en général, de la production de connaissances du lien réseau/territoire en particulier, dont les régimes de vérité seraient différents de ceux des sciences de la nature. Dans ce sens, notre discipline doit prendre en compte la diversité des réalités géographiques. En s'inscrivant dans une *« exigence épistémologique de contextualisation de l'explication »* (Pumain, 2014), cette approche vient ainsi en contre-point de la critique systématique des démarches empiristes et inductives décriées par Karl Popper au XXème siècle. Ainsi, Passeron affirme dans le titre de son livre que le raisonnement sociologique opère dans un espace *« non-poppérien »* (Passeron, 1991). La construction de nouvelles connaissances, dans ce cadre, n'est pas nécessairement nomologique. On pose alors la question de l'accumulation des connaissances à l'intérieur des sciences historiques. Questionner les effets structurants des infrastructures de transport sur le territoire nous confronte nécessairement à des réalités locales et des variations temporelles. Alors, le recours à un raisonnement naturel *« n'implique pas que les sciences humaines et sociales soient condamnées à l'exceptionnalisme, à l'affirmation réitérée de chaque cas »* (Pumain, 2014). Ainsi, on peut identifier un *« autre processus, plus continuiste, qui permettrait la réinterprétation d'une théorie ancienne dans une nouvelle, plus englobante, ou une évolution contrôlée du contenu des concepts »* (Pumain, 2005). Plutôt qu'une opposition entre induction positiviste et déduction poppérienne, entre réalisme et nomologie, nous plaçons notre travail dans une posture intermédiaire, dans laquelle le constructivisme trouve un écho sensible dans notre question de recherche.

La définition d'un champ constructiviste se trouve dans ce que certains appellent la *« guerre des sciences »* (Hacking, 2001). Les grands principes sont évoqués dans l'ouvrage de Peter Berger et Thomas Luckmann intitulé *« La construction sociale de la réalité »* à partir des rapports sociaux (Berger et Luckmann, 1966). Dès lors, on pourrait attester d'un rapprochement entre réalisme et nomologie quand on considère qu'il existe effectivement des objets dans le monde réel, lesquels font l'objet de catégorisations et donnent naissance à des concepts ou de grandes idées. Il faut signaler que *« les catégories utilisées par les scientifiques sont toujours relatives à un état global de leurs connaissances et construites socialement, dans une interaction complexe entre l'état du monde et son interprétation par la science »* (Pumain, 2014). Mais dans une perspective longue, la multitude de contextes dans l'espace et dans le temps rendrait d'après Passeron les comparaisons et la cumulativité des connaissances impossibles (Passeron, 1991). Pour autant, Denise Pumain souligne que *« puisque l'histoire du monde social est finie, le savoir acquis sur cette énumération de cas doit en principe être cumulable, même s'il s'agit d'interprétations partielles, voire erronées, toujours révisables »* (Pumain, 2005). Questionner les effets structurants des infrastructures de transport sur la longue durée est alors possible quand *« toute connaissance énoncée devrait nécessairement se référer au cas ou à la série d'exemples cautionnant l'interprétation proposée »* (Pumain, 2014). Une démarche constructiviste et géohistorique permet de définir un contexte *« non pas tant par le lieu ou la date exacte de l'observation, mais par le choix d'un certain niveau de résolution du découpage spatial et de la séquence de durée pour lesquelles les conclusions de l'observation sont énoncées »* (Pumain, 2014).

Dans ce cadre, nous défendons une position abductive, telle qu'elle a été introduite dans la pensée constructiviste par le pragmatique Charles Peirce, comme seul mode de raisonnement permettant d'aboutir à de nouvelles connaissances. C'est ainsi qu'un certain nombre d'auteurs parlent de *« théories d'arrière-plan »* (Quine, 1977) qui constituent un référentiel dans lequel on place les connaissances géohistoriques. L'abduction est alors à replacer dans des perspectives empiriques : elle n'est pas purement inductive parce que le point de départ n'est pas une simple série d'observations. Elle n'est pas plus purement déductive dans la mesure où l'expérience participe à la construction d'une conjecture. L'abduction relève de deux étapes : *« de l'hypothèse, on tire par déduction des conséquences que l'on soumet à l'épreuve »*. Par la suite, *« dans une phase inductive, qui n'est pas ici un processus de construction d'hypothèses par généralisation, mais une évaluation empirique »*, on participe à la construction de nouvelles connaissances.

Si Jean-Claude Passeron compare le raisonnement sociologique à un *« va et vient »*, alors nous l'appliquons ici au raisonnement géohistorique, qui *« se distingue du récit historique par des moments de raisonnement expérimental, mais ces moments de pureté méthodologique alternent nécessairement dans son travail interprétatif avec d'autres moments du raisonnement naturel. […] Le raisonnement statistique est bien un raisonnement expérimental mais il ne le reste qu'autant qu'il n'énonce rien sur le monde historique : dès qu'on met du sens dans l'énoncé de ces corrélations formelles, les phrases se chargent de contexte »* (Passeron, 1991). Ainsi, l'auteur propose trois phases d'énonciation de la connaissance *(Figure 2. 7)*, que nous adaptons à notre travail de recherche, au croisement de la géohistoire et des *Spatial Humanities* :

- *« Enonciation informative »* : il s'agit ici de la collecte de données empiriques, identifiée comme le challenge traditionnelle des SIG-H ;
- *« Effets de connaissances »* : ils sont le challenge émergent des SIG-H, parce qu'il se trouve au cœur du système socio-spatial, entre le lieu, son environnement et la trajectoire individuelle ;
- *« Effets d'intelligibilité »* : il s'agit de la dernière phase d'un projet SIG-H, qui permet la formulation de généralités théoriques, portant sur un monde empirique. Le contrôle empirique est rendu possible par un retour aux données géohistoriques, et laisse la possibilité de reformuler des hypothèses.

Alors, là où notre interprétation d'un projet de SIG-H montrait une succession des différentes étapes du projet *(Figure 2. 6)*, le débat épistémologique traitant des sciences historiques comment le processus de constructions de connaissances n'est pas linéaire et comment le mouvement de « va et vient » est une partie intégrante d'un projet de SIG-H.

Aussi, à travers ces trois niveaux de formulation de la connaissance géohistorique, une démarche constructiviste suppose la garantie de deux critères complémentaires, ce qu'Alain Desrosières appellent les critères de **précision** et les critères de **pertinence**. Dans une perspective de *« division sociale du travail »* (Desrosières, 2010), cette combinaison peut alors passer par l'interdisciplinarité. Mais au cœur de notre question de recherche, la géohistoire intègre en elle-même cette division. Les deux sections suivantes ont pour objet de définir les critères de précision et de pertinence à établir.

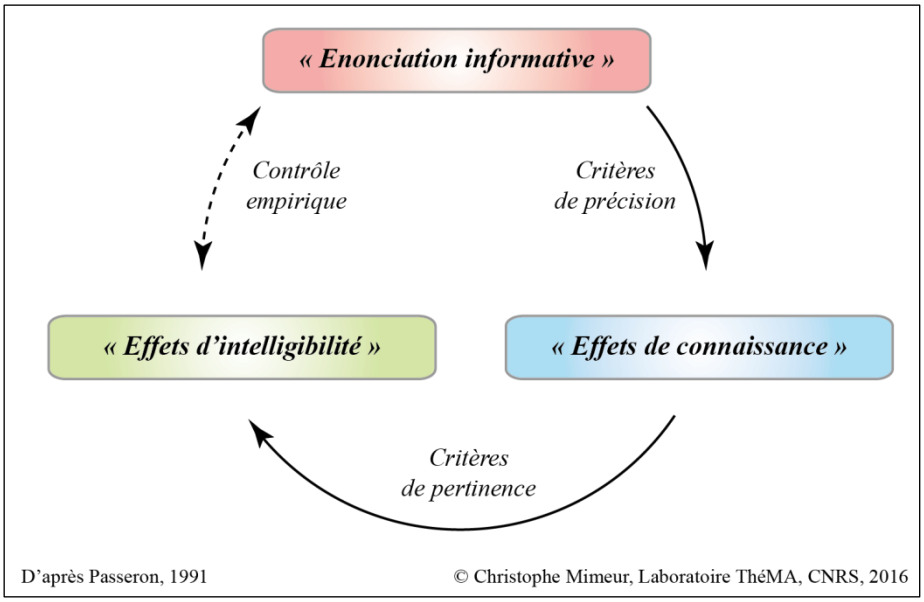


Figure 2. 7. Le processus de construction de connaissances géohistoriques

2.3.2. **Quantifier le lien entre réseau et territoire : des critères de précision entre convention et mesure**

Alors que nous avons discuté des considérations épistémologiques nécessaires à la production de connaissances géohistoriques, il faut voir ici comment elles s'imbriquent dans des conditions techniques (Pumain, 2005). La démarche constructiviste participe de cette articulation : comme mode d'observation du réel, la catégorisation permet d'adopter une méthode réflexive. Il s'agit ici d'objectiver le réel, quand *« une partie de l'activité scientifique se résout dans la production d'informations, archivables et donc cumulables : résultats d'enquêtes, dénombrements, inventaire, typologies, jurisprudences, archives »* (Pumain, 2005). Il faut voir aussi dans le développement des moyens techniques et informatiques une accumulation des informations qui influe sur les évolutions épistémologiques de la connaissance.

Nous plaçons ce travail dans la géographie théorique et quantitative, pour la construction de connaissances géohistoriques du lien entre réseau et territoire. Le sociologue Alain Desrosières a largement étudié l'histoire de la statistique. Nous reprendrons sa définition de la quantification : *« quantifier, c'est convenir puis mesurer »* (Desrosières, 2008). Les définitions classiques, telles celle du dictionnaire Larousse retiennent la deuxième acception : *« traduire quelque chose en une quantité mesurable »*. La première acception relève quant à elle de la définition des objets conceptuels qu'il s'agira seulement ensuite de mesurer, sous forme numérique le plus souvent. Notre travail de recherche se trouve ainsi pleinement au cœur de ces deux acceptions : la définition des objets conceptuels est fondamentale, dans un contexte géohistorique et une démarche constructiviste, pour ensuite envisager une mesure du lien entre réseau/territoire. Dès lors la quantification passe par la définition des objets pour mieux ensuite analyser leurs interactions : on rejoint aussi ici la démarche abductive, parce que notre questionnement n'est pas uniquement guidé par l'information géohistorique, mais doit s'inscrire aussi dans des théories d'arrière-plan.

La conclusion de notre premier chapitre a permis de constater une permanente controverse des effets structurants des grandes infrastructures de transport. Il s'agit ici d'établir des conventions, et de se doter de théories d'arrière-plan permettant d'investiguer cette question à la lumière des éclairages épistémologiques que nous avons identifiés. La quantification est donc ici issue de notre construction sociale de la réalité, même si souvent, les mesures qui en découlent *« tendent à devenir la réalité par un effet de cliquet irréversible »* (Desrosières, 2008).

D'un côté, le réseau est le premier concept de notre travail de recherche. Il est défini comme un *« système de transport »*. Dans notre perspective géohistorique et constructiviste, nous choisissons de nous concentrer sur le réseau ferroviaire français, depuis ses prémices. Par sa logique, sa structure et sa dynamique *(cf. Chapitre 1)*, nous faisons l'hypothèse que l'évolution du réseau est un marqueur des *« traces du temps »* (Durand-Dastès, 1990). En revanche, la géohistoire fait davantage référence aux temporalités qu'au temps. En bornant notre étude à la charnière des

XIXème et XXème siècles, la temporalité adoptée n'est en rien imposée par les historiens, mais par l'innovation ferroviaire et la mise progressive en réseau. Une seconde temporalité est à imbriquer à la première, et a trait à l'histoire des mobilités (Flonneau et Guigueno, 2009).

De manière conventionnelle, nous posons l'hypothèse que l'extension du réseau et le développement de ses performances sont caractéristiques des effets d'un système de transport. A ce moment, on peut envisager la mesure de ces deux éléments :

- l'extension du réseau peut être appréhendée par la mesure linéaire des chemins de fer en France ;
- les performances du réseau peuvent être traduites par les vitesses moyennes sur les traverses qui sillonnent la France.

Alors, cette quantification du réseau permet de *« réfléchir non pas sur le court terme et sur l'impact d'un tronçon, comme ont tendance à le faire la plupart des acteurs politiques, mais avec la prudence et le recul de l'histoire »* (Bretagnolle, 2014). La première information permet d'appréhender la couverture spatiale du réseau *(Figure 2. 8)*. La seconde information permet de questionner le concept de *« grande vitesse » (Figure 2. 9)*. Jusqu'au début du XIXème siècle, les déplacements sont caractérisés par l'*« immense piétinement des Hommes »* (Bavoux et al., 2005), qui ne dépassent pas 5 km/h. Les premiers systèmes de transport organisés sont les diligences, qui permettent de relier des relais postaux, et de multiplier la vitesse par deux par la même occasion avec la traction hippomobile (Bretagnolle, 2003). L'irruption ferroviaire marque une réelle rupture, et permet la *« révolution de la vitesse »* (Studeny, 1995). Ainsi, la grande vitesse est relative au contexte dans lequel s'inscrit le développement du réseau ferroviaire. La vision sur la longue durée que nous avons ici du réseau n'est pas statique mais bien dynamique et est l'expression d'une *« convergence spatio-temporelle »* (Janelle, 1969).

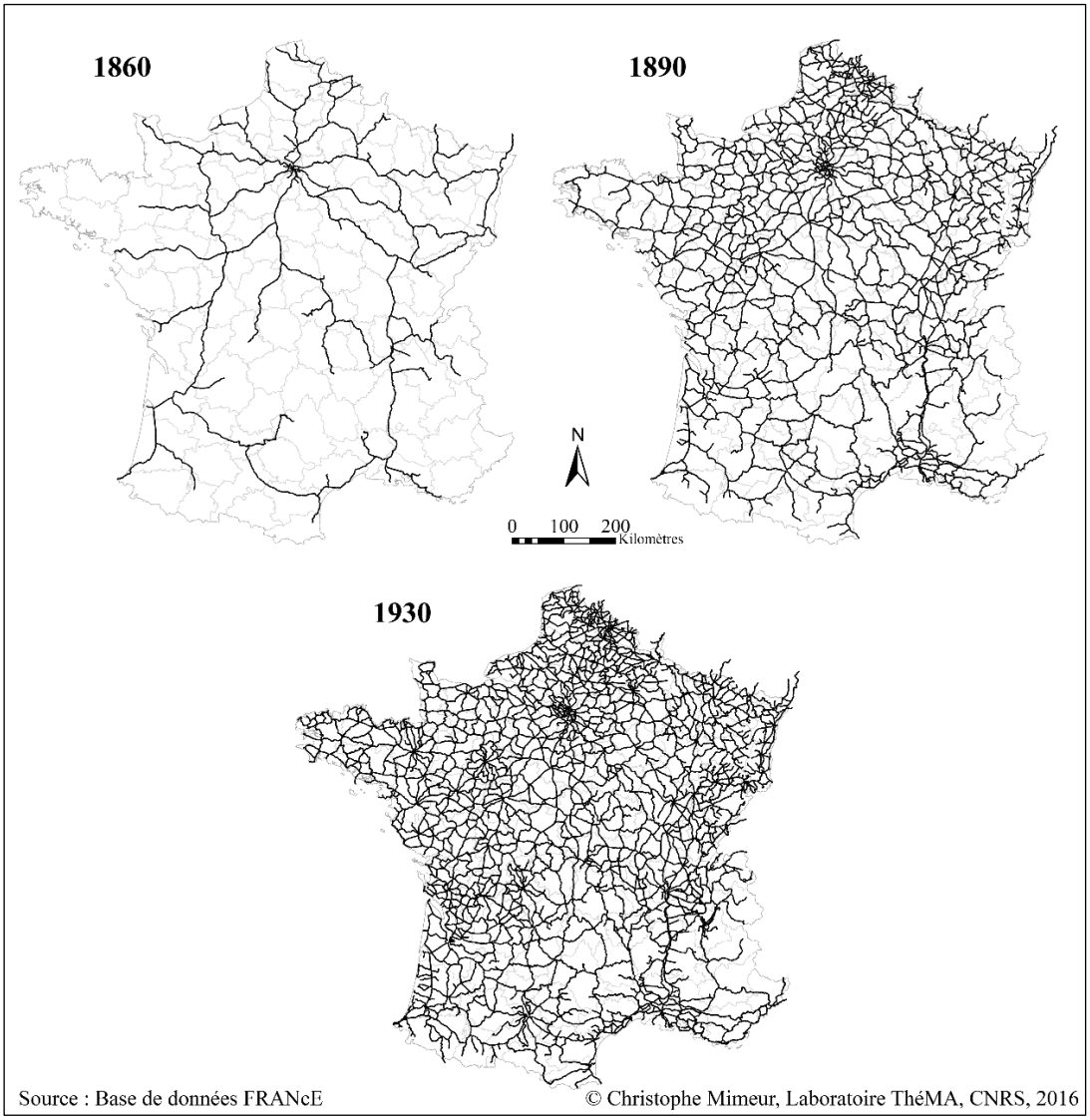


Figure 2. 8. Evolution du réseau ferroviaire français entre 1860 et 1930

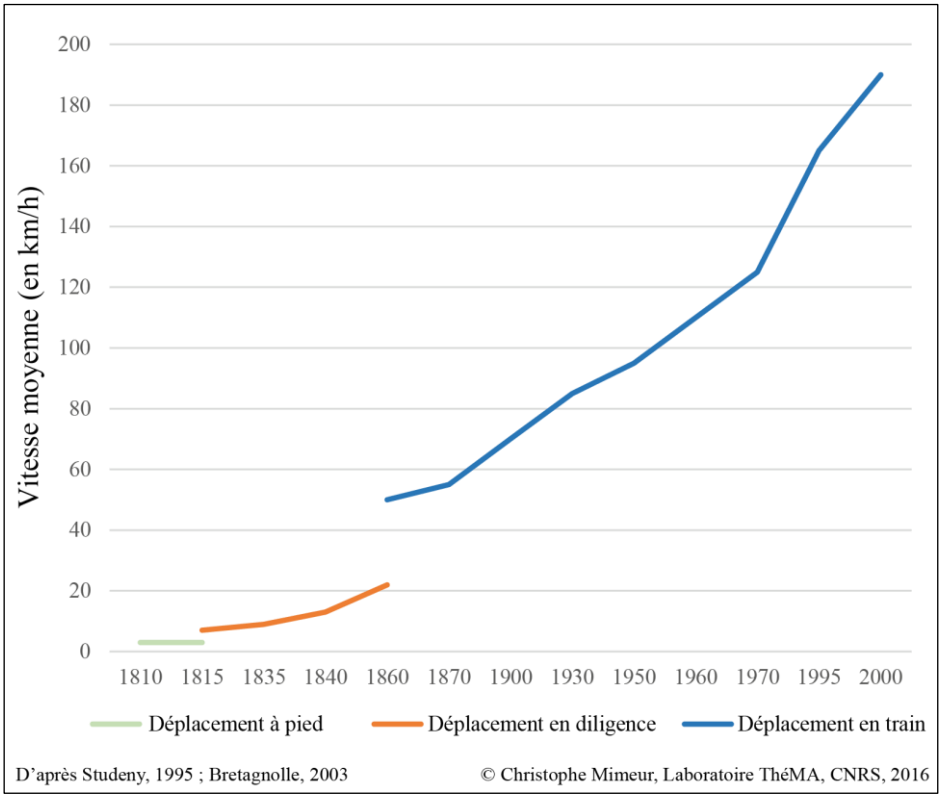


Figure 2. 9. Evolution de la vitesse moyenne de déplacement entre 1815 et 2000

D'un autre côté, notre second objet de recherche est le développement des territoires, nous confrontant aussi ici aux contraintes techniques de l'accumulation de connaissances. Il s'agit là aussi d'établir des conventions et des mesures. Les géographes ont montré comment l'étude des villes et de leurs évolutions individuelles ont largement été traitées de manière statique (Pumain, 1997). Initiés au début des années 1990 en France, des modèles dynamiques tentent d'expliquer *« la forte différenciation des villes par leur taille et leurs fonctions […] à partir des relations entre les villes et entre les villes et leur environnement »* (Pumain, 1997). Telle est donc la posture conventionnelle que nous adoptons ici. Elle entre ici en résonance avec le système socio-spatial que nous avons défini dans la première section, comme intersection entre un lieu, son environnement et sa trajectoire individuelle.

Pour leur mesure, nous retenons également une description simplifiée des villes : *« la variable de la taille démographique est un excellent et le meilleur résumé de très nombreuses propriétés fonctionnelles. […] En outre, sur de très longues durées, la taille a le mérite de pouvoir être évaluée assez simplement, de façon comparable »* (Pumain, 1997). Pour autant, de manière également conventionnelle dans l'étude des dynamiques démographiques, la tendance est de séparer ce qui relève de l'accroissement naturel de ce qui relève des migrations. On pose alors l'hypothèse que *« l'aléa des décisions individuelles : faire naître un enfant, migrer là plutôt qu'ailleurs, installer une entreprise peut aboutir à une évolution globale relativement stable et déterminée de la ville »* (Guérin-Pace, 1993). Il a été également démontré comment la taille des villes est souvent corrélée à *« beaucoup de descripteurs quantitatifs, comme le nombre d'emplois, d'établissements*

ou de logements mais aussi qualitatifs, comme la diversité ou la rareté des activités et la variété des populations présentes » (Pumain, 1997).

La mesure de la taille des villes repose alors sur le recensement de la population. A partir de 1801, le recensement devient systématique dans l'espace et dans le temps. L'unité spatiale de référence est la commune, avec des critères de domiciliation qui ont évolué dans le temps mais dont la stabilité remonte à 1841. Ici encore, la granularité spatiale n'est pas imposée par l'historien mais par l'objet de notre recherche, permettant d'avoir une finesse à l'échelle des quelques 36 000 communes françaises. Ainsi, alors que des travaux de géographes ont traité de la croissance urbaine en lien avec le chemin de fer (Bretagnolle, 2009 ; Pumain, 1982), notre travail prend en compte l'ensemble de la population française, qu'elle soit urbaine ou rurale. Le raisonnement peut donc être conduit à l'échelle d'un système de peuplement national, où chaque sous-système représente alors une commune.

Enfin, dans une démarche quantifiée, le compromis entre convention et mesure nous permet d'*« accepter des étapes qui simplifient la description tout en gardant l'essentiel de la dynamique. La variable population répond bien à ces deux critères »* (Guérin-Pace, 1993). Parmi les principales difficultés, la définition du seuil urbain est déterminante : *« l'INSEE a adopté la définition des communes urbaines un seuil de 2 000 habitants agglomérés, mais de nombreux historiens en récusent la validité »* (Guérin-Pace et Pumain, 1990). D'autres parlent d'un seuil à 3 000 habitants. Ainsi, le choix d'un seuil à 2 500 habitants paraît intermédiaire et a été retenu par de nombreux géographes (Pumain, 1982, 1997), et fait l'objet d'une catégorisation issue de notre propre construction sociale de la réalité. Nous insistons ici sur la prise en compte, non pas seulement des villes, mais de l'ensemble des communes. C'est là que l'on identifie la nécessaire combinaison de critères de précision et de critères de pertinence.

2.3.3. **Des cadres de validité dans des critères de pertinence**

Les critères de pertinence évoqués par Alain Desrosières relèvent d'après nous de deux ordres de considérations différents. Le premier réside dans les *« lunettes »* que l'on porte pour construire des connaissances géohistoriques du lien entre réseau et territoire, discutant d'une démarche introduisant la complexité. Le second réside dans l'appropriation de théories d'arrière-plan appliquées à notre question de recherche.

On peut d'abord se poser la question d'un nécessaire recours à la *« systémique »* quand on parle d'un *« système de transport »*. Or, il apparaît plutôt *« sectaire »* de confondre *« systémisme avec l'esprit de système »* (Pumain, 2003). Défini comme un *« ensemble d'éléments qui sont reliés et coexistant [le système] constitue une totalité organisée distincte de son environnement »* (Lévy et Lussault, 2003). L'étude systémique relève d'une démarche constructiviste et cherche à identifier la multitude des composants d'un système par leurs définitions et leurs relations. Le système socio-spatial dont il a été question dans la première section de ce chapitre en est l'illustration. Cette démarche semble alors tout à fait adaptée pour étudier les effets structurants des grandes

infrastructures de transport sur le développement territorial, d'autant qu'ils permettent de dépasser des causalités simples. Pour autant, se limiter à une démarche systémique comporte selon nous plusieurs écueils, qu'il convient de signaler, lorsqu'il est confronté à une démarche géohistorique :

- Le recours à la longue durée ne permet pas d'obtenir de *« données suffisamment précises »* (Bretagnolle, 2014) nécessaires à la description d'un système et de son fonctionnement ;
- La multiplication des cas particuliers pour appréhender l'ensemble des composantes d'un tel système vient flouter la clarté des hypothèses et des conclusions énoncées ;
- Définir ce qui est à l'intérieur d'un système de ce qui est à l'extérieur est une entreprise complexe, ce qui rend l'identification des perturbations (Elissalde, 2000) au système tout aussi difficile.

Si la scène du raisonnement systémique stimule la réflexion de la définition des objets et de leurs interactions, la multiplication des composantes et des possibilités d'interrelations d'un système participe à la complexité de sa compréhension. Alors, se limiter à cette analyse viendrait selon nous alimenter la diversité sémantique et la boîte noire associées aux effets structurants des infrastructures de transport. Loin d'opposer systémique et complexité, nous identifions de manière complémentaire les apports du paradigme de la complexité, défini par Edgar Morin, comme part entière de la science des systèmes. Alors, *« la complexité sociale peut-être rendue plus intelligible si l'on accepte d'énoncer des propositions, lois ou modèles, regroupant ces processus et leurs résultats les plus fréquents, selon une opération dite de simplexité »* (Pumain, 2014). Le recours à la complexité peut-être rationnel en dépassant ce qui relève de l'idiographie et de la nomothétie (Pumain, 2005). Elle trouve aussi toute sa place dans une démarche empirique, quand elle est définie comme *« le nombre d'interprétations non équivalentes qu'un observateur peut se faire d'un système »* (Livet, 1983). Elle est valide dans une démarche intermédiaire, abductive.

Les objets de recherche tels qu'ils ont été définis dans la section précédente trouvent ici leur place dans le cadre de réflexion proposé par la systémique et dans le cadre d'expérimentation de la complexité. Compte-tenu de ce positionnement intermédiaire, notre démarche de recherche s'oriente vers la construction de connaissances étudiant l'évolution de la distribution d'un système de peuplement dans un espace en rétraction permis par l'expansion d'un système de transport. La démarche systémique suppose la définition de propriétés des objets et la nature des relations qu'ils entretiennent. Le recours à la complexité permet leur simplification à partir de variables robustes. La suite de cette section propose d'énoncer les principales questions de notre recherche, selon deux ordres : ceux qui portent sur les propriétés de l'objet, ceux qui portent sur les relations d'un objet sur l'autre. Fidèles à notre positionnement, nous enrichissons nécessairement ces réflexions par des théories d'arrière-plan. Ces questions montrent comment notre travail de recherche se trouve selon nous au cœur de la scène de l'aménagement du territoire *(Figure 2. 10)*.

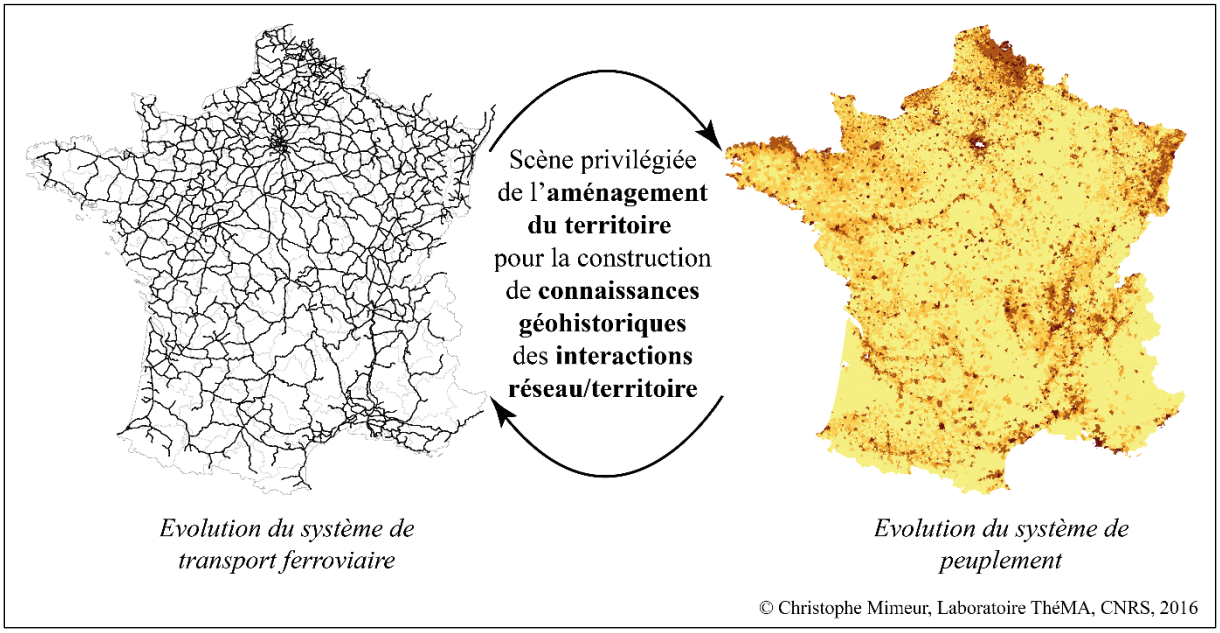


Figure 2. 10. La scène de l'aménagement du territoire face à notre enjeu de recherche

Le premier ordre des critères de pertinence porte sur la nature et les propriétés des objets de notre recherche. Ainsi, notre démarche incite à la définition de propriétés propres au réseau. Quel processus préside à la diffusion de l'innovation ferroviaire ? Il s'agit ici de qualifier le phénomène d'évolution du réseau ferroviaire, à travers la dimension logistique propre aux innovations. Par ailleurs, la question d'une diffusion homogène se pose sur le territoire français et passe par l'identification de critères venant comprendre et expliquer l'hétérogénéité du développement du réseau ferroviaire. Un des objectifs de ce travail de recherche est aussi d'éclairer comment le développement du réseau est le fruit d'une hiérarchisation du réseau. Ici, c'est l'enrichissement de l'information géohistorique d'un point de vue méthodologique qui permettra de la qualifier et de la confronter ensuite aux grands plans d'aménagement, tels qu'ils ont été pensés et votés au moment du développement du réseau ferré.

Quel processus préside à l'évolution du système de peuplement ? Le raisonnement que l'on propose se concentre sur la distribution des populations et ses reconfigurations, si elles existent. Dans ce cadre, l'apport des études sur les systèmes de villes, les modalités de la hiérarchisation des unités de peuplement (Bretagnolle, 2003 ; Guérin-Pace et Pumain, 1990) sont un bagage sérieux pour étudier de manière plus fine la distribution du peuplement sur l'ensemble du territoire cette fois-ci et sur le temps long. Ainsi, plus que ce processus propre aux dynamiques de population, qui ne sont pas le cœur même de nos recherches, ce sont bien leurs relations avec le développement du réseau qui nous intéressent au premier chef.

Le second ordre des critères de pertinence réside bien dans les relations à établir entre système de transport et système de peuplement. Les débats sur les effets structurants des infrastructures de transport ont longtemps traité d'une relation dans un seul sens. Il ne s'agit pas de l'occulter définitivement mais de l'analyser à la lumière du temps long. Confronté à la distribution

des populations, le développement du réseau peut obéir à une lecture oscillant entre l'efficacité territoriale, qui consiste à irriguer les zones les plus peuplées, et l'équité territoriale, qui consiste à irriguer l'ensemble du territoire, comme *« instrument du bien collectif »* (Dupuy, 1989).

Alors, il s'agit aussi de confronter les hiérarchies de nos deux objets de recherche, fruit d'une construction sociale dans un cas, d'une construction inductive dans l'autre. L'identification de rythmes, différenciés ou non, permet d'évaluer la force de la congruence ou la force d'une hiérarchie sur une autre. De manière plus guidée, il s'agit aussi de questionner le rôle de la structure urbaine préexistante (Pumain, 1982) dans le dessin du réseau ferroviaire. Alors, on inverse ici le sens de la causalité. L'identification de quelques grands principes de développement des réseaux, construite comme hypothèse, peut alors être confrontée de manière empirique à l'expérimentation, permettant d'évaluer leur degré de pertinence, dans le contexte dans lequel ces grands principes s'inscrivent.

Si notre positionnement géohistorique a permis une réflexion sur les échelles spatiales et temporelles et sur les nombreuses possibilités que l'on a à les combiner, les critères de pertinence permettent d'associer le processus de connaissances dans une *« exigence épistémologique de contextualisation »* (Pumain, 2014). Le sociologue Michel Grossetti, par ailleurs très sensible à la question géographique, propose la définition d'un *« espace à trois dimensions des phénomènes sociaux »* (Grossetti, 2011). Il combine de manière plutôt classique les échelles spatiales et temporelles, en rajoutant dans une troisième dimension, celle de la *« variété des contextes »*. Elle se caractérise alors par les théories d'arrière-plan. Un petit nombre de contextes implique une spécialisation des connaissances construites tandis qu'une pluralité des contextes participe à une montée en généralité *(Figure 2. 11)*.

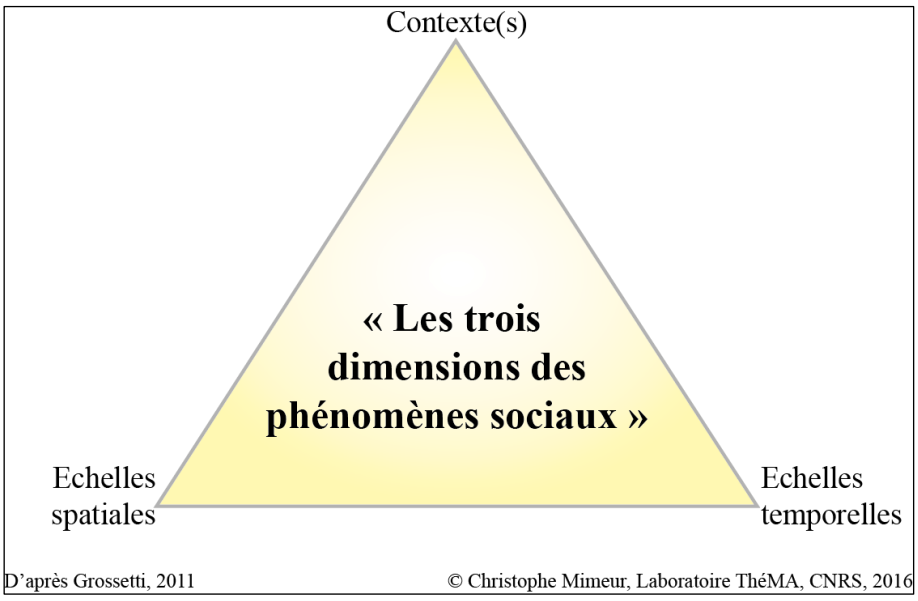


Figure 2. 11. Les trois dimensions des phénomènes sociaux

La définition de cadres d'exploration du lien entre réseau et territoire a permis un regard épistémologique sur l'acquisition de connaissances, appliquée au champ géohistorique. En s'insérant dans le permanent débat des effets structurants des infrastructures de transport, nous plaçons ce travail dans une logique constructiviste, parce que la définition des contextes dans lesquels nous investiguons cette question est fondamentale pour dépasser les simples cas d'études. Cette démarche suppose donc la pertinence des contextes : l'usage de la parcimonie n'est pas, selon nous, réductionniste quand nous explorons ce débat dans la perspective de la longue durée. Elle permet la formulation de questions à mi-chemin entre des hypothèses issues de théories d'arrière-plan et de nouvelles connaissances issues des traitements empiriques de l'information géohistorique. Enfin, cette démarche suppose des critères de précision, parce que le recours à la complexité nécessite l'utilisation de variables robustes, qui sont les porteurs d'une réalité, certes objectivée, mais qui tâchent d'englober de la manière la plus complète possible les composants du système de relations du couple réseau/territoire.

Conclusion

Ce deuxième chapitre a permis de montrer comment il est possible d'enrichir le débat des effets structurants des infrastructures de transport sur le développement des territoires par une réflexion sur les échelles d'analyse. L'espace est défini comme l'objet par excellence de la discipline géographique, en écho à l'émancipation progressive à laquelle elle a dû recourir au XIXème siècle vis-à-vis de l'histoire. Ainsi, l'objet de l'histoire, le temps, a été délaissé par le géographe. Longtemps considérés comme marginaux de la discipline, les géographes qui introduisent le temps ne le font souvent que dans une perspective rétrospective. Il faut chercher du côté de Fernand Braudel pour comprendre comment la géographie a pourtant intérêt à prêter attention à cette dimension : ce n'est plus le temps mais les temporalités qui participent à la construction de nouvelles réflexions. Malgré le succès des temporalités courtes, les temporalités longues sont encore timides dans les travaux des géographes français, *a fortiori* dans le champ de la géographie des transports. La géohistoire contemporaine tend à émerger sans que les chercheurs ne s'y réclament forcément.

La géographie a d'autant plus intérêt à introduire les temporalités dans ces analyses qu'elle se trouve au cœur du *« spatial turn »* qu'ont connu les sciences humaines et sociales depuis les années 1980, couplé au tournant technologique permis par l'informatique. La géographie connaît alors un renouveau, surtout dans les universités américaines, quand son outil largement répandu, le SIG, s'adapte à de nouvelles problématiques spatio-temporelles en devenant un SIG-H. Nous plaçons ainsi l'information géohistorique au cœur de notre travail de recherche, dans une démarche de projet de SIG-H. Les deux principaux challenges montrent comment la collecte et la digitalisation de l'information géohistorique prennent une part importante et comment la construction de nouvelles connaissances à partir de la première étape tend aujourd'hui à devenir prioritaire. C'est en tout cas dans ce second challenge émergent que nous situons notre positionnement épistémologique.

Conscients de la dimension mythique, au sens scientifique, que porte le débat sur les effets structurants, nous positionnons la construction de connaissances géohistoriques sur le lien entre réseau et territoire dans une démarche constructiviste, passant par la définition de critères de précision et de pertinence. Dans le premier cas, nous défendons une approche quantifiée du lien entre réseau/territoire où nous avons donné un statut à l'information géohistorique. Dans le second cas, nous défendons une approche par la science de la complexité, qui permet de définir les propriétés de nos objets et les relations qu'ils tissent dans un espace qui lie échelles spatiales, temporelles et contextes d'énonciation des théories d'arrière-plan.

Ce positionnement épistémologique, méthodologique et disciplinaire nous conduit à envisager un processus de modélisation des relations entre système de transport et système de peuplement. Il s'agit de revenir dans un premier temps sur la première étape du projet de SIG-H dans le *Chapitre 3* afin de structurer l'information géohistorique. Alors, nous verrons dans la deuxième partie de ce travail de recherche comment nous envisageons la modélisation des relations entre nos deux objets, qui nécessite un retour à la structuration des données.

Chapitre 3. LES TRACES DU LIEN RESEAU / TERRITOIRE PAR LA BASE DE DONNEES FRANCE

Introduction

Défini comme *« une représentation spatiale, ou un modèle, de données utilisées pour caractériser une portion de la surface terrestre »* (Frank, 1992), le Système d'Information Géographique est utilisé dans de nombreux domaines, des transports et de l'histoire comme nous l'avons déjà vu, mais aussi dans les domaines de l'environnement, de la sécurité civile. Quand il doit introduire des informations temporelles, l'information géohistorique prend place alors dans la matrice spatio-temporelle, obligeant à penser la conception des bases de données contenues à l'intérieur d'un SIG-H (Favory, Nuninger et Sanders, 2012). Il s'agit de la tâche du challenge traditionnel d'un projet de SIG-H, telle que nous l'avons envisagée dans le chapitre précédent. La constitution d'une telle base de données nécessite des moyens conséquents, davantage encore quand le travail est effectué à l'échelle nationale (Gregory et Ell, 2007) : nous verrons en quoi la démarche ne peut guère être automatisée et nécessite l'intervention d'opérateurs de saisie géomatique.

Dans ce cadre, la construction de ce projet doit s'appuyer sur les quelques grands principes dévolus à un SIG. Dans notre cas, nous retenons d'un SIG sa capacité à représenter l'espace à partir des trois primitives définies en géographie, parce que *« les agencements peuvent être discriminés en quelques grands types, en « espèces d'espace » génériques qui permettent d'indexer toutes les formations spatiales spécifiques observables »* (Lussault, 2007). Ces trois espèces trouvent une résonance particulière dans les objets de notre recherche : quand la population est répartie dans plus de 36 000 communes, quand le chemin de fer dessine des lignes sur le territoire français, et quand les gares sont des points d'entrée sur le réseau ferroviaire.

Ce chapitre présente comment la démarche de constitution d'un SIG-H, à l'intérieur duquel nous plaçons le projet FRANcE (French RAilway Network), est le résultat d'une démarche largement contrainte par les objets de recherche. Après que les objectifs de la recherche ont été définis dans le précédent chapitre, nous verrons que la qualité, la diversité et l'hétérogénéité de nos sources historiques sont liées à l'histoire (3.1.), celle des recensements de population d'une part, celle du déploiement du réseau d'autre part. Elle aboutit à une conceptualisation des données adaptée aux sources identifiées. La suite du chapitre permettra de voir comment la « révolution de la vitesse » est envisagée dans la base de données FRANcE. Enfin, nous verrons comment notre démarche permet de cerner la croissance du réseau à l'image de la théorie des réseaux, tout en pointant les premières limites d'un seul outil SIG-H dans le processus de construction de connaissances géohistoriques (3.2.).

3.1. L'information géohistorique face à l'hétérogénéité des sources

La démarche quantifiée du lien réseau/territoire a permis de définir des critères de précision. La population, c'est-à-dire la taille des villes – et des communes –, semble un bon proxy sur le temps long du développement de ces villes, lui-même corrélé au niveau de fonctions et de services disponibles. D'autre part, nous envisageons les dynamiques liées au développement du réseau par l'évolution de son déploiement sur le territoire, dans un premier temps. Il s'agit là d'une phase de collecte importante, qui a bénéficié de la constitution d'une équipe et de financements. Les principaux contributeurs sont la Fondation américaine pour les humanités (NEH) et l'Agence nationale de la recherche (ANR). Cette section vise à montrer quelles sont les méthodes d'implémentation du temps puis ensuite comment nous nous approprions la source historique, à travers l'exposé successif des sources de nos deux objets principaux.

3.1.1. **L'information démographique, un « classique mondial »**

Depuis le début du développement du SIG, de nombreux travaux se sont attachés à décrire l'évolution des limites des unités administratives, en lien avec des grandes bases de données statistiques (Gregory et Healey, 2007b). Il s'agit dans la plupart des cas de digitaliser les sources qui détiennent les recensements de population : le SIG national britannique est le résultat de l'exploitation de registres d'état civil et de données issues de la Poor Law du début du XIXème siècle jusqu'aux années 1970, combinée à des cartes anciennes et des sources textuelles pour les modifications des limites paroissiales, niveau le plus fin du découpage territorial au Royaume-Uni. De la même manière, le NHGIS américain propose des données de population à l'échelle du comté, car les municipalités, bien qu'à un niveau inférieur, ne couvrent pas l'ensemble du territoire. Les découpages américains reprennent ceux qui existent depuis le début des années 2000, et sont diffusés en ligne.

La démarche française est récente. Si pendant longtemps les études démographiques ont été réalisées au niveau national, parfois régional, mais surtout départemental ou cantonal, l'observation au niveau communal ne couvrait pas la totalité du territoire français jusqu'au début des années 2000 (Motte, Séguy et Théré, 2003). Regroupées au sein d'un dictionnaire intitulé *« Communes d'hier, communes d'aujourd'hui »*, ces observations sont le résultat de la compilation de trois sources principales :

- Les sources officielles : elles détiennent les arrêtés des découpages communaux et portent les informations sur la toponymie des communes françaises. En 1943, le Code Officiel Géographique a succédé à la nomenclature des communes de France créée en 1806 ;
- Les sources officielles secondaires : la publication des recensements de la population font l'objet d'un décret au Journal Officiel depuis 1801, servant d'authentification de la part du gouvernement ;

- Les sources secondaires : Claude Motte a ainsi utilisé des dictionnaires nationaux, des publications locales ainsi que des cartes historiques pour venir compléter les informations des deux sources précédentes.

Cette source a été utilisée dans le cadre du programme ANR Corpus *e-geopolis*. Ce projet d'envergure mondial avait pour but de constituer une base de données statistique sur les dynamiques démographiques à l'échelle mondiale. Ce travail porte néanmoins uniquement sur les aires urbaines de plus de 10 000 habitants depuis 1800. Le travail a été par la suite étendu à l'ensemble du territoire français. Soulignons ici les difficultés inhérentes à ce colossal travail de collecte à l'échelle des plus de 36 000 communes françaises, dont les auteurs de l'ouvrage sont pleinement conscients, mais qui n'altèrent qu'à la marge notre démarche :

- La fiabilité des recensements est stabilisée à partir de 1831 : auparavant, les données n'étaient officialisées qu'à un niveau agrégé alors que les données les plus fines restaient manuscrites ;
- Les trois guerres entre les XIXème et XXème siècles sont venues perturber le rythme des recensements, en modifiant les périodes intercensitaires ;
- Les auteurs relèvent près de 4 800 changements de frontières administratives, qui se caractérisent par la baisse du nombre de communes, de 40 609 en 1801 à 36 565 en 2010. Pour autant, les mêmes auteurs soulignent la globale stabilité du découpage communal.

La collecte de ces informations revient à s'emparer d'un problème récurrent en SIG, celui du Modifiable Area Unit Problem. Dans le cadre des recensements paroissiaux en Grande-Bretagne et pour prendre en compte les évolutions des découpages des paroisses, Ian Gregory et Paul Ell ont proposé une méthode d'interpolation spatiale qui permet d'attribuer de manière optimale la distribution des populations (Gregory et Ell, 2006). Mais dépourvus de cette information, les chercheurs ont privilégié l'association des populations issues de l'ouvrage de Claude Motte au chef-lieu communal. L'IGN propose le *Répertoire Géographique des Communes (RGC)* qui contiennent les toponymes, les coordonnées et l'altitude des chefs-lieux de communes définis par le bâtiment de la mairie (IGN, 2015). De plus, compte tenu de la stabilité des découpages communaux, les données ont été rapportées aux limites administratives définies dans la base de données *GEOFLA* (IGN, 2015). Dans le cas d'une imbrication de limites, une règle de répartition a été adoptée afin que la population de la commune supprimée soit distribuée proportionnellement aux autres communes concernées par la redistribution. D'après les estimations des auteurs, seulement 823 communes sont concernées, et ne concernent donc que 2 % de l'échantillon. Ainsi, l'information spatiale des populations françaises représentée dans le SIG-H est à la fois ponctuelle, sur le chef-lieu de commune, et aussi surfacique, adaptée aux limites administratives actuelles.

Implémenter l'information temporelle suppose en revanche l'arbitrage entre plusieurs méthodes. Initiées dans les années 1980 (Langran, 1992), les méthodes de représentations spatio-temporelles dans un SIG ont évolué autour de trois modèles. On reprend ici la *Figure 3. 1* proposée par Thomas Thévenin, qui rend compte de la manière de représenter l'avancée d'un front urbain. Compte tenu de l'absence d'information sur l'évolution des découpages, la troisième méthode n'est

pas réutilisable et n'est pas représentée ici. Si la première est évidemment réalisable, nous retenons le modèle spatio-temporel composite, parce que le volume des données collectées est optimisé, et parce que les requêtes spatio-temporelles associées peuvent être complexifiées. Ainsi, à partir du découpage actuel, il s'agit de remonter le temps : l'information de référence repose sur les 36 565 communes françaises, auquel on attribue ensuite les attributs de population de chaque recensement.

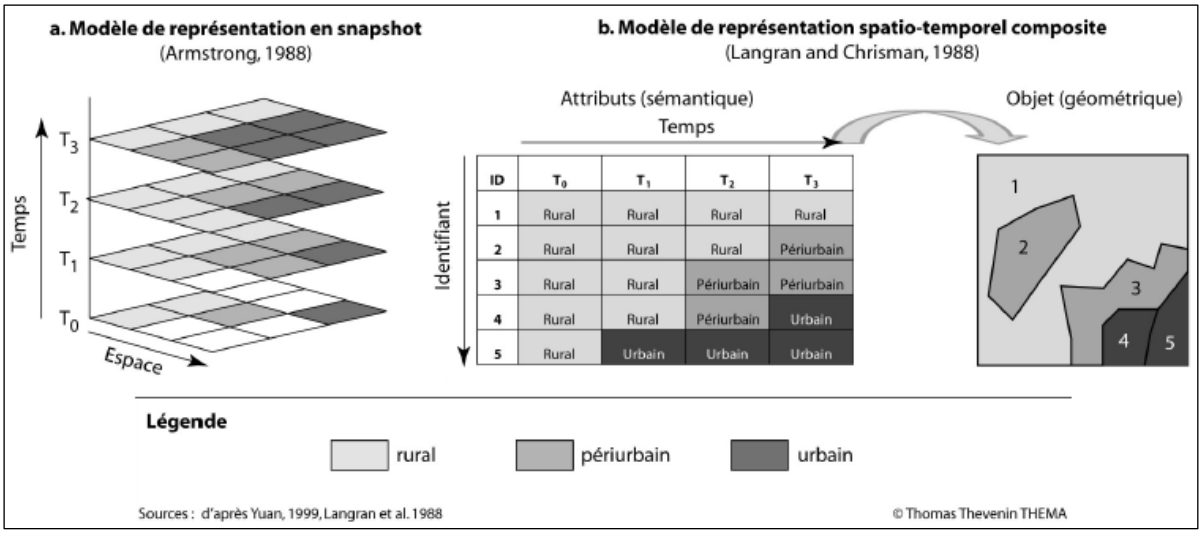


Figure 3. 1. Trois exemples de représentation spatio-temporelle dans un SIG

On propose maintenant d'exposer comment cette information démographique est implémentée dans la base de données FRANcE, en reprenant le modèle Pyramid (Peuquet, 2002) qui permet de reprendre d'une part la définition de l'objet, et d'autre part la donnée qui s'y rattache *(Figure 3. 2)*. La donnée est découpée selon la manière dont elle est localisée et en fonction de la représentation du temps. Cette structure constitue une *« carte d'identité »* de l'information géohistorique, qui sera réutilisée pour les autres objets de notre recherche. La question « Comment » permet de qualifier l'objet spatio-temporel. Dans le cas de l'information démographique, la donnée est numérique quand elle mobilise la source historique, elle est sémantique quand elle qualifie le statut administratif de chaque commune. Ainsi, si la collecte de l'information est communale, la structure hiérarchique de l'administration française permet l'agrégation d'informations à des niveaux supérieurs : dans un SIG-H, plus la granularité spatiale est fine, plus il est facile de traverser les échelles, il en est de même pour la granularité temporelle.

Ainsi, l'information démographique recueillie à l'échelle européenne est réinvestie ici dans notre perspective française. De la même manière, une initiative européenne a participé au lancement de la collecte de l'information ferroviaire, qui requiert une conceptualisation des données plus complexe, encore plus adaptée à l'objet.

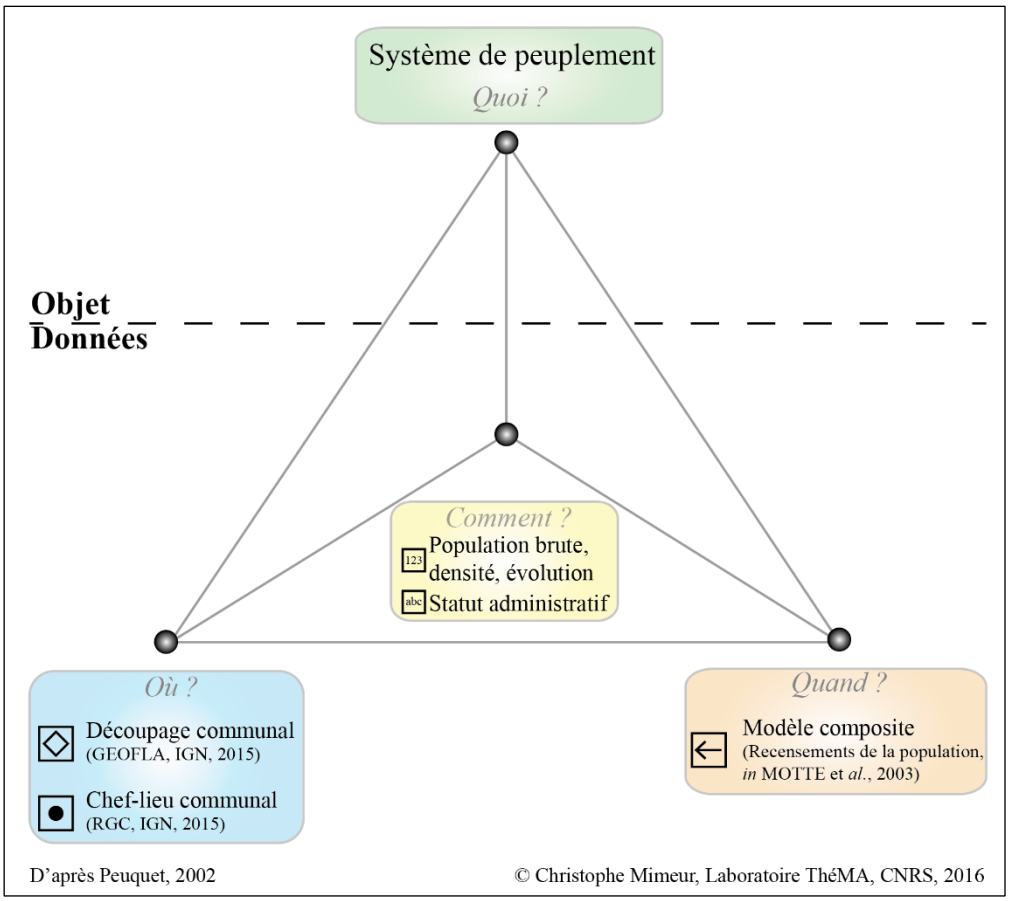


Figure 3. 2. L'information démographique dans la base de données FRANcE

3.1.2. **L'information ferroviaire, reflet de la complexité du système ferroviaire**

La collecte de l'information ferroviaire trouve sa place dans une initiative européenne, dont l'objectif est de numériser les infrastructures de transports depuis le XIXème siècle. Neuf pays participent au projet financé par la Fondation Européenne pour la Science (ESF : Inventing Europe). Les objectifs du projet, par la confrontation à des informations démographiques et économiques, définissent une granularité temporelle par décennie et l'utilisation du concept d'accessibilité. Dans le cas français, la base a été constituée par une équipe du laboratoire ThéMA. La collecte des informations en Grande-Bretagne a été permise par l'extraordinaire atlas historique de Michael H. Cobb : à l'intérieur d'une même source, le réseau bénéficie d'une grande précision spatiale, alors que lignes et gares sont renseignées par leurs dates d'ouverture et de fermeture (Morillas-Torné, 2012). Il n'existe pas en France une unique source dans laquelle nous pouvons collecter l'information compatible aux enjeux présentés dans notre positionnement. Nous posons l'hypothèse que cette difficulté est inhérente à la complexité du système ferroviaire français, depuis ses débuts.

Le travail de collecte est le résultat d'une combinaison entre sources primaires et sources secondaires. Pour définir la géométrie du réseau ferroviaire français, un fond cartographique de la SNCF a été acquis auprès de la Bibliothèque Nationale de France. Datant de 1944, il correspond à une période où la densité ferroviaire est presque maximum, car les premières suppressions de lignes ont commencé après 1936, mais se sont généralisées après-guerre. De la même manière que l'information démographique a été implémentée dans les limites administratives actuelles, l'information vectorielle de base du SIG-H est celle issue de Route500 (IGN, 2010). Pour autant, cette couche de base ne comporte que les 28 000 kilomètres actuels de réseau. Il a donc fallu en reconstituer autant à partir du document de 1944, au 1/800 000ème. Il s'agit là d'une limite inhérente au SIG-H, quand les cartes d'origine et les bases actuelles n'ont pas été levées à la même échelle (Knowles, 2005).

Par ailleurs, l'introduction de la dimension temporelle a requis l'utilisation de sources secondaires, souvent constituées par des passionnés. Parmi elles, le CD-ROM *« Histoire chronologique des chemins de fer de 1827 à 2000* », par les Editions Cima, a permis de renseigner quelles étaient les dates d'ouvertures et de fermetures des lignes et gares du réseau principal, à l'image de la source britannique. Par ailleurs, les précisions nécessaires pour le réseau local ont été renseignées grâce à la série de volumes *« Les Petits Trains de Jadis »*, publiée par José Banaudo et Henri Domengie. En 10 volumes, répondant à un découpage par grandes compagnies ferroviaires, ils retracent l'histoire de toutes ces lignes disparues, à l'aide de sources très variés mais précises. L'exemple de la *Figure 3. 3* montre comment à partir de la base Route500, de nouveaux tronçons ont été tracés dans la région périphérique de la gare du Creusot TGV aujourd'hui. Ainsi, conformément au scan de la SNCF et à la source locale des lignes, on voit comment la ligne à grande vitesse suit aujourd'hui une ancienne ligne ferroviaire locale : tunnels et viaducs caractérisent la ligne ancienne alors que de vastes déblais et remblais ont permis l'installation de la LGV.

La méthode par laquelle la temporalité est implémentée dans le SIG-H est en revanche différente de celle de l'information démographique : une simple approche par *snapshot* ou composite vient amoindrir la précision des sources. Travaillant sur le cadastre, Vrana propose d'autres approches (Vrana, 1989), parmi lesquelles nous retenons celle du *« data stamping »*, que l'on pourrait qualifier d'horodatage : les dates d'ouverture et de fermeture sont des attributs de chaque tronçon et de chaque gare *(Figure 3. 4)*. Quand une gare ou une ligne sont toujours existantes, la date de fermeture est codée « 2999 ». La technique de l'horodatage a été étendue dans le cas d'une réouverture de ligne, avec un champ spécifique, de même que dans le cas d'une électrification de la ligne.

Par ailleurs, la combinaison des sources a permis la diversification des attributs des tronçons. Outre la chronologie des ouvertures et de l'électrification, chaque tronçon est classé selon le type de voie : métrique, classique, grande vitesse. Le nom de l'opérateur lors de la mise en service est également renseigné, de même que le type de train qui circule, selon trois grandes périodes : 1831-1959, 1960-1980, 1981-2016. De leur côté, la nature des gares qualifiée par l'accueil de voyageurs et/ou de marchandises.

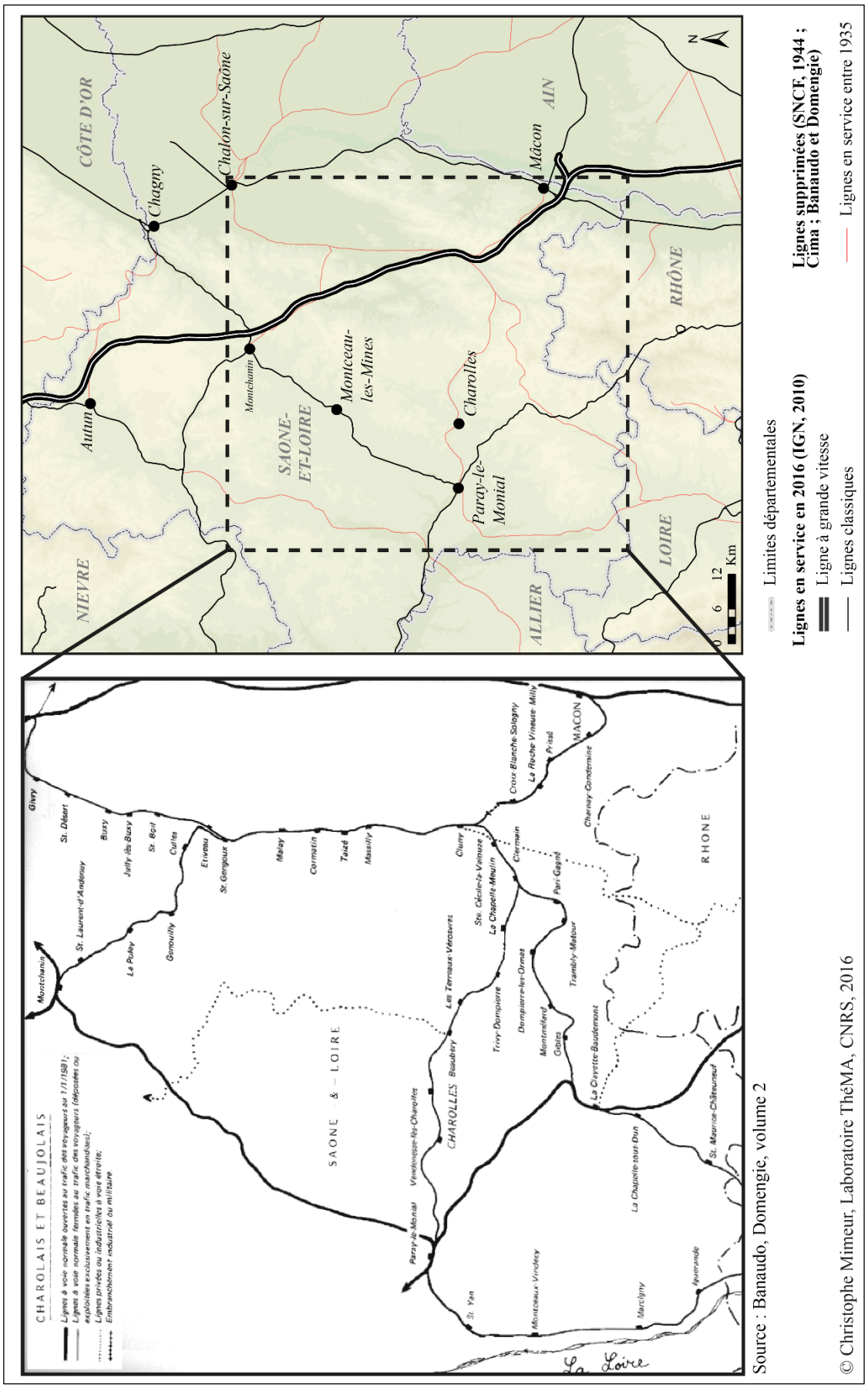


Figure 3. 3. Sources primaires et secondaires pour la spatialisation du réseau dans le SIG-H

a. La ligne

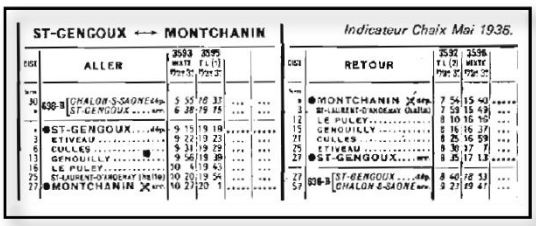
ST.GENGOUX – ÉTIVEAU – MONTCHANIN (25,8 km)

C'est dès 1872 que le Conseil-Général de Saône-&-Loire accorde une concession à MM. PARENT-PECHER et RICHE frères, pour une ligne de Roanne à Chalon-sur-Saône avec embranchement sur l'important centre ferroviaire et minier de Montchanin, dans le but d'assurer au secteur industriel de Montceau-les-Mines, Montchanin et Le Creusot un débouché vers la région lyonnaise. Après quelques études sur le terrain, le groupe financier, titulaire par ailleurs de concessions similaires dans l'est de la France, est déclaré en faillite et déchu de ses concessions en 1877. Les deux voies concernées sont alors incorporées au réseau P.L.M. par les conventions de 1883, et l'antenne de Montchanin ouverte au printemps 1889 par les soins du «grand réseau».

Après la libération, une voiture a été adjointe au train de marchandises jusqu'au 8 Mai 1948. Depuis lors, la desserte de Genouilly, alimentée par un trafic de bétail, de bois et de pierres concassées (chargées à Culles) s'est poursuivie en traction vapeur (140-B ou 040-TB de Chalon-sur-Saône) puis avec des engins diesel 030-DA (actuelles C-61000 à bielles), Y-51100 ou BB-63000, avant d'être supprimée au printemps 1969. Déclassée l'année

Ligne	Ouverture	Fermeture	Energie	Type de voie	Opérateur	Type de réseau
Saint-Gengoux - Montchanin	1889	1969	Non électrifié	1	PLM	Omnibus

b. La gare


ST-GENGOUX ⟷ MONTCHANIN — Indicateur Chalx Mai 1936.

ALLER / RETOUR

Gare	Ouverture	Fermeture	Nature
Saint-Gengoux	1889	1969	Voyageurs
Etiveau	1889	1969	Voyageurs
Culles	1889	1969	Voyageurs
Genouilly	1889	1943	Voyageurs
Le Puley	1889	1943	Voyageurs
Saint-Laurent-d'A	1889	1942	Voyageurs
Montchanin	1861	2999	Voyageurs / Fret

Source : Banaudo, Domengie, volume 2 © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Figure 3. 4. De la source au SIG-H : sources pour la chronologie du réseau

La compilation des différentes sources permet donc de répertorier non seulement l'ensemble des voies ferrées qui existent ou qui ont existé, mais aussi le type de voie *(Figure 3. 5)*. Ainsi, la base de données FRANcE montre comment le nombre de kilomètres de réseaux s'est accru en moyenne de 623 kilomètres chaque année entre 1840 et 1930, alors qu'il a diminué de 475 kilomètres chaque année entre 1930 et 2000. On voit également comment les ouvertures et les fermetures sont largement liées au cycle de vie des lignes. Si le rythme de la création de voies ferrées est plus élevé que le rythme de fermeture, on peut voir comment la fermeture des lignes à voie métrique et classique simple a contribué à la rétractation du réseau. Le réseau qualifié de « classique double » semble alors le plus stable depuis 1890, qui marque une rupture de la construction de ce réseau. Les voies métriques ont eu une heure de gloire provisoire, entre 1890 et 1960. Enfin, les premiers tronçons à grande vitesse inaugurés en 1981 apparaissent dans l'état des lieux de 1990, et connaissent une croissance depuis.

La représentation de l'information ferroviaire à l'aide du modèle Pyramid de Peuquet permet de résumer comment le réseau ferroviaire est l'association d'informations linéaires et ponctuelles, qui utilise la méthode de l'horodatage pour représenter la dynamique temporelle *(Figure 3. 6)*. Elle rassemble plus de 50 000 kilomètres de tronçons et plus de 10 000 gares. Pour autant, il faut envisager la manière dont laquelle le projet SIG-H peut permettre d'explorer les deux objets de recherche. Si un SIG permet de manière traditionnelle de combiner les différentes primitives de l'information géographique, la combinaison des différentes manières avec laquelle nous représentons le temps pose la question de l'harmonisation des objets.

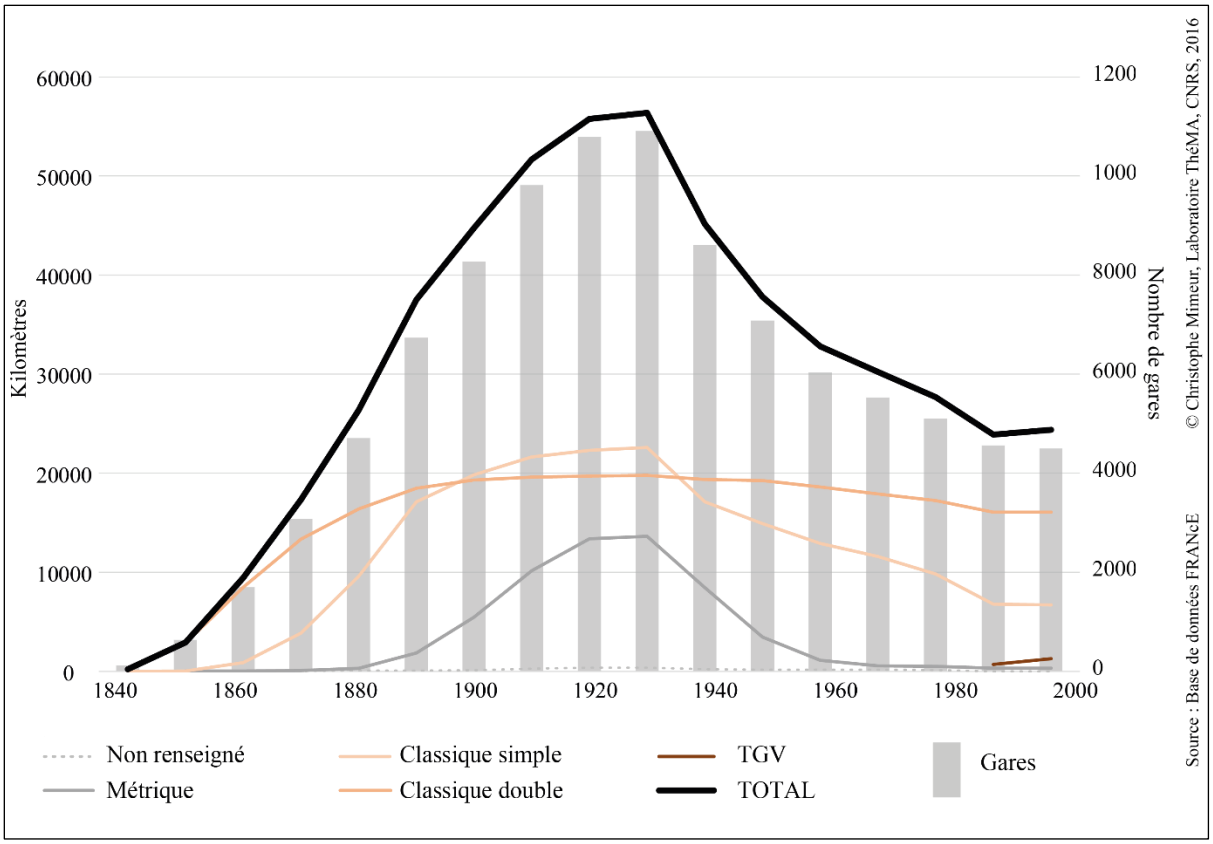


Figure 3. 5. L'évolution du réseau ferroviaire français en fonction du type de voie entre 1840 et 2000

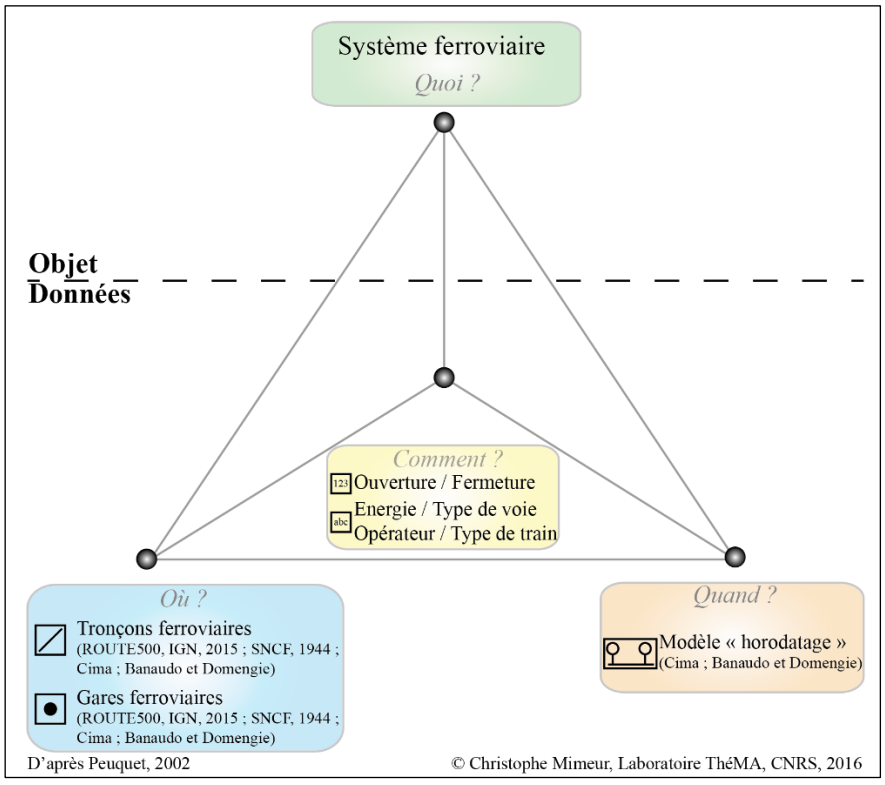


Figure 3. 6. L'information ferroviaire dans la base de données FRANcE

3.1.3. **La construction du modèle conceptuel PONT**

La collecte des informations géohistoriques a été conduite de manière séquentielle. Pour autant, notre questionnement de recherche requiert la mise en relation des deux objets collectés à l'intérieur du SIG-H. Dans ce type de projet, il semble désormais évident qu'il n'y a pas de démarche intégrée permettant la création d'un modèle conceptuel de données génériques (Gregory et Ell, 2007). Pour autant, dans le challenge émergent de l'analyse et de la production de nouvelles connaissances, la conceptualisation des relations entre nos objets requiert un dialogue entre informaticiens et géographes. Dans ce cadre, il s'agit de lier l'ensemble des éléments issus des modèles Pyramid des deux objets de notre recherche. La conceptualisation d'un modèle de données est alors *« une représentation graphique [qui] doit être réalisée avec un formalisme simple et clair, ayant une grande richesse d'expression »* et qui est accompagnée d'une *« description textuelle »* (Proulx, Larrivée et Bédard, 2002). Ce formalisme a fait l'objet de nombreuses contributions (O'Sullivan, 2005 ; Parent et al., 1997), dans lequel le langage UML (Unified Modeling Language) tient une place importante, car il répond à la simplicité énoncée ci-dessus et au dialogue avec les sciences informatiques. Il entre dans le courant de la géomatique, définie comme *« une discipline ayant pour objet la gestion des données à référence spatiale et qui fait appel aux sciences et aux technologies reliées à leur acquisition, leur stockage, leur traitement et leur diffusion »* (Office de la langue française, 1993).

La construction d'un modèle conceptuel de données a pour objectif de décrire les données dans un système d'information de manière formelle. Elle suppose la définition des objets à travers un langage standardisé, mais suppose également la définition des relations que les objets entretiennent. Dans notre cas, il s'agit d'établir les relations à l'intérieur de chaque objet mais aussi entre les objets. De manière traditionnelle, à l'intérieur d'un SIG, les relations entre les entités sont envisagées de manière spatiale, en reprenant les trois primitives de la géographie. Malgré l'importance des temporalités dans la structure de nos objets, le réflexe spatial des modèles conceptuels de données en SIG nous pousse à conserver cette entrée, qui nécessite néanmoins de déconstruire les représentations Pyramid isolées de chaque objet. Notre démarche géohistorique nous conduit à adopter une seconde entrée thématique : le réflexe serait alors d'isoler d'un côté les informations relatives au réseau, de l'autre les informations relatives au territoire. Pour autant, nous faisons ici l'hypothèse que cette structure double pourrait devenir restrictive et nous adoptons une vision évolutive de notre base de données. En effet, sur la longue durée, le système de transport ferroviaire est un élément parmi d'autres des dynamiques de réseaux : d'autres réseaux, navigables, routiers coexistent avec le réseau ferroviaire. De la même manière, nous posons l'hypothèse que, si la taille de communes semble *« le meilleur »* (Pumain, 1997) indicateur de la dynamique d'un territoire sur le temps long, d'autres variables sont également des marqueurs des évolutions économiques, sociales voire politiques des territoires.

La suite de la sous-section montre la formalisation du modèle conceptuel de données PONT (PeOple Network Territory). Il s'agit d'abord de définir la *« grammaire »* qui permet de qualifier les objets et leurs relations *(Table 3. 1)*. La symbolique est en partie adaptée du formalisme UML

développé dans un logiciel, nomme Perceptory (Bédard et al., 2004 ; Parent et al., 1997 ; Proulx, Larrivée et Bédard, 2002). Dans ce langage, nous ajoutons deux symboles, qui permettent de rendre compte du temps dans la qualification des objets. De cette sorte, nous qualifions dans le modèle conceptuel de données la méthode d'implémentation du temps : l'horodatage, quand les tronçons sont définis par une date d'ouverture (début) et une date de fermeture (fin) ; l'approche composite, quand la nature des données suggère la prise en compte de clichés à intervalles réguliers, comme les recensements décennaux par exemple.

		Représentation	Concept	Définition
Objets	Espace		Polygone	Objet surfacique
			Ligne	Objet linéaire
			Point	Objet ponctuel
	Temps		*Composite*	*L'information temporelle est un attribut, à partir d'une couche de référence.*
			Horodatage	*Date de début, date de fin (ou 2999)*
	Attribut		Numérique	Attribut numérique
			Sémantique	Attribut sémantique
Relation			Association	Lien entre deux ou plusieurs objets
			Agrégation	Association dans laquelle une des deux extrêmités joue un rôle prépondérant sur l'autre
			Composition	Association dans laquelle une extrêmité compose l'autre extrêmité
			Spécialisation	Association dans laquelle un élément est une sous-catégorie d'un élément plus général
			Généralisation	Association dans laquelle plusieurs éléments forment un élément plus général

D'après Parent et al., 1997, Proulx et al., 2002 © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 3. 1. La grammaire UML dans le modèle PONT

La seconde figure expose l'agencement des objets et informations *(Figure 3. 7)*. La suite du texte facilite la lecture des enjeux d'une telle formalisation. Ainsi, fidèles aux principes et aux opportunités laissées par la géohistoire, nous isolons les découpages spatiaux comme une catégorie d'objets. Elle est la base des relations entre nos objets à ce stade, mais nous verrons plus tard de quelle manière on peut envisager d'autres découpages. Celui envisagé ici est encore une fois guidé par les objets et les représentations que l'on a de nos objets de recherche. La hiérarchie de l'administration française est alors une catégorisation classique en SIG : les relations sont alors des

agrégations et des compositions. Il s'agit d'un couplage fort que l'on peut verbaliser : un département *a* plusieurs communes, et plusieurs communes *composent* un département.

Par ailleurs, l'information démographique est alors la plus simple à formaliser parce qu'elle est unique dans cette grande catégorie du système de peuplement, du moins pour l'instant. La population brute est alors rattachée à chaque chef-lieu de commune et peut facilement être agrégée au canton, au département ou à la région.

Le système de transport, ici appréhendé par le réseau ferroviaire, est constitué des lignes et de gares. Les lignes sont composées de différents segments, ici nommés arcs, dont la jonction s'effectue grâce à des points. On parle alors d'association entre le point de départ et le point de fin de chaque segment d'une part et un nœud du réseau ferré. Parmi ces nœuds, une catégorie est dite « spéciale » parce qu'elle renferme les gares. Le lien se fait avec les découpages spatiaux par composition : si elle existe, alors la commune a une gare. De manière imbriquée, ce lien vaut aussi pour les niveaux agrégés des découpages.

Finalement, on voit comment les découpages font ici le pont entre nos deux objets de recherche. Alors qu'ils sont au cœur des problématiques géohistoriques, nous adoptons donc ce modèle PONT (PeOple Network Territory), dans la mesure où une classe d'objets particulière permet la liaison entre le système de transport et le système de peuplement. Par ailleurs, cette structure est évolutive parce que sa construction permet l'ajout ultérieur d'autres objets permettant de qualifier les réseaux, d'autres permettant de qualifier le système de peuplement. Aussi, l'isolement des découpages dans une seule classe d'objet peut permettre d'en envisager d'autres, qui devront toutefois pouvoir lier les deux autres classes d'objet. Une représentation graphique simple, assortie de la description que nous faisons d'en faire, fait du modèle PONT un modèle conceptuel de données pour la base de données FRANcE, répondant aux critères d'objectivation, associant les contraintes de l'information géohistorique. Pour autant, on peut déplorer ici les rares investigations qui visent à introduire de manière plus globale la conceptualisation du temps dans les modèles conceptuels de données SIG (Gregory et Healey, 2007b ; Knowles, 2005 ; O'Sullivan, 2005). Cette remarque est d'autant plus vraie que la valeur ajoutée de la base de données FRANcE réside sans doute dans la prise en compte généralisée des vitesses le long du réseau ferroviaire français, qui nécessite de penser à son tour la structuration de cette nouvelle information dans la base de données.

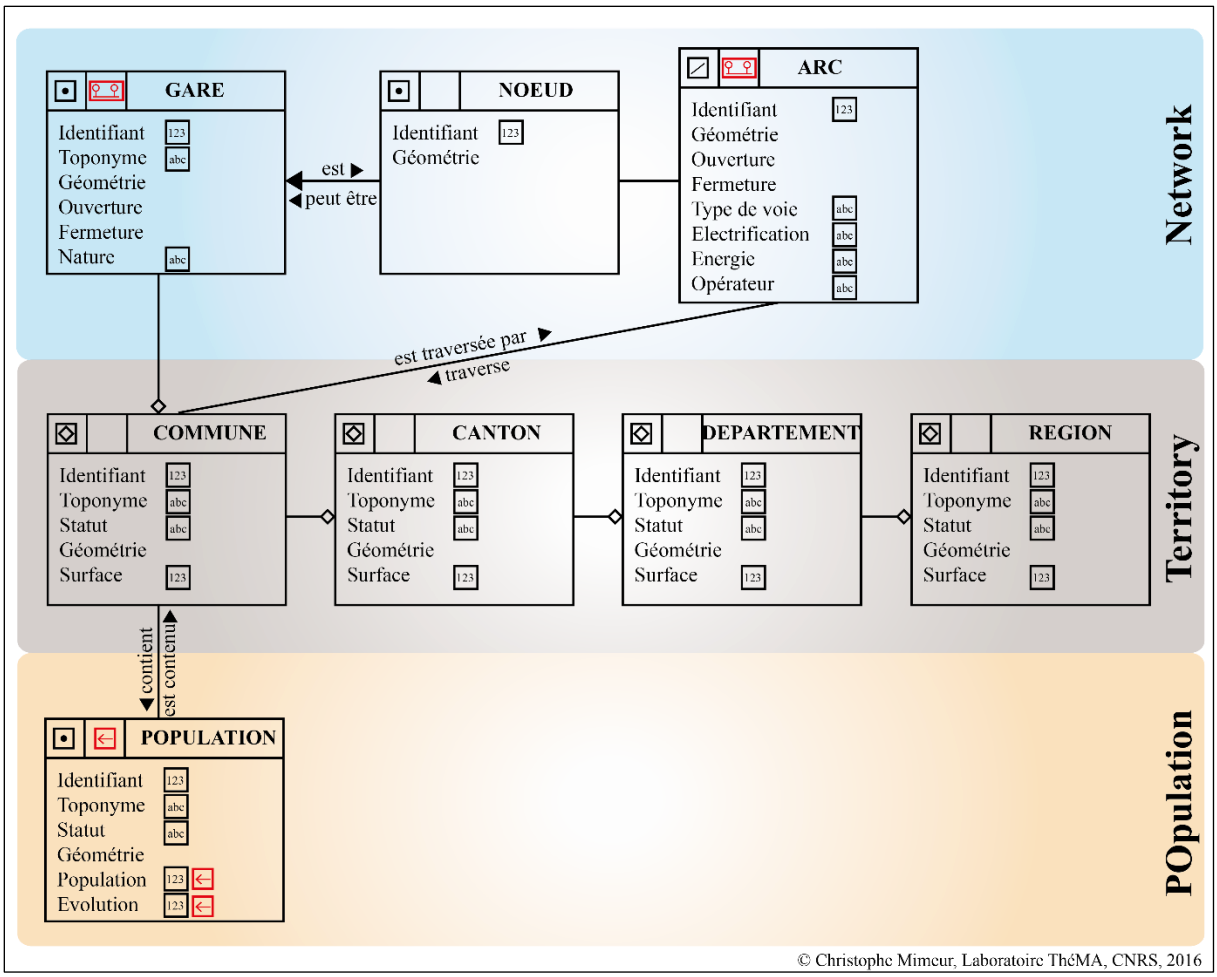


Figure 3. 7. Le modèle PONT dans la base de données FRANcE

3.1.4. **La vitesse dans FRANcE**

Au commencement du réseau, la vitesse et la performance ne sont pas recherchées (Studeny, 1995). L'objectif est alors de cheminer et d'élargir l'horizon : *« l'intérêt pour de telles machines [la triade vapeur-fer-charbon], c'est-à-dire pour une vitesse mécanique, reste très marginal au XIXème siècle »* (Studeny, 1995). Nous posons aussi l'hypothèse que l'irruption ferroviaire est une rupture par rapport aux diligences d'autrefois, parce que *« la voie ferrée supprime souvent les fonctions d'étapes, auberges, hôtels, et métiers du cheval »* (Studeny, 1995). Toutefois, notre définition d'une théorie des réseaux a montré comment le réseau est lui-même générateur de sa structure, sa dynamique et ses lois d'évolution. Il s'agit donc de qualifier ces trois états de fait, en complexifiant l'information géohistorique à travers l'introduction des vitesses de déplacement.

L'introduction de la vitesse dans le SIG-H répond aux enjeux développés dans le chapitre précédent. L'objectif est double : il s'agit de rendre compte de la mise en place progressive d'une hiérarchie du réseau à travers les vitesses, qu'il s'agira de rapprocher du système de peuplement,

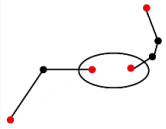
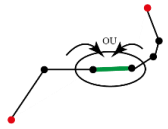
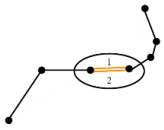
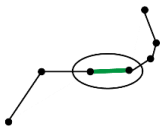
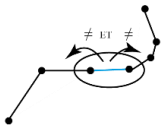
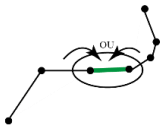
mais il s'agit surtout de rendre compte des temps de parcours nécessaires aux nouveaux usagers du train. La dimension temporelle des réseaux de transport est alors double : comment se propagent-ils sur le temps long ? Comment permettent-ils de se déplacer sur le temps court ? Le recours à la notion d'accessibilité devient central. Elle a d'abord été définie par Walter G. Hansen en 1959 comme un *« potentiel d'opportunités d'interactions »* (Hansen, 1959). Elle a été plus tard précisée dans le champ de la géographie des transports à travers deux définitions complémentaires :

- *« rendre compte de la possibilité et de l'effort à consentir pour parcourir l'espace, dans le but d'atteindre un lieu, une ressource ou un équipement »* (Kwan, Janelle et Goodchild, 2003) ;
- *« mesure du système de transport permettant à une personne ou une marchandise d'atteindre une activité ou une destination au moyen d'un mode ou d'une combinaison de modes de transport »* (Geurs et Ritsema von Eck, 2001).

Cette mesure, qui rentre dans le positionnement quantifiée de ce travail, est communément utilisée dans la géographie de transports, et passe par la mobilisation de la théorie des graphes, *« un outil particulièrement utilisé pour les représenter et pour résoudre un certain nombre de problèmes classiques comme les chemins les plus rapides entre une ou plusieurs origines et destinations, la capacité d'un réseau »* (Mathis, 2003). Porteur d'un paradoxe entre simplicité de la représentation et complexité des relations à établir, le recours au graphe dans la base de données FRANcE pose deux contraintes : la première est d'avoir un réseau connexe – où il *« existe un chemin […] pour aller de tout point à tout autre point »* (Lévy et Lussault, 2003). La deuxième est d'avoir une couverture totale de l'information sur la vitesse sur l'ensemble du réseau de transport.

L'exigence d'obtenir un graphe connexe nous confronte une nouvelle fois aux enjeux d'un projet SIG-H, où l'information peut-être incomplète. Il faut donc imaginer un dispositif de correction des éventuels décalages. Ainsi, si des outils intégrés au SIG permettent de corriger ces erreurs topologiques, le recours aux corrections manuelles a souvent été préféré, d'autant qu'il faut garder à l'esprit qu'un tronçon n'est pas toujours en service tout au long de notre période d'étude. Nous retenons ici trois types d'erreurs topologiques les plus rencontrés dans notre travail. La *Table 3. 2* détaille le type d'erreur, sa recherche, sa description et la solution apportée pour garantir la connexité du réseau dans son ensemble. La méthode de correction est semi-automatique, venant confirmer que l'intervention manuelle du géomaticien est encore majoritaire dans un projet géohistorique.

Si cette première étape n'entrave pas la modélisation conceptuelle des données, le cas de l'enrichissement de FRANcE par les vitesses est différent. Là encore confrontés aux enjeux d'un projet SIG-H, nous devons avoir recours à de nouvelles sources secondaires, en ayant en tête la première exigence, celle d'une couverture complète et connexe du réseau. La meilleure des solutions aurait été de mobiliser les guides horaires de passages de train pour l'ensemble de notre période, comme dans le cas britannique (Crafts et Leunig, 2005). Pour autant, cette information n'est pas disponible dans la granularité spatiale et temporelle que nous avons choisie. Dès lors, la démarche défendue ici est qualifiée de discrète et est résumée en trois étapes.

Défaut de topologie	Description du défaut		Recherche du défaut	Réparation du défaut	
Tronçon manquant	Une ligne est scindée en deux alors qu'elle est bien une seule et même ligne		Identification des "terminus" et recherche d'un autre point dans un voisinage immédiat Investigations visuelles	Création manuelle d'un tronçon Rapatriement des attributs des tronçons contigus	
Tronçons superposés	Deux lignes se superposent et ont les mêmes attributs		Recherche des tronçons ayant la même géométrie et les mêmes attributs	Suppression d'un des deux tronçons	
Attribut erroné	Certains attributs sont aberrants		Recherche des tronçons dont les attributs sont différents de ceux des tronçons contigüs	Correction manuelle des attributs en fonction de la situation du tronçon sur le réseau	

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 3. 2. La correction des erreurs de topologie dans FRANcE

L'information sur la vitesse ferroviaire ne peut répondre à la méthode de l'horodatage utilisée pour digitaliser le chemin de fer. Par son caractère discret, la vitesse est alors implémentée selon la méthode composite, par décennie. Trois sources principales permettent de renseigner les vitesses sur trois grandes périodes de la construction du réseau ferroviaire : la construction initiale du réseau (Caron, 1997 ; Studeny, 1995), la phase d'électrification des principales lignes (Blier, 1996 ; Bretagnolle, 2003) puis la construction des lignes à grande vitesse. Par croisement avec le type de voie recensé dans la section précédente, différentes classes de lignes ont été définies en fonction des périodes *(Figure 3. 8. a.)*.

La seconde étape passe par la prise en compte des contraintes topographiques pour affiner les vitesses. Ainsi, la calibration a été opérée grâce à trois sources à trois dates différentes : le Guide horaire de l'almanach Hachette de 1908, la carte des vitesses de Wiener de 1934 et la carte actuelle fournie par SNCF Réseau. La troisième étape consiste alors à obtenir un graphique de l'évolution des vitesses moyennes sur l'ensemble de la période qui permet de renseigner et de couvrir l'ensemble du réseau *(Figure 3. 8. b.)*.

On propose désormais d'évaluer l'estimation de vitesses obtenues en les confrontant d'une part à la source secondaire de Christophe Studeny sur de grandes lignes en 1870. Puis ensuite, nous confrontons un deuxième exemple à l'échelle très locale, en 1930. Les figures suivantes n'ont qu'une valeur illustrative, mais permettent de voir comment la généralisation des vitesses permet tout de même de se rapprocher de la réalité des déplacements. Il ressort d'abord des grands parcours une surestimation de nos données *(Table 3. 3)*. Pour autant, les valeurs de Studeny sont issues des trains les plus rapides, alors que notre base de données FRANcE ne permet pas de différencier les

différents types de service sur une même ligne, que ce soient des express, des rapides ou de simples omnibus. Il faut souligner tout de même que les ordres de grandeur sont conservés, venant confirmer la validité de notre approche par les vitesses.

Une étude de la grille horaire de la *Figure 3.* 4 montre la diversité des services. Si la ligne voit le passage de trois omnibus quotidiens en une heure environ jusque dans les années 1930, la ligne Montchanin-Saint-Gengoux-le-National ne comporte plus que deux circulations par jour : un train léger omnibus et un train mixte marchandises-voyageurs, ce qui explique des arrêts parfois plus longs dans certaines gares. Ainsi, nous avons pris en compte une fourchette de temps de parcours entre chaque gare, selon les deux types de trains qui circulent en 1936. Nous observons alors une proximité entre la fiche horaire de la ligne Montchanin-Saint-Gengoux et les valeurs obtenues grâce à la base de données FRANcE *(Table 3. 3)*. Cet exemple n'a bien sûr qu'une visée illustrative. Les grilles horaires des chemins de fer au XIXème siècle sont contenues dans l'indicateur Chaix, du nom du propriétaire de l'imprimerie en charge de l'impression des grilles horaires des grandes compagnies. Cette importante source est en revanche largement inégale et hétérogène, même si sa richesse participe à la géohistoire du réseau ferroviaire, nous aurons l'occasion d'y revenir.

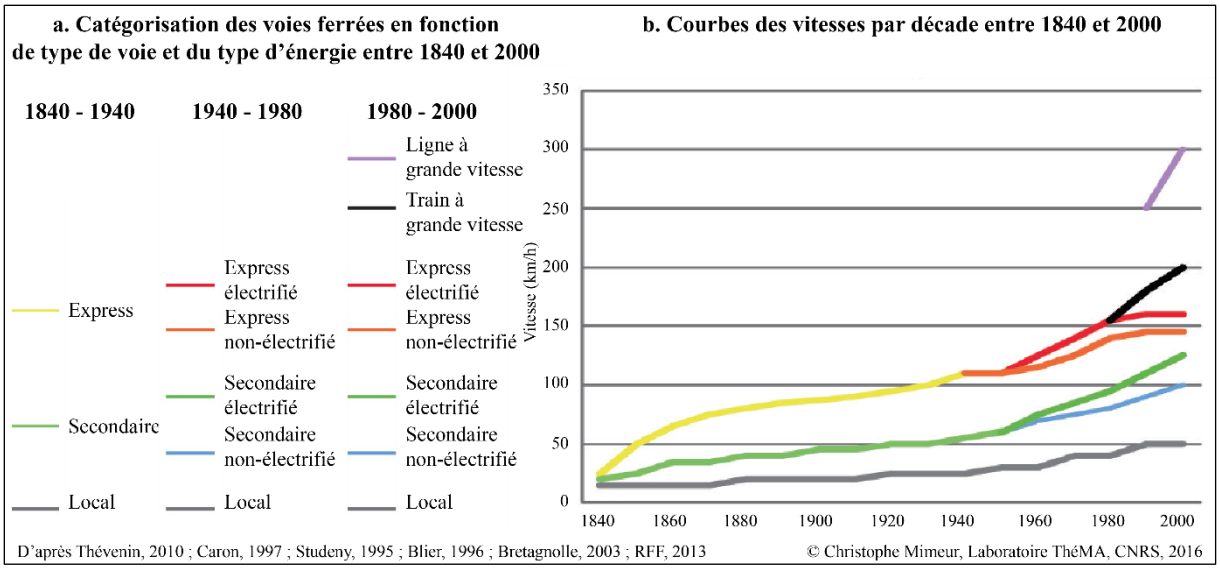


Figure 3. 8. La vitesse dans la base de données FRANcE

En revanche, la confrontation avec les valeurs obtenues par Christophe Studeny participe davantage à la phase de validation pour l'ensemble du réseau et l'ensemble de la période d'étude. Elle vient alors complexifier le modèle PONT par la combinaison de l'implémentation temporelle *(Figure 3. 9)*. Elle permet aussi d'établir une nouvelle classification des voies ferrées, non plus uniquement de voies, mais aussi par catégories de vitesse. Alors, pour chaque décennie, une classification évolutive rend compte des lignes rapides, intermédiaires et lentes. La base de données

FRANcE possède une information inédite à de telles échelles spatiale et temporelle, ce qui en fait son principal intérêt, alors que l'implémentation temporelle n'apporte pas ici de nouvelles avancées.

Grandes lignes en 1870	Temps de parcours en heures	
	Studeny	Base de données FRANcE
Paris-Dijon	6	7
Paris-Lille	4,5	5,5
Paris-Rennes	8	10
Paris-Bordeaux	11	14
Paris-Strasbourg	10	11
Paris-Lyon	9	11

Ligne Montchanin-Saint-Gengoux en 1936	Temps de parcours en minutes	
	Temps intergare d'après 4 services	Base de données FRANcE
Montchanin		
Saint-Laurent-d'Andenay	Entre 5 et 9 minutes	7
Le Puley	Entre 9 et 16 minutes	10
Genouilly	Entre 4 et 7 minutes	4
Culles	Entre 9 et 22 minutes	10
Etiveau	Entre 5 et 9 minutes	6
Saint-Gengoux	Entre 4 et 7 minutes	7

Source : Base de données FRANcE
D'après Studeny, 2009 ; Banaudo et Domengie

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 3. 3. Comparaison des temps de parcours entre les sources secondaires et la base de données FRANcE

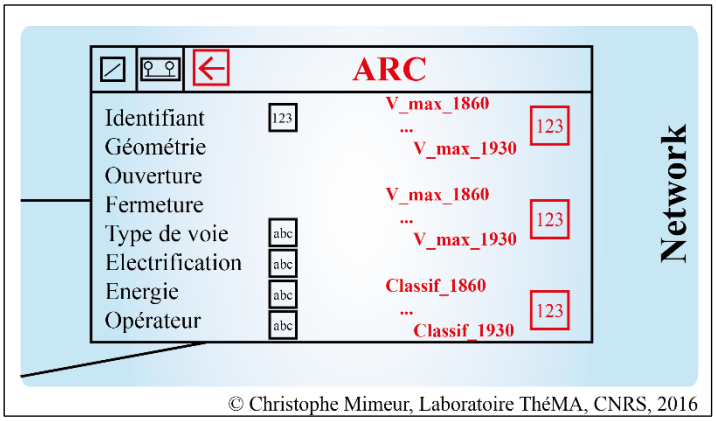


Figure 3. 9. La vitesse dans le modèle PONT

Ainsi, cette section a mis l'accent sur l'articulation des questionnements thématiques mais surtout méthodologiques, inhérents au processus de collecte dans un projet de SIG-H. Cette étape est sans doute la plus grosse consommatrice de ressources : les sources historiques, leur expertise, le passage à la formalisation dans le langage propre au SIG, la digitalisation. Elle jette ainsi les premières pierres de la nécessaire collaboration entre thématiciens et informaticiens. Dans le cadre de la constitution de la base de données FRANcE, la phase de collecte a bénéficié du soutien d'institutions pour la recherche à l'échelle française et européenne. Aussi, nous avons vu comment nous adaptons un modèle conceptuel de données PONT, où le système de transport et le système de peuplement sont liés par les découpages spatiaux. Cette séparation en trois classes d'objets est évolutive, parce qu'elle permet de complexifier l'information sans l'altérer. De nouvelles

investigations permettent d'enrichir l'information ferroviaire pour appréhender la *« révolution de la vitesse »* (Studeny, 1995). Il s'agit alors désormais d'étudier dans quelles mesures le SIG-H permet de revisiter la phase de croissance du réseau, et comment les limites de cet instrument apparaissent.

3.2. De l'exploitation à la complexification de l'information

Même si le SIG-H semble porter en son sein une capacité de dialogue entre les différentes disciplines, pour croiser des informations donnant naissance à de nouvelles connaissances, le cadrage géohistorique élargit les possibilités d'analyses et d'ajout de nouvelles informations. Ainsi, il s'agit maintenant d'étudier comment la base de données FRANcE peut retracer de manière à la fois empirique et théorique la croissance du réseau ferroviaire à partir de premiers traitements à l'intérieur du SIG-H. Pour autant, les perspectives de complexification de l'information pointent les limites de cet outil, si on le considère de manière isolé. Il s'agit de qualifier plus précisément la phase de développement du réseau depuis les premières traverses posées au début du XIXème siècle jusqu'à nos jours. Notre approche quantifiée s'inscrit dans la théorie des réseaux, appliquée au développement du réseau ferré français : *« l'idée d'une aptitude intrinsèque du réseau à croître sans soucis de quelconques obstacles »* (Dupuy, 1987a) est alors confrontée au(x) contexte(s) dans lequel la croissance du réseau s'inscrit. Nous projetons alors ici d'analyser la phase de naissance du réseau ferroviaire et la phase de croissance du réseau ferroviaire français, à rapprocher de la théorie des cycles de produits adaptée à la théorie des réseaux *(cf. 1.1.3)*.

3.2.1. La mise en connexité du réseau ferroviaire dans la base de données FRANcE

Les premières lignes sont surtout le fruit d'initiatives ponctuelles et privées. Inaugurée en 1827, la ligne entre Saint-Etienne et Andrézieux est concédée sur une longueur de 18 kilomètres à la « Compagnie du Chemin de Fer de Saint-Etienne à la Loire », principalement pour les besoins d'exploitation de la houille. Considérée comme la première ligne de chemin de fer de l'Europe occidentale, la traction est hippomobile jusqu'en 1844 et se substitue à l'arrivée des premières locomotives (Caron, 1997). Cinq ans plus tard, le prolongement de la ligne jusqu'à Lyon marque les premiers projets d'ouvrages d'art nécessaires pour poser la voie ferrée sur une route préexistante, dédiée à la fois au transport des voyageurs et à celui des marchandises. Pour autant, le retard français par rapport au Royaume-Uni et à l'Allemagne n'est encore pas rattrapé avec l'ouverture de la première ligne de banlieue par les frères Pereire et Eugène Flachat entre Paris-Saint-Lazare et Saint-Germain-en-Laye. Son succès, avec près de 18 000 voyageurs le premier jour, marque selon François Caron le début d'une prise de conscience qui aboutira en 1842 à la « Loi relative à l'établissement des grandes lignes de chemin de fer en France », corrigée en 1844.

Elle édicte la *« cession des terrains par les communes, construction par l'Etat et exploitation par des compagnies – fortune générale, fortune locale, fortune privée –, tels sont les trois éléments mis en jeu pour arriver à la réalisation »*. Elle prend le nom du directeur général des Ponts et Chaussées de l'époque : l'**« Etoile de Legrand »** est largement centrée sur Paris *(Figure 3. 10)*. Loin de rattraper son retard rapidement, *« en 1850 encore la France offre le spectacle d'un réseau éclaté, en lignes éparpillées, dispersées, sans cohérence »* (Studeny, 1995). Il faut attendre la fin de la décennie 1850 pour obtenir un réseau connexe. En 1842, on peut ainsi identifier sur la figure huit sous-réseaux, qui persistent encore en 1850. C'est d'abord à proximité de Paris que le réseau a grandi d'une part, de même que le long de la vallée de la Saône que du Rhône, réduisant le nombre de sous-réseaux à trois. Ce n'est qu'après 1855 que l'ensemble s'est lié, alors que la transversale entre l'océan Atlantique et la mer Méditerranée a été la dernière dessinée dans le projet de 1842. La finalisation de ce projet montre un décalage entre l'établissement de la loi et sa concrétisation presque vingt ans plus tard *(Table 3. 4)*. Ainsi, en 1859, les conventions entérinent l'oligopole à six têtes autour des six grandes compagnies *(Figure 3. 11)* qui resteront jusqu'à la nationalisation de 1937.

	1842			1850		
	Longueur du réseau	Vitesse moyenne	*Taux de croissance annuel moyen*	Longueur du réseau	Vitesse moyenne	*Taux de croissance annuel moyen*
Voies lentes	77	19	-	148	19	*8,51*
Voies intermédiaires	367	27	-	1507	27	*19,31*
Voies rapides	181	34	-	1267	35	*27,54*

	1855			1860		
	Longueur du réseau	Vitesse moyenne	*Taux de croissance annuel moyen*	Longueur du réseau	Vitesse moyenne	*Taux de croissance annuel moyen*
Voies lentes	192	24	*5,34*	933	24	*37,19*
Voies intermédiaires	3043	33	*15,09*	5094	33	*10,85*
Voies rapides	2474	43	*14,32*	3378	43	*6,43*

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 3. 4. La mise en connexité du réseau par classes de vitesses entre 1842 et 1860

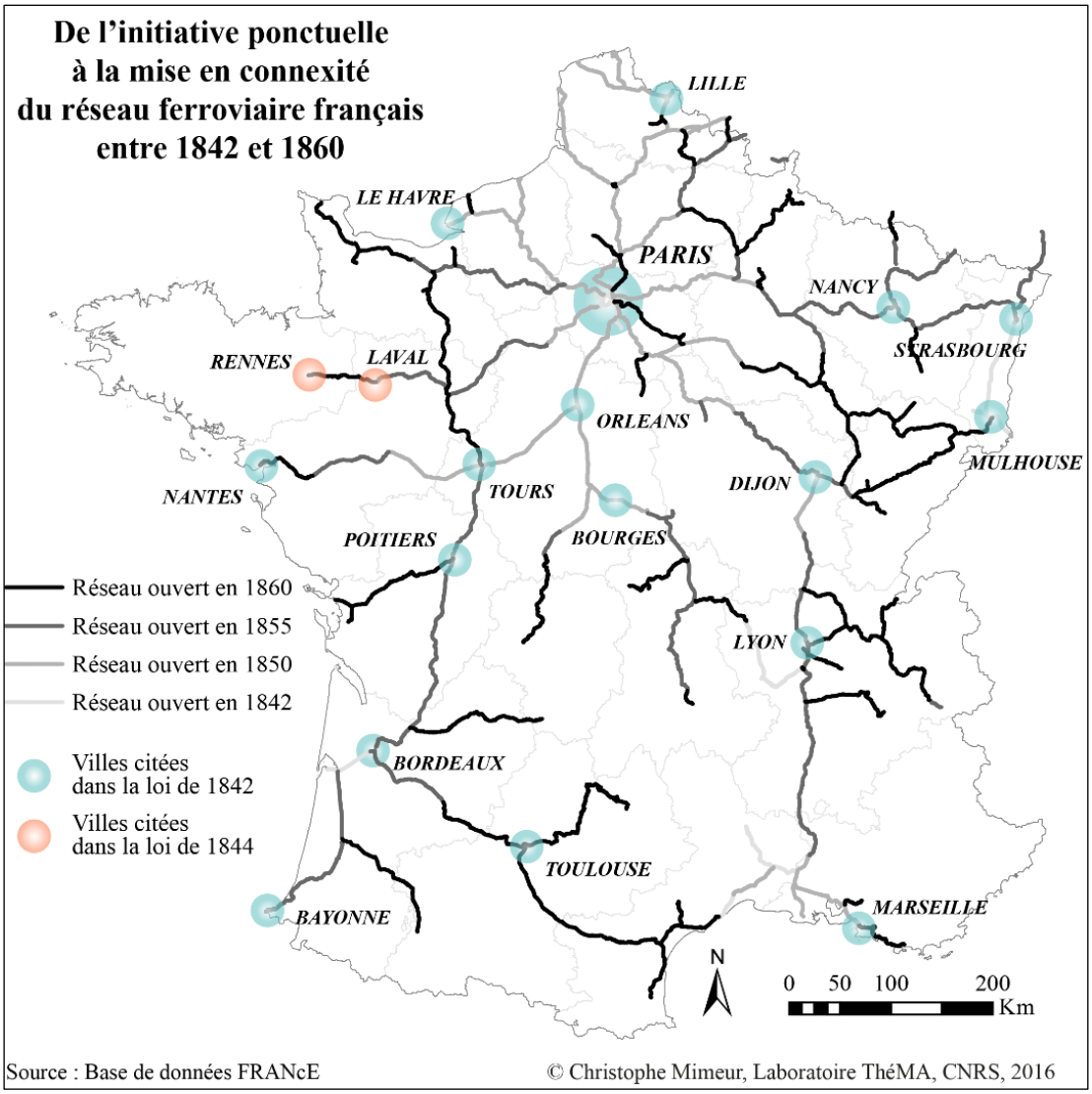


Figure 3. 10. De l'initiative ponctuelle à la mise en connexité du réseau ferroviaire français entre 1842 et 1860

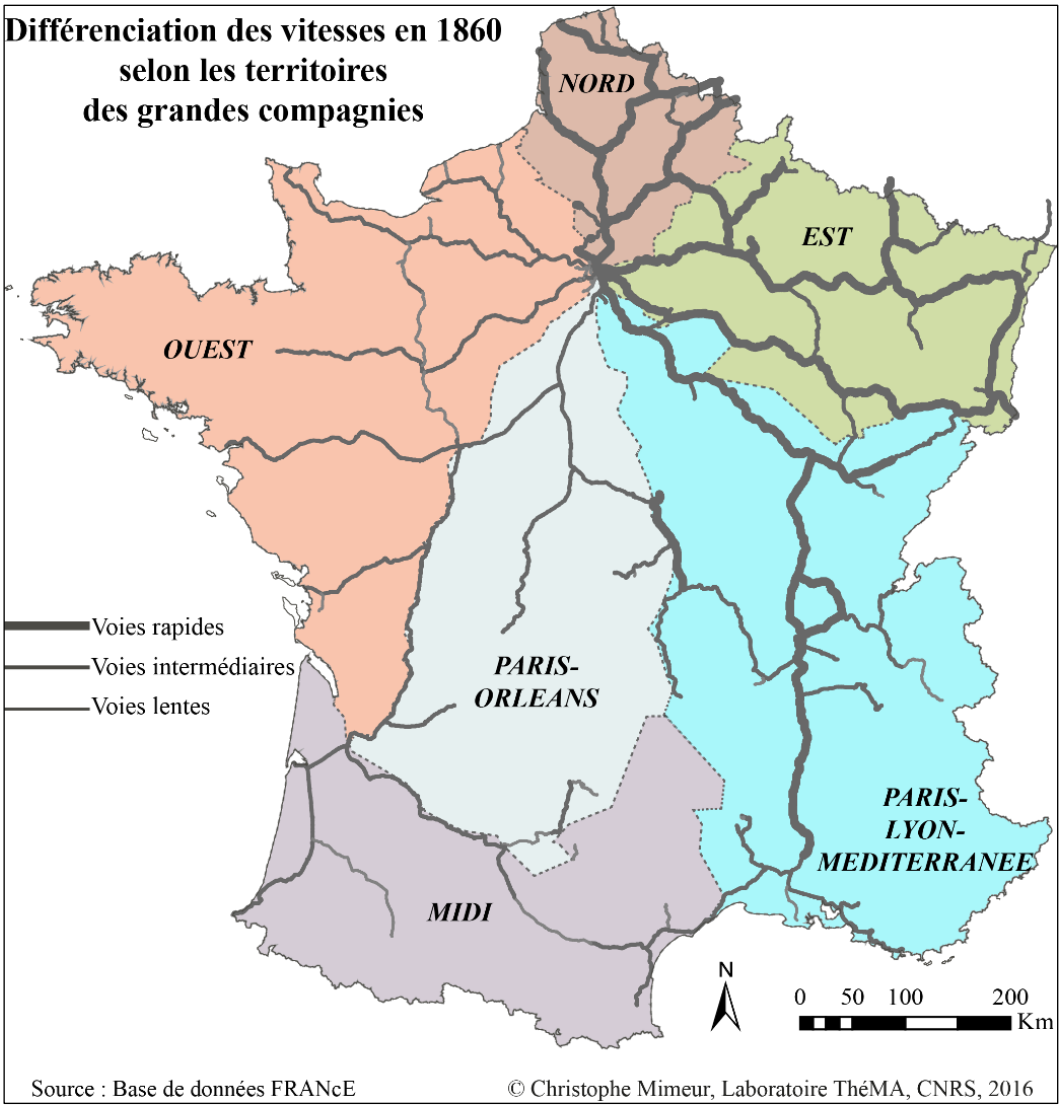


Figure 3. 11. Différenciation des vitesses en 1860 selon les territoires des grandes compagnies

Pour autant, la nouvelle information sur les vitesses nous permet de complexifier l'analyse de la mise en connexité du réseau ferroviaire jusqu'en 1860 *(Figure 3. 10)*. La mise en connexité du réseau, surtout opéré jusqu'en 1855, est le fait de la construction de voies rapides et intermédiaires, valeurs à nuancer avec le relatif écart entre les deux types de vitesse. Elles enregistrent les taux de croissance annuel moyen les plus fort jusqu'en 1855, d'abord avec les voies rapides, surtout développées au nord et à l'est de Paris *(Table 3. 4)*. Le développement plus tardif à l'ouest est ainsi surtout le fait de voies intermédiaires. Ainsi, entre 1855 et 1860, seuls quelques raccordements viennent terminer la mise en connexité du réseau alors que dans le même temps, la longueur des constructions de voies lentes augmente de 37 % chaque année. Le constat est encore plus saisissant quand on confronte les classes de vitesses et les « territoires » des grandes compagnies ferroviaires *(Figure 3. 11)*. Ainsi, les voies les plus rapides dessinent trois branches de l'Etoile de Legrand, vers Lille, Strasbourg et Marseille, dont les lignes appartiennent à trois compagnies différentes. Les branches vers la Manche, l'Ouest et le Sud-Ouest sont des voies sur lesquelles la vitesse est intermédiaire. Enfin, la transversale Atlantique-Méditerranée se fait pour partie sur voies lentes.

Suivant le processus de mise en connexité du réseau, cette représentation permet de rendre compte d'un processus rapide selon lequel la croissance du réseau participe d'un balancier entre complémentarité et concurrence (Dupuy, 1987a). Evoquée pour faire le parallèle entre la croissance du chemin de fer et l'essor de la télégraphie électrique, la complémentarité montre ici la possibilité de liaisons transversales entre les territoires des grandes compagnies. Les liaisons entre la Compagnie du Nord et celle de l'Ouest sont inexistantes en 1860, alors que cette dernière est bien reliée à l'axe Paris-Bordeaux exploité par la Compagnie Paris-Orléans. La concurrence transparaît quant à elle, dans la présence d'une grande ligne de Legrand dans chaque compagnie, alors même que Mulhouse peut être rejoint *via* Vesoul (Compagnie de l'Est) ou *via* Dijon (PLM), dès 1860.

3.2.2. **Le cycle de vie du réseau ferroviaire dans la base de données FRANcE**

La première phase de connexité étant achevée en 1860, on se propose désormais de qualifier la phase de croissance *(cf. 1.1.3, p. 30)*, en appliquant le modèle du cycle de vie des produits *(Figure 1. 3, p. 34)*. Le recours à la théorie permet de dépasser le simple inventaire patrimonial pour dégager des processus de plus long terme. Le modèle décrit alors trois périodes (Garrison et Levinson, 2014) :

- Une phase de *développement* : il s'agit des prémices du réseau, dans une période de découverte de l'innovation technique ;
- Une phase de *croissance*, où l'innovation est rôdée et efficace, à partir d'une année d'inflexion ;
- Une phase de *maturité*, où l'innovation n'apporte plus que des effets marginaux.

La succession de ces trois phases décrit une courbe en S *(S-curve) (Figure 1. 4, p. 35)*, que de nombreux chercheurs ont cherché à estimer à partir de données sur les systèmes de transport (Garrison et Souleyrette, 1996). Il s'agit ici de voir dans quelle mesure l'extension du réseau ferroviaire entre 1840 et 1930 répond au modèle de cycle de vie des produits, décrits par la théorie des réseaux et celle des innovations. Cette modélisation permet de mesurer également la date à laquelle la croissance du réseau s'est considérablement infléchie.

A partir des données sur l'évolution du linéaire du réseau entre 1840 et 1930, estimer la courbe logistique de la diffusion d'un réseau est donné par l'équation suivante :

$$\frac{N_t}{N_{max} - N_t} = e^{at+b} \quad (1) \qquad \text{ou} \qquad ln\frac{N_t}{N_{max} - N_t} = at + b \quad (2)$$

où :

- N_t est la taille du réseau à un temps t
- N_{max} est la taille maximale du réseau, 1930 dans notre cas
- t est l'année d'observation
- a et b sont les paramètres du modèle.

La méthode d'estimation de ce modèle consiste à définir a et b à l'aide d'une régression linéaire, où la variable explicative est le temps. Son application passe par la définition du point d'inflexion de la courbe *(t_i)*, donnée par :

$$t_i = \frac{b}{-a} \tag{3}$$

Enfin, la courbe est obtenue par l'équation suivante :

$$\widehat{N}_t = \frac{N_{max}}{1 + e^{(-a(t-t_i))}} \tag{4}$$

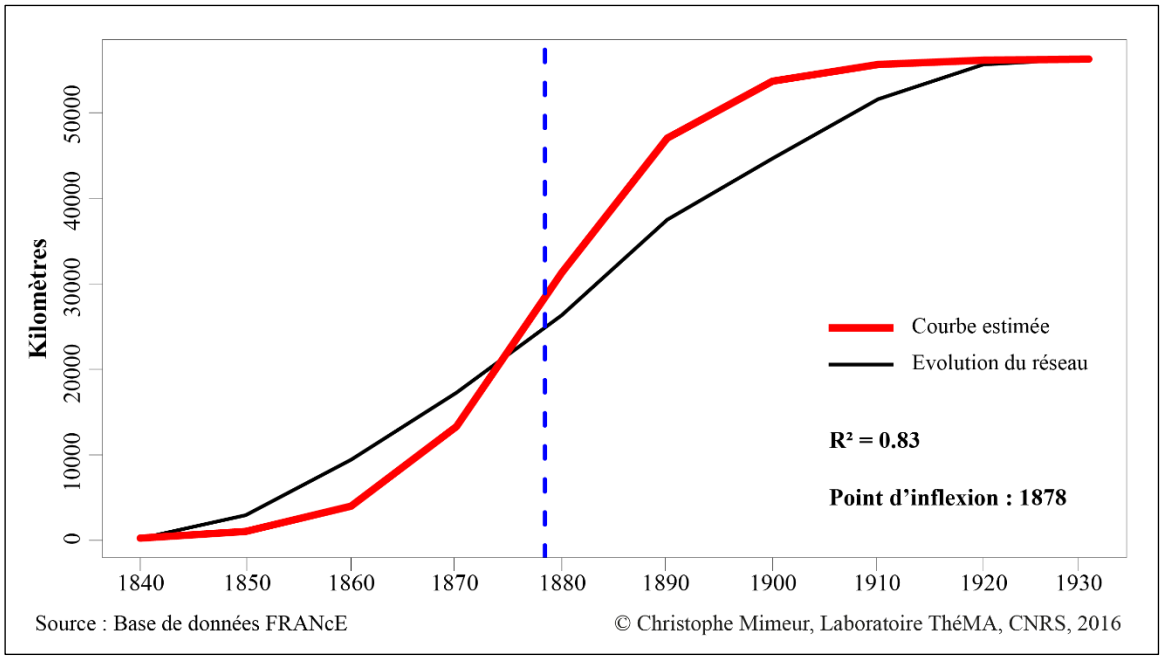


Figure 3. 12. Estimation de la courbe tendancielle du réseau ferroviaire entre 1840 et 1930

Ainsi, la *Figure 3. 12* montre la croissance globale du réseau entre 1840 et 1930, accompagnée de la courbe en S qui modélise les différentes phases de croissance, ainsi que l'identification d'une date d'inflexion du réseau. L'estimation globale donne une date d'inflexion du réseau en 1878, date à laquelle le plan Freycinet a été voté, il s'agit de la barre verticale dans le graphique. Le R^2 mesure quant à lui la corrélation entre les points issus de la base de données FRANcE et les points de la courbe tendancielle dessinée par le modèle. Ainsi, le processus de croissance du réseau, issu de la théorie des réseaux, identifie à lui seul l'ambitieux programme de travaux publics initié par le ministre des travaux publics Charles de Freycinet. En organisant le rachat de petites compagnies non rentables, l'objectif est désormais de donner un accès au chemin de fer à tous les Français, afin de favoriser le développement économique et désenclaver les régions les moins favorisées (Caron, 1997). Non seulement les sous-préfectures doivent avoir un accès au réseau, mais aussi un maximum de chefs-lieux de canton. Pour rendre compte de l'expansion du réseau sur le territoire français et des contrastes d'équipement (Verdier et Bretagnolle, 2007), nous étudions l'évolution de la densité ferroviaire entre 1860 et 1930, exprimée par département en mètres par kilomètres carrés *(Figure 3. 13)*.

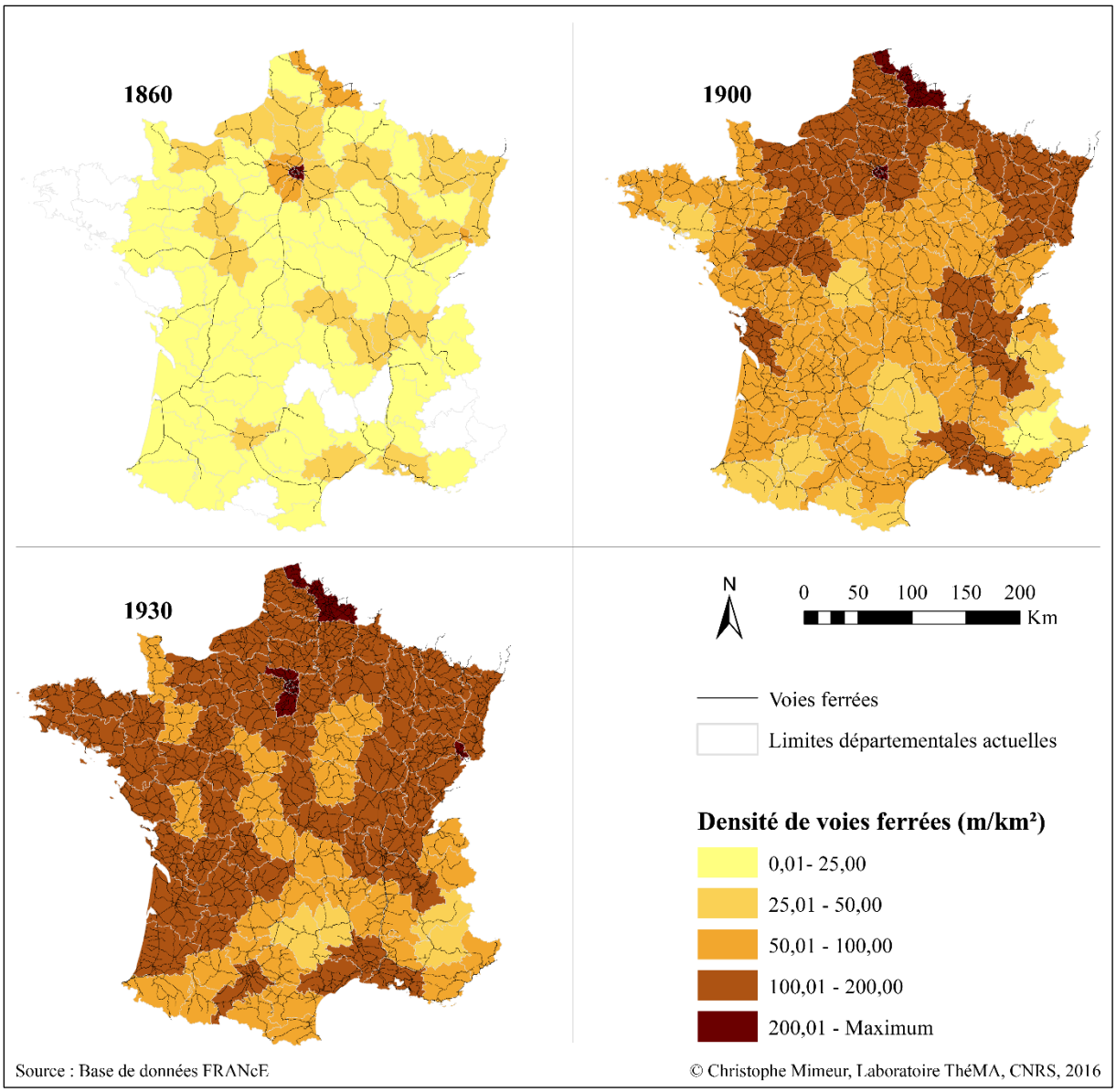


Figure 3. 13. Evolution de la densité ferroviaire entre 1860 et 1930

Lors de la mise en connexité du réseau en 1860, l'accès au réseau est largement lié eu dessin de l'étoile de Legrand. Les densités les plus fortes suivent ainsi les axes vers le Nord et vers l'Est d'une part, ainsi que les points clairement énoncés dans la loi d'autre part. Par ailleurs, la Bretagne, à l'extrémité ouest du pays n'est encore pas desservie par le train, de même que les zones les plus élevées – les Alpes, les Pyrénées et le Massif Central. En 1900, tous les départements sont desservis par le réseau et la densité exprime la même logique de celle d'il y a quarante ans, en dépit de l'augmentation globale des densités. Le retard pris par certains départements en 1860 n'est donc que partiellement comblé en 1900.

De plus, la disparité Est/Ouest est accentuée par une couverture du réseau dans les bassins industriels traditionnels du XIXème siècle. A l'Ouest, la densité est forte autour du Mans, et tend à

diminuer à mesure que l'on se dirige vers le Sud. A l'Est, on souligne une discontinuité entre la région parisienne, le Nord et l'Est d'une part, le Centre-Est largement doté et le Sud d'autre part, malgré le doublement de l'axe Lyon-Marseille de part et d'autre du Rhône.

En 1930, date à laquelle le réseau est le plus dense, la répartition des densités peut être rapprochée des logiques réticulaires des grands axes pourtant dessinés presque un siècle plus tôt. Elle laisse entrevoir un couloir de passage à proximité du Mans, vers Rennes, Nantes et Bordeaux. Les développements les plus forts touchent le Grand Ouest, autour de l'axe Paris-Bordeaux. Le même constat peut être fait autour d'un axe Paris-Clermont. Etonnamment, on constate un vide entre Paris et Dijon et entre Lyon et Marseille, malgré l'importance croissante de la ligne Paris-Lyon-Marseille. Ainsi, malgré un niveau d'équipement supérieur à 100 mètres par km² dans une grande majorité des départements, des inégalités subsistent encore, surtout dans le Centre de la France et les principales montagnes.

Pourtant, en dépit des gains d'équipement sur l'ensemble du territoire, il faut souligner que *« les vitesses élevées ne valent que pour les grandes liaisons, dès que l'on quitte les lignes principales, un net ralentissement s'opère »* (Studeny, 1995). Ainsi, parmi les 181 lignes inscrites dans le projet, de nombreuses sont en réalité empruntées par des « tortillards » sur des voies métriques, alors que le réseau classique double tend à stagner durant cette phase *(Figure 3. 5, p. 110)*. Ainsi, la courbe tendancielle de la croissance du réseau peut être également évaluée individuellement pour chaque classe de vitesses, afin d'étudier les rythmes de construction du réseau pendant sa phase de croissance *(Figure 3. 14)*.

Les traits pleins représentent la croissance réelle du réseau alors que les courbes en pointillés sont une estimation de la courbe tendancielle. La corrélation entre les deux représentations est traduite par le coefficient de corrélation, tandis que les dates d'inflexion sont représentées par les barres verticales. Si les prémices du réseau voient les mêmes tendances pour les trois types de réseaux, la mise en connexité marque le début de la différenciation des tendances. L'estimation de la courbe tendancielle du réseau rapide s'infléchit en 1881, même si son développement reste modeste par rapport aux réseaux intermédiaires et lents. Les voies les plus rapides ont déjà été mises en service à l'occasion de la mise en application de l'Etoile de Legrand. En revanche, c'est le réseau intermédiaire qui va croître le plus entre 1840 et 1920, puisque les premières fermetures interviendront en 1926. L'inflexion du réseau intermédiaire se situe en 1878, comme la concrétisation des lois de 1865, qui voient la naissance des chemin de fer d'intérêt local (Caron, 1997). La croissance de nombreuses compagnies aux mains de notables locaux ne fut que temporaire et les rachats par les grandes compagnies s'opèrent à la fin des années 1870. Pour autant, la courbe tendancielle montre que la « fièvre » du réseau intermédiaire ne connaît qu'un point d'arrêt pour mieux redémarrer ensuite. En effet, le plan Freycinet instituant le développement des lignes locales se concrétise par l'extension du réseau secondaire, en partie, mais surtout par l'inflexion du réseau lent en 1883. Alors que la littérature identifie la date de 1914 comme celle de l'achèvement du développement Freycinet, la courbe tendancielle du réseau traduit la phase de maturité du réseau lent entre 1910 et 1920. Celle des réseaux rapides et intermédiaire se situe à la charnière des XIXème et XXème siècle.

Alors que la décroissance des réseaux commence après 1930, les courbes tendancielles montrent comment la phase de maturité a été très éphémère. La courbe globale l'identifie entre 1910 et 1930 mais elle est également décalée quand on discrétise le réseau par les classes de vitesse. Le réseau rapide est celui qui connaît la plus longue phase de maturité à partir de 1900. Celle du réseau intermédiaire est concomitante à celle du réseau rapide, à la différence qu'il est le premier touché par les fermetures. Enfin, la phase de maturité du réseau lent est celle qui arrive le plus tard après 1910.

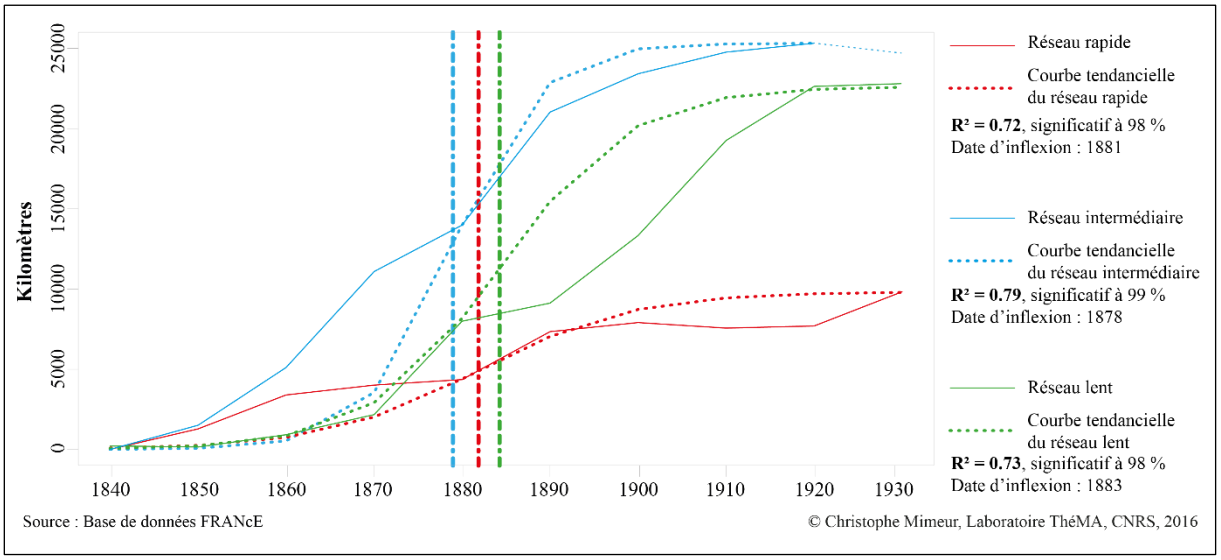


Figure 3. 14. Estimation des courbes tendancielles du réseau ferroviaire entre 1840 et 1930 selon la classe de vitesses

Quand on confronte les courbes globales et différenciées d'une part, et les dates d'inflexion d'autre part, on voit comment la phase de naissance du réseau est à rapprocher de l'expansion du réseau rapide sur les axes de Legrand à l'Est de la France. Par la suite, l'inflexion du réseau, tendanciellement identifiée en 1878 pour le réseau global est le fait du développement des réseaux intermédiaires. Enfin, sa phase de maturité, identifiée tardivement entre 1910 et 1930, est le résultat d'une arrivée à maturité du réseau lent à la veille de la Première Guerre Mondiale. Il s'agira dans la suite de notre travail d'identifier les effets de ces ruptures dans les dynamiques de peuplement. Que ce soit dans les phases de naissance et de croissance, mais aussi celle de la maturité, l'étude des différenciations spatiales est envisagée pour dégager des contrastes entre les phases et à l'intérieur de ces phases, en dépit d'une stabilité observée d'un simple œil dans la phase de maturité.

Ainsi, cette première approche de la base de données FRANcE montre la richesse de l'information quand on la confronte aux propriétés de la théorie des réseaux. Cet outil de recherche, le SIG-H, semble répondre aux enjeux de la démarche épistémologique que nous avons envisagée dans le chapitre précédent. Pourtant, sa complexification semble indispensable pour répondre aux enjeux de nos questionnements.

3.2.3. **Se déplacer en train : quelle approche pour le déplacement dans la base de données FRANcE**

L'approche par l'accessibilité que nous défendons dans ce travail requiert la complexification de l'information. L'intensité des possibilités d'interactions qu'elle permet de traduire est l'approche privilégiée. Dans son ouvrage sur la production des mobilités, Tim Cresswell souligne que pendant longtemps, les approches de la mobilité *« ne sont généralement pas définis spatialement et de ce fait, peuvent suggérer des mouvements sans contrainte spatiale et sans limites – puisque toute la question est la résistance à soi-même, aux observateurs, aux sujets. Mais suggérer que la liberté et l'égalité en matière de mobilité en découlent est en soi une illusion puisque nous n'avons pas tous le même accès à la route »* (Cresswell, 2001). Dès lors, notre approche vise à comprendre quelles sont les inégalités en matière d'accès aux ressources, alors même que le réseau tend à se diffuser sur l'ensemble du territoire, de manière toutefois hiérarchisée. Ainsi, parmi les principes géohistoriques les plus forts (Grataloup, 2015), les dynamiques de changement sont d'autant plus intenses qu'agissent des forces de proximité, traduites par la formule suivante par Christian Grataloup :

Densité + Connexité = Historité

Traduire les évolutions et l'intensité des densités revient dans notre cas à analyser le système de peuplement et ses dynamiques, par son évolution sur le temps long et sa distribution dans l'espace. L'historité est alors définie comme un *« processus propre à une société. […] Lorsqu'il s'agit surtout de logique spatiale, l'historité est mutante, dynamique et relève surtout de la transformation »* (Grataloup, 2015). D'autre part, objectiver l'évolution et l'intensité de la connexité relève dans notre travail de la mobilisation de la théorie des graphes. Elle permet non seulement de rendre compte de la proximité mais aussi de la mobilité (Chapelon, 1997). Sans entrer dans les détails méthodologiques développés dans la seconde partie, nous montrons ici comment le seul outil SIG ne permet pas de rendre compte des évolutions que nous souhaitons voir dégager. La modélisation des réseaux consiste dans la réalisation d'algorithmes qui définissent des chemins optimaux en fonction d'un critère de coût (Mathis, 2003). Cette mesure a pour objectif de prendre en compte la complexité des contraintes de déplacements : ils sont traditionnellement la traduction d'une distance, d'une valeur monétaire, d'un effort physique. Quand on replace cette exigence dans le cadre géohistorique, ces contraintes de déplacements sont de plusieurs ordres et leur maîtrise devient un *« vecteur des inégalités sociales »*, qui vient remettre en cause *« l'évidence trompeuse d'une vitesse généralisée »* (Flonneau et Guigueno, 2009).

Alors que notre étude vise les 36 000 communes françaises, les quelques 11 000 gares françaises à l'apogée du réseau ne couvrent pas l'ensemble du territoire. On souligne le caractère nécessairement discontinu de l'accès au réseau ferroviaire, par opposition au réseau routier ou piéton dans lequel on peut entrer à n'importe quel point du réseau (Stathopoulos, 1997). Ainsi, *« les liaisons rurales se heurtent longtemps encore à l'obstacle des lieux et des distances »* (Studeny, 1995). Notre approche passe alors par la prise en compte d'un nouveau segment dans la chaîne de déplacement dans laquelle l'individu doit d'abord rejoindre une gare avant de pouvoir profiter des

bénéfices de la vitesse. Alors que de nombreux discours de l'époque, techniques ou politiques, portent sur les records de vitesses (Studeny, 1995), notre approche privilégie les *« moyennes usuelles »* qui résultent le plus souvent de *« réseaux multiples »* (Flonneau et Guigueno, 2009).

Une autre contrainte de déplacement dans l'appréhension de l'accessibilité réside dans le caractère évolutif du réseau ferroviaire dans notre période d'étude, rendant le graphe qui le sous-tend tout aussi évolutif. Ainsi, si les outils SIG commerciaux revêtent déjà des limites dans la prise en compte de l'information temporelle (Thévenin, 2010), alors la complexification de l'information les multiplie. Le logiciel SIG doit alors s'accommoder d'un double langage : celui issu de la théorie des graphes, qui est implémentée dans l'extension Network Analyst du logiciel ArcGIS par exemple, celui de la manière dont est implémentée l'information temporelle dans les couches d'informations géohistoriques. Par ailleurs, notre positionnement dans les humanités numériques visant à diversifier l'information et la valoriser par la diffusion nous conduit à nous tourner vers des outils libres, sans avoir l'obligation de posséder un logiciel au préalable pour consulter et traiter l'information géohistorique.

Pour rendre compte de ces difficultés, nous reprenons l'exemple de la zone du bassin industriel du Creusot. Il montre comment les contraintes de déplacements sont à appréhender de manière dynamique pour rendre compte des contextes dans lesquels nous objectivons la construction de connaissances géohistoriques. Ainsi, pour un voyage entre Genouilly et Dijon, l'évaluation des temps de parcours est ici à appréhender de deux manières différentes selon que l'on se place en 1880 ou en 1890 *(Figure 3. 15)*. En 1890, l'ouverture de la ligne entre Montchanin et Saint-Gengoux-le-National un an plus tôt a permis la création d'une gare à Genouilly. En revanche, en 1880, rejoindre Dijon en train requiert d'abord de se rendre à la gare de Montchanin pour emprunter ensuite le réseau ferroviaire en suivant la côte viticole jusqu'à la capitale des ducs de Bourgogne. Ainsi, notre approche par l'accessibilité doit rendre compte de cette double contrainte de déplacement sur le temps long : le segment nécessaire pour rejoindre le réseau ferroviaire, l'évolution des vitesses moyennes le long de ce même réseau. Cette complexification de l'information ne permet pas de recourir directement au graphe et suggère de faire des choix méthodologiques dont il sera question dans la seconde partie.

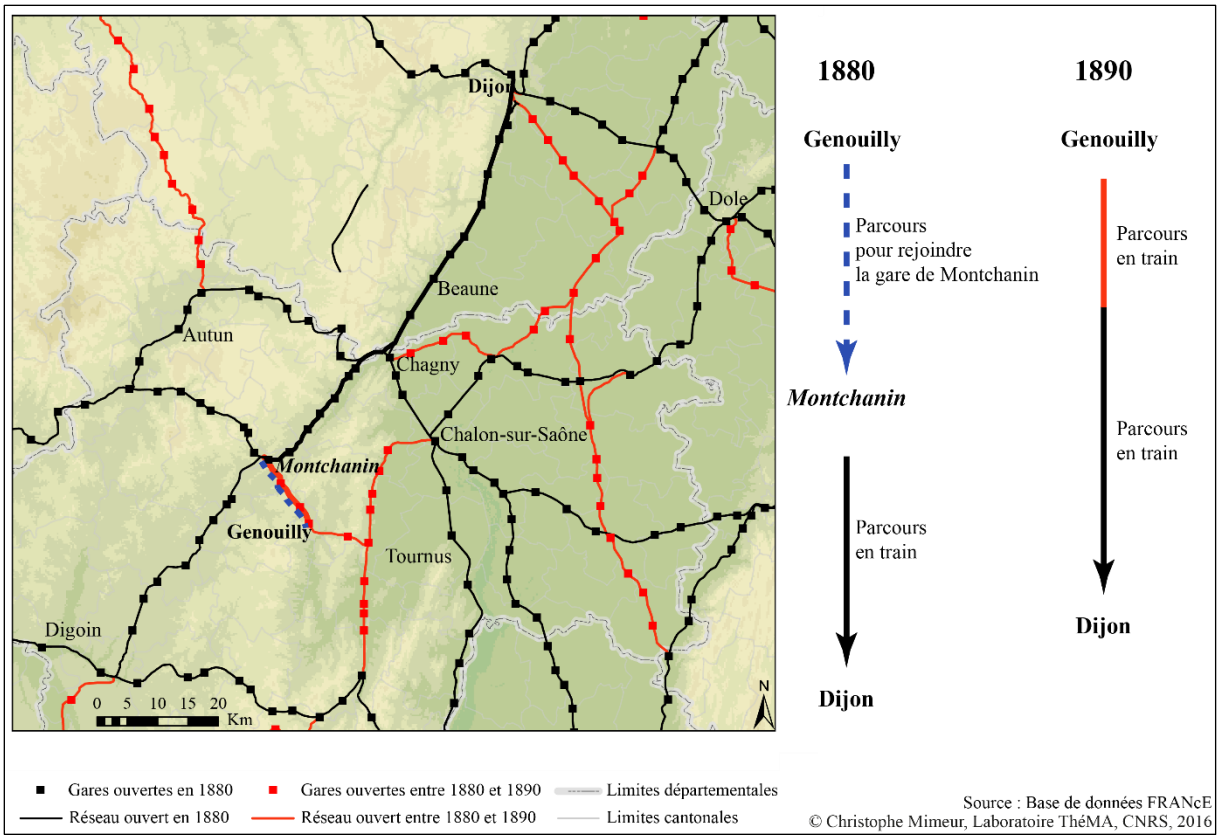


Figure 3. 15. La complexification des contraintes de déplacement pour mesurer l'accessibilité dans la base de données FRANcE

Le calcul d'indicateurs d'accessibilité généralisée requiert une méthodologie applicable dans tous les cas de manière opérationnelle. De tels chiffres démontrent que l'utilisation d'un logiciel SIG commercial atteint ses limites dans les possibilités de traitements d'une part, ainsi que dans les capacités de calculs et de stockage d'autre part. Dès lors, la délocalisation des données vers des outils plus optimisés pose la question de la restructuration des données de la base de données FRANcE, sans pour autant perdre en qualité et finesse d'informations.

Ainsi, cette dernière section a permis de voir dans quelle mesure l'information inédite contenue dans la base de données FRANcE participe à la construction de connaissances, entre empirisme issu de la donnée géohistorique et théorie issue de la conceptualisation du réseau. Elle permet de mettre en exergue la différenciation des rythmes et des croissances du réseau ferroviaire français. Pour autant, la prochaine étape de complexification qui passe par la prise en compte des contraintes de déplacements dans les contextes définis par ailleurs montre comment l'outil SIG classique pointe ses premières limites quant aux possibilités de traitements et de calculs.

Conclusion

Ce troisième chapitre a permis de montrer comment notre travail de recherche s'insère dans un premier temps dans le challenge traditionnel identifié pour la constitution d'un SIG-H. Les tiraillements thématiques et méthodologiques permettent de faire le choix des sources historiques à mobiliser, qui est toutefois le plus souvent contraint par la rareté et l'hétérogénéité des sources primaires. Dans le cadre de notre positionnement méthodologique et épistémologique, nous avons opté pour une modélisation conceptuelle des données largement évolutive, qui place d'un côté l'information liée au réseau, de l'autre l'information liée à la population. Les découpages spatiaux sont une catégorie à eux seuls dans la mesure où la démarche géohistorique nous autorise, sinon à les multiplier, mais à diversifier et à s'affranchir des découpages liés aux traditionnelles bornes fixées par les historiens.

Parallèlement, nous postulons que cette base de données FRANcE participe à revisiter le patrimoine ferroviaire français. Cette base de données géolocalisées vient répondre aux enjeux des humanités numériques, comme potentiel outil de recherche pouvant être mobilisé par de nombreuses disciplines. Le croisement avec d'autres outils est alors rendu possible dans la mesure où la valorisation et la diffusion des données se développent. Pour cela, nous défendons une approche par les logiciels libres, permettant le partage et l'élargissement du dialogue avec d'autres chercheurs et ainsi les possibilités d'analyse.

Si la constitution de la base de données FRANcE ne présente pas de réelles innovations méthodologiques dans l'implémentation de l'information géohistorique à l'intérieur du SIG-H, la place que nous souhaitons donner à la *« révolution de la vitesse »* (Studeny, 1995) dans nos analyses vient enrichir de manière inédite l'inventaire des voies ferrées depuis 1830. Pour couvrir l'ensemble du territoire, un croisement de sources secondaires a abouti à la généralisation des vitesses sur l'ensemble du réseau depuis ses débuts. Les premières exploitations de la base de données FRANcE ont montré comment le SIG-H pouvait nous permettre de qualifier la phase de croissance du réseau ferroviaire français entre 1840 et 1930 dans une démarche abductive, en s'emparant des principales propriétés dégagées par la théorie des réseaux. Elles ont permis de voir dans quelles mesures la hiérarchisation progressive du réseau se traduit dans les rythmes de croissance du cycle de vie du réseau ferroviaire jusqu'à son apogée en 1930.

L'approche que nous défendons par l'accessibilité comme indicateur des inégalités d'accès aux réseaux d'une part, à l'ensemble des ressources d'autre part passe par la théorie des graphes. Son utilisation nécessite toutefois une couverture spatiale et temporelle complète, qui doit être garantie par ailleurs par une topologie parfaite du réseau, où chaque nœud du réseau est associé à un arc, afin d'appliquer des algorithmes mathématiques. La prise en compte des contraintes de déplacement dans le contexte géohistorique adopté passe alors par une complexification de la chaîne de déplacement depuis une origine vers une destination. Alors, l'approche par l'accessibilité que nous défendons montre les premières limites de l'usage d'un seul outil SIG dans ses capacités de calculs et de stockage d'une information sur la longue durée.

CONCLUSION DE LA PREMIERE PARTIE

L'objectif de cette première partie était de replacer le concept d'effet structurant des infrastructures de transport dans une perspective historique et épistémologique. Cette étape est nécessaire dans la définition d'une réflexion géohistorique des liens entre dynamiques d'un système de transport et dynamiques d'un système de peuplement sur le temps long, celui de l'innovation ferroviaire depuis les premiers rails de la première moitié du XIXème siècle. Alors ce travail semble avoir sa place dans la valorisation du patrimoine du chemin de fer français, dont l'apogée se situant en 1930, a longtemps oublié les traces des voies métriques qui sillonnaient les campagnes et traversaient chaque chef-lieu de canton.

Le premier chapitre a permis de replacer la genèse du vocable de l'effet structurant dans la construction d'un mythe largement véhiculé par le corps des ingénieurs identifié dans le courant saint-simonien, à la tête de la plupart des projets d'infrastructures du XIXème siècle. La remise en cause du mythe participe cependant à une controverse scientifique, dans laquelle la prise en compte de l'espace est inégale, souvent abstraite, mais où la profondeur temporelle semble quant à elle bien identifiée comme un levier d'analyse dans l'évaluation des effets structurants. Il vient donc justifier la nécessité d'une réflexion sur les échelles spatiales et temporelles dans lesquelles nous envisageons de revisiter le concept d'effet structurant des infrastructures de transport.

Le second chapitre a participé à ces réflexions en fixant un cadre épistémologique à l'intérieur duquel nous envisageons la construction de connaissances géohistoriques. En mobilisant la géohistoire contemporaine, ce travail s'inscrit dans une démarche constructiviste dans la géographie théorique et quantitative, dans laquelle l'abduction permet la construction de connaissances par la donnée, autant que par l'utilisation de théories d'arrière-plan, participant à la définition des contextes dans lesquelles les conclusions de nos analyses s'inscrivent. Ce positionnement vient donner une place centrale à l'information géohistorique, valorisée depuis le *« spatial turn »* opéré dans les sciences humaines et sociales. A cette occasion, l'utilisation du Système d'Information Géographique dans le champ émergent des humanités numériques a ouvert de nombreuses pistes.

Le troisième chapitre a permis d'identifier les enjeux issus du challenge traditionnel des projets de SIG-H, qui ont été par ailleurs soulevés dans la construction de la base de données FRANcE. Sans apporter de nouveautés méthodologiques dans la constitution d'un SIG-H, nous identifions en revanche une information inédite portant sur les vitesses moyennes sur l'ensemble du réseau ferroviaire français, sur l'ensemble de notre période d'étude. Si les premières exploitations montrent la richesse de l'information contenue dans la base de données FRANcE, ouvrant de nouvelles pistes d'investigations, alors la complexification inhérente aux enjeux de notre recherche montre vite les limites d'un seul outil SIG-H.

C’est ici que nous envisageons le challenge émergent dans les *Spatial Humanities*, où nous défendons une posture qui vise à dépasser la quête toujours plus longue de nouvelles données pour la valorisation de la donnée existante. Ainsi, l’approche par la *« révolution de la vitesse »* est selon nous celle qui permet de valoriser la nouvelle information géohistorique d’une part, celle qui promet d’apporter de nouvelles connaissances géohistoriques pour revisiter les effets structurants des infrastructures de transport d’autre part. Une approche par l’accessibilité semble revêtir un enjeu largement investi dans la géographie des transports, mais encore peu représentée sur le temps long. Compte tenu des difficultés à appréhender les contraintes de déplacements dans une perspective géohistorique, nous identifions à la fin de cette première partie la nécessité de retravailler ou redéfinir un cadre méthodologique, permettant d’intégrer ces nouvelles informations et de procéder à de nouvelles analyses. Parce qu’un seul outil SIG ne peut répondre aux enjeux dégagés tout au long de ces trois chapitres, nous défendons ici la construction d’un instrument de recherche qui permet de revisiter les progrès techniques et le déploiement du réseau ferroviaire français. La définition de cet instrument est à resituer dans le projet SIG-H comme la passerelle entre le challenge traditionnel et le challenge émergent : c’est l’objet de la seconde partie de cette thèse.

DEUXIEME PARTIE :

L'ACCESSIBILITE DANS UN CONTEXTE GEOHISTORIQUE : UNE MODELISATION CROISSANTE DE LA DONNEE, DE L'EFFET DU RESEAU A SA DYNAMIQUE DE CROISSANCE

INTRODUCTION DE LA DEUXIEME PARTIE

Si les études de long terme entre infrastructures de transport et croissance de la population sont nombreuses (Kasraian et al., 2016), les relations sont souvent envisagées à partir de données historiques fragmentées, venant réduire les fenêtres de temps et d'espaces des analyses (Bretagnolle, 2014). Les jalons de notre positionnement posés dans la première partie de la thèse visent au contraire la largesse de la fenêtre spatio-temporelle, faisant ainsi apparaître des enjeux transcalaires, dans lesquels le géographe aime à dégager des structures à partir de données pourtant collectées à partir de la maille élémentaire de la commune ; mais aussi des enjeux temporels, où l'approche par le temps de parcours vise à identifier des rythmes dans la relation entre transport et territoire. Les enjeux géohistoriques sont aussi bien identifiés : dépasser le culte de la collecte de nouvelles données historiques pour l'émergence de connaissances structurantes, sans occulter les dimensions patrimoniales d'une telle entreprise. Cela veut dire que si le recours aux sources est souvent un acte chronophage et n'est pas à privilégier systématiquement, le recours à des données déjà existantes s'inscrit pleinement dans les humanités numériques. Il s'agit là d'un choix de réutiliser des données collectées au prix de moyens conséquents par ailleurs, qui passe alors par leur intégration dans la structuration des données déjà stabilisée. Le recours à l'outil SIG est apparu indispensable, dans la lignée des nombreux travaux géohistoriques, occupés par le challenge traditionnel identifié plus haut (Gregory et Ell, 2007 ; Gregory et Geddes, 2014) et reste encore indispensable dans la mise en relation de données spatio-temporelles de sources et structures diverses.

Pour autant, nous identifions à ce stade de multiples raisons nous incitant à ne plus se limiter à ce seul type d'outil (Thévenin et al., 2016), mais plutôt à dessiner les contours d'un instrument de recherche au centre duquel on place la mesure, l'objet de la mesure, de sorte de questionner ses apports et ses limites. Il s'agit aussi d'estimer sa valeur ajoutée dans le débat sur les effets structurants du transport, en exploitant la diversité des outils de la géomatique. Cette partie vise donc à doter notre travail d'une réflexion sur la manière d'envisager les effets de l'infrastructure de transport sur l'évolution du système de peuplement. Aussi, ce travail vise à explorer les opportunités de la modélisation en géographie pour envisager le pendant de la controverse, c'est-à-dire en étudiant le rôle de la relativité des localisations dans la construction cumulative des réseaux de transport (Dupuy, 1993), qui n'est en revanche que très peu étudié d'un point de vue empirique (Levinson et Yerra, 2006), et encore moins dans un contexte géohistorique (Kasraian et al., 2016 ; Pumain, 2014).

Dès lors, la seconde partie de la thèse s'appuie sur la *« révolution de la vitesse »* (Studeny, 1995) pour enrichir les études de long terme en faisant de l'accessibilité par les temps de parcours, une nouvelle variable participant au débat sur les effets structurants des infrastructures de transport sur la longue durée. Comment la modéliser dans un contexte géohistorique, en mobilisant l'arsenal des possibilités laissées par la géomatique ? Comment la relier dans l'explication de ses relations avec un système de peuplement ? Comment la valoriser dans le cadre d'une modélisation dynamique d'un réseau de transport ? La seconde partie de la thèse s'inscrit pleinement dans le challenge émergent des SIG-H, où la donnée géohistorique a vocation à s'extraire du cadre d'un

seul outil pour diversifier les potentiels de modélisation face à la question de recherche des relations entre transport et territoire. Pour cela, si les rapprochements vers l'histoire sont parus naturels dès la première partie, le rapprochement vers d'autres disciplines participe à la construction de notre instrument de recherche.

L'objectif ultime de cette seconde partie est la définition d'un plan d'expérimentation qui permet de définir la stratégie d'analyse du lien entre réseau et territoire sur le temps long à partir de la construction de l'instrument de recherche qui sera détaillée au cours de cette partie. Ce plan d'expérimentation devra ainsi mettre à l'épreuve les données et méthodes auxquelles nous faisons appel dans une démarche d'ensemble cohérente pour une réponse à notre question de recherche.

La première étape vise à adapter la modélisation des données pour la construction d'une accessibilité géohistorique. Pour cela, nous adaptons les quatre piliers de l'accessibilité (Morris, Dumble et Wigan, 1979) établis dans les années 1970 dans notre contexte géohistorique, où le formalisme des graphes viendra remplir les quatre conditions suivantes (Chapitre 4) :

- *« les mesures d'accessibilité doivent incorporer un élément de séparation spatiale qui traduit les changements en termes de performance du système de transport »* :c'est la relativité de la notion de grande vitesse qui est à explorer ici ;
- *« les mesures d'accessibilité doivent être cohérentes avec les modèles de comportements »* : le recours aux travaux d'historiens est indispensable ;
- *« les mesures d'accessibilité doivent être techniquement réalisables et leur mise en œuvre aisée »* : le recours à la science informatique prendra alors tout son sens ;
- *« les mesures d'accessibilité doivent être facile à interpréter et à être traduites en termes de politiques urbaine ou de transport »* : l'interprétation géo-(historique) doit être systématiquement assortie aux formalisations mathématiques.

La modélisation de la donnée géohistorique permet alors de s'inscrire dans la lignée des travaux empiriques portant sur l'explication de la relation entre dotation en réseau et croissance locale. En s'appuyant des techniques de *« plus en plus sophistiquées »* (Kasraian et al., 2016), nous identifierons dans quelle mesure un rapprochement vers l'économétrie participe à l'enrichissement du débat. Ainsi, si la dotation en réseau sera enrichie dans cette thèse par les temps de parcours comme une variable explicative, alors les prudences méthodologiques (Offner, 1993a) nous poussent à diversifier l'information géohistorique dans la lignée du challenge émergent, en mettant à l'épreuve notre modèle conceptuel de données, pour le doter de nouvelles informations, participant à la prise en compte de contextes locaux dans la mesure des relations entre transport et territoire (Chapitre 5).

Enfin, nous investiguerons les potentialités de la modélisation dynamique en géographie pour investir d'autres hypothèses posées dans le débat sur les effets structurants des infrastructures de transport. Nous envisagerons une démarche de simulation de l'évolution d'un réseau de transport par l'identification de règles simples. Le recours à l'informatique permet d'envisager des comparaisons entre des situations simulées et des situations réelles passées, à partir de mécanismes simples à dimension heuristique et didactique, en donnant des pistes d'analyse de la *« causalité inverse »* des effets structurants des infrastructures de transport (Chapitre 6).

Chapitre 4. CAPITALISER L'INFORMATION GEOHISTORIQUE : ENTRE PRECISION ET GENERALISATION

Introduction

Sans se rapprocher de perspectives utilitaristes quant à l'exploitation de la base de données FRANcE, l'objectif de ce chapitre est la capitalisation de l'information géohistorique contenue dans notre corpus, indépendamment de toute autre information. Dans notre cadre, elle consiste alors à l'identification de leviers pour envisager de nouvelles approches, de nouvelles *« manières de faire »*. De ce point de vue, la scène de dialogue offerte par la géohistoire contemporaine participe au dialogue entre les disciplines des sciences humaines et sociales, mais pas seulement. Par ailleurs, cet objectif répond ici à notre positionnement dans le challenge émergent des projets de SIG-H, dans lequel nous souhaitons dépasser les enjeux liés à la collecte de l'information pour s'imprégner de ceux de son traitement, de leur enrichissement par eux-mêmes, participant à des processus se rapprochant de l'induction.

Naturellement, l'information géohistorique s'est trouvée déjà au centre de la construction de l'instrument de recherche : ainsi les premiers jalons de notre projet se sont rapprochés de l'outil SIG. Pour autant, nos larges échelles d'analyses, les critères de pertinence, de précision et de vraisemblance de l'appréhension des déplacements de l'époque participent à l'identification des limites de l'usage d'un seul outil, alors même qu'il possède des méthodes indispensables pour l'étude de phénomènes dans l'espace. Ainsi, *« en modélisation des réseaux, domaine issu de la recherche opérationnelle, une certaine forme d'empirisme tend à dominer, notamment dans les disciplines intermédiaires entre sciences humaines et sciences dures comme l'aménagement, disciplines qui, pour l'essentiel, importent leurs outils. Or, leurs besoins spécifiques ne sont que peu pris en compte par les disciplines fondamentales comme les mathématiques ou plus appliquées comme l'algorithmique, sans doute simplement parce que les dynamiques de recherche sont très différentes »* (Mathis, 2003). Le dimensionnement et l'ambition de notre recherche nous incitent à nous rapprocher de l'informatique et des mathématiques, suspectant dans la théorie des graphes un formalisme réunissant les deux ambitions que nous identifions dans les termes de capitalisation de l'information géohistorique.

Ce chapitre vise ainsi à identifier les contours de notre instrument de recherche, comme scène d'articulation des objets de recherche, des données, des méthodes et des outils de la géomatique pour la construction de nouvelles connaissances géohistoriques. Alors, nous relevons comment les enjeux traditionnellement dévolus à la science de l'information géographique peuvent être transposés dans le cadre plus large et plus riche de la *GeoComputation* (4.1.). Ce même cadre est celui de la transformation de l'information de la base de données FRANcE par un graphe ferroviaire fonctionnel (4.2.), qu'il convient d'enrichir pour rendre compte d'une chaîne de déplacement complète sur l'ensemble du territoire, venant contribuer aux débats sur l'arbitrage entre le recours à la source historique et le recours à la généralisation théorique de l'information face à l'incomplétude inhérente à notre démarche (4.3.).

4.1. Introduire la base de données FRANcE dans la *Geocomputation*

Face aux enjeux épistémologiques et méthodologiques que nous avons exposés dans le deuxième chapitre de ce travail, l'objectif de cette thèse est bien de construire un instrument permettant d'étudier une information spatio-temporelle. Aussi, nos premières investigations nous ont orientés vers l'outil SIG, que nous avons défini préalablement, en pointant les limites « matérielles » dans notre approche de l'accessibilité. Pour autant, face à nos questions de recherche, nous pointons ici les apports et limites de l'outil SIG d'un point de vue scientifique, qui ouvrent de nouvelles perspectives à la frontière entre l'information géographique et les sciences informatiques (Goodchild et Longley, 2014).

Plus qu'un outil pour le géographe, le Système d'Information Géographique questionne des objets et des méthodes dans la science de l'information géographique (Pickles, 1995). Ses dénominations sont en revanche diverses : les Américains la nomment volontiers *GIScience* (Goodchild, 1995), quand les francophones – Français et Québécois plus spécifiquement – la nomment *Géomatique* (Collet, 2005). De nombreux Britanniques ont quant à eux diffusé le terme de *Geocomputation* (Openshaw et Abrahart, 1996). Il s'agit ici de montrer dans quelle mesure nous nous emparons de ces concepts, ces méthodes et ces outils, en réponse à notre question de recherche. Loin de rejeter l'outil SIG appliqué à notre raisonnement, il s'agit au contraire de l'enrichir et de l'entourer d'autres outils participant à la construction d'un instrument géohistorique.

4.1.1. **Des enjeux de la science de l'information géographique …**

Nécessairement impliqués dans le tournant quantitatif de la géographie (Haggett, Frey et Cliff, 1977), les développements de l'informatique des années 60 et 70 ont multiplié les possibilités de calcul, ont complexifié les modèles les plus simples et ont attisé le besoin de données toujours plus nombreuses (Gahegan, 1999). Pour autant, l'auteur y voit aussi un asservissement du géographe à l'ordinateur, qui imposait une structure appauvrie des données, où les concepts des géographes avaient du mal à se retrouver et où les géographes peinaient à adapter leur vocabulaire. Pour autant, un peu plus tard, les premiers SIG ont été discrètement développés à l'écart des géographes, parfois devenus sceptiques à l'ordinateur. Gahegan suggère alors un rapprochement entre le géographe et l'informaticien, participant à la large utilisation des SIG depuis le milieu des années 1980.

En introduction de la définition d'une *GIScience*, Michael Goodchild et Paul Longley posent trois enjeux scientifiques (Goodchild et Longley, 2014). Ils peuvent être appliqués dans le cadre simple d'un SIG à de nombreux, sinon tous les objets de recherche, montrant ainsi que les implications sont avant tout scientifiques et fondamentales, avant d'être appliquées et opérationnelles. Nous identifions de même une proximité entre ces enjeux et ceux de notre travail de recherche. Il s'agit :

- D'un *enjeu pratique* : comment aborder l'incertitude dans l'information géographique (Openshaw et Taylor, 1979) ;

- D'un *enjeu empirique* : il répond à la première loi de la géographie établie par Walter Tobler, qui dit que *« tout interagit avec tout, mais deux objets proches ont plus de chances de le faire que deux objets éloignés »* (Tobler, 1970) ;
- D'un *enjeu théorique* : il permet de qualifier les relations entre les objets géographiques sur le plan, de manière topologique, connues sous le terme de *« 9-intersection of topology »* (Egenhofer et Franzosa, 1991).

Ils sont devenus d'autant plus prégnants que les géographes ont de plus en plus à se confronter à de grands ensembles de données spatio-temporelles : l'enjeu pratique réside dans l'extraction de l'information géographique et pose la question des échelles, tandis que les enjeux empiriques et théoriques posent la question de l'établissement de modèles de relation et d'interaction des objets représentés dans le SIG. Ces trois états de fait concordent également avec les questionnements de notre recherche et sont donc liés dans un raisonnement sur la longue durée. L'enjeu pratique l'est quand l'information est géohistorique (Knowles, 2005) : l'incertitude est liée alors à l'incomplétude, et à la méthode d'implémentation de l'information temporelle. L'enjeu empirique l'est quand l'accessibilité est selon nous une variable clé des relations entre réseau et territoire. L'enjeu théorique l'est tout autant quand il s'agit de croiser des informations surfaciques à des informations linéaires, *a fortiori* si l'on considère des découpages spatiaux et temporels évolutifs.

Ces trois enjeux de la science de l'information géographique répondent exactement et successivement à ceux que nous avons posés pour la construction de l'instrument de recherche de notre seconde partie. Il s'agit bien d'extraire l'information géohistorique contenue dans la base de données FRANcE, constituée à de larges échelles. Ici, l'objectif est la capitalisation de cette information pour en faire ressortir des connaissances géohistoriques inédites. Dans un second temps, les autres enjeux participent à la construction de modèles d'analyse spatiale (Miller et Goodchild, 2014), qui s'attachent aussi bien à définir des relations que des interactions.

Si l'adéquation entre notre question de recherche et le cadre des *GISciences* est éclairant, le recours à de larges bases de données spatio-temporelles pose la question de l'utilisation du seul outil SIG pour nos analyses. D'abord, les protocoles de collecte qui visent à consigner l'information spatio-temporelle participe aux limites des SIG quand ils sont confrontés aux temporalités (Thévenin, 2010). De plus, face à l'émergence de grandes bases de données spatio-temporelles, que beaucoup appellent aujourd'hui *big data* (Kitchin, 2014a), ce sont les performances de l'outil SIG qui sont remises en cause. Une des réponses à ces limites réside sans doute sur le passage progressif de l'utilisation d'un logiciel SIG commercial sur sa propre machine à l'utilisation de serveurs, qui ont permis la multiplication des processeurs nécessaires au calcul ainsi que des architectures nécessaires au stockage. Aujourd'hui, la question de la gestion de flux de données instantanés issus des réseaux sociaux, de l'utilisation des téléphones mobiles relève selon certains auteurs d'un nouveau paradigme scientifique (Kitchin, 2014b), largement guidé par les données traitées de manière intensive. Sans entrer dans un quatrième paradigme scientifique, nous identifions toutefois dans notre cas des similitudes entre certaines propriétés dédiées aux *big data* d'un côté (Kitchin, 2014b), et les caractéristiques de la base de données FRANcE d'un autre. L'objectif n'est pas d'inscrire notre base de données dans les *big data* mais de montrer en quoi les problématiques qui en sont issues peuvent être appliquées dans notre cas :

- Le *volume*, quand il faut gérer des matrices d'accessibilité entre les 36 000 communes françaises ;
- La *vélocité* : elle est toutefois relative, malgré les mises à jour possibles tant pour les données de population que pour l'évolution du réseau ferroviaire. De même, les contraintes liées à la collecte de la donnée géohistorique fait qu'elle n'est pas continue ;
- La *variété des données* : elle est plus ou moins structurée et pas forcément homogénéisée ;
- L'*exhaustivité* : elle capte l'ensemble de la population, l'ensemble du réseau ;
- La *résolution* : le niveau administratif le plus fin permet une identification par indexation, il en est de même pour les sections de réseau ;
- La *relation* : il existe des champs permettant d'adjoindre d'autres larges bases ;
- La *flexibilité* : les extensions par les champs et par la multiplication des individus sont rendues possibles par le modèle de données PONT.

Si certains auteurs se contentent de comparer les *big data* à des données qui ne peuvent entrer dans les formats d'une feuille de calcul de logiciels bureautiques (Kitchin, 2014b), le volume n'est qu'une des caractéristiques des grands ensembles de données spatio-temporelles. Elles rencontrent un vif succès aujourd'hui dans les colloques de géographie, tels l'AAG ou l'ECTGQ[20], où des sessions spéciales sont organisées de manière systématique. Aussi, la plupart des publications tendent aujourd'hui à rassembler les approches qui relèvent de la géographie quantitative, de la *GIScience* de celle qui relève de l'analyse spatiale (Cuyala, 2014). Malgré cette propension à les rassembler, on pourra identifier par la suite quelques nuances.

Ainsi, face aux enjeux que nous venons de soulever, Mark Gahegan identifie depuis la fin des années 1990 une *« tentative consciente de faire avancer la recherche par le retour à l'analyse géographique, que ce soit avec ou sans l'outil SIG »* (Gahegan, 1999). Il y voit là l'agenda de la *GeoComputation*, dont l'objectif est d'enrichir la discipline géographique par une boîte à outils et de méthodes pour modéliser et analyser des problèmes géographiques complexes, parfois non-déterministes. L'effort est alors à accomplir dans la double perspective de notre discipline et celle des sciences informatiques. En dépit des intentions clairement énoncées, la définition de la *GeoComputation* est en revanche sujette à de nombreux débats, qui persistent à chaque rassemblement. Symptomatique des vocables entre Etats-Unis et Angleterre, sur les treizes rassemblements organisés par le *Centre for Computational Geography*, les deux tiers l'ont été dans un pays du Commonwealth. Ils ont lieu en alternance avec la conférence internationale des *GIScience*, montrant à la fois les proximités et les frontières floues qui les entourent.

Plus qu'une agglomération d'outils et de logiciels, Paul Longley définit la *GeoComputation* comme *« ce que ses chercheurs et pratiquants font, rien de plus, rien de moins »* (Longley, 1998). Derrière ce consensus, se cachent des définitions complémentaires : Stewart Fotheringham la définit comme celle qui se réfère à des *« méthodes d'analyse spatiale dans lesquels l'ordinateur y joue un rôle pivot »* (Fotheringham, Brunsdon et Chariton, 1998) alors que Stan Openshaw la définit comme l'application de calculs à haute performance (HPC) mobilisant l'intelligence artificielle pour résoudre des problèmes géographiques complexes ou invisibles (Ehlen, Caldwell et Harding, 2002 ;

[20] AAG : Annual Meeting of the Association of American Geographers. ECTQG : European Colloquium in Theoretical and Quantitative Geography.

Openshaw et Abrahart, 1996), à l'origine d'un nouveau paradigme scientifique. A l'opposé, Helen Clouclelis voit dans cet ensemble d'outils et de méthodes *« une application éclectique de méthodes et techniques computationnelles pour dresser des propriétés spatiales, pour expliquer un phénomène géographique et pour résoudre des problèmes géographiques »* (Couclelis, 1998). Cette pluralité des définitions nous permet de définir dans un premier temps la *GeoComputation* comme la rencontre de la géographie, l'analyse spatiale et l'informatique.

4.1.2. **… à la *GeoComputation* dans une perspective géohistorique**

Nous proposons ici de définir dans quelle mesure et jusqu'où nous inscrivons notre travail dans la *GeoComputation*. Le rapprochement est d'abord aisé par le préfixe *Geo*. De la même manière que l'information spatiale est au cœur du raisonnement géohistorique, elle l'est aussi dans cette démarche-là. La base de données FRANcE y entre alors naturellement par l'information localisée, et *a fortiori* complétée par une information temporelle. Dès lors, nous souhaitons aussi mettre en place une démarche géo-méthodologique : en identifiant le besoin de se démarquer du seul outil SIG, la démarche par la *GeoComputation* paraît donc se justifier. Pour cela, nous nous appuyons sur les trois piliers définis plus haut, ainsi que l'exemple d'un module d'enseignement destiné au cycle master à l'Ecole Polytechnique de Lausanne[21], qui définit son objectif par l' *« analyse d'informations géo-référencées qui requiert l'utilisation de méthodes informatiques »*.

La question de recherche que nous avons développée fait naturellement appel à l'analyse spatiale, qui permet une *« étude formalisée de la configuration et des propriétés de l'espace des sociétés »* (Lévy et Lussault, 2003). Elle est à rapprocher des méthodes statistiques. Il faut cependant associer différentes méthodes mobilisées, selon la définition donnée par Denise Pumain dans l'encyclopédie en ligne *Hypergéo*. Elle identifie un rôle actif ou passif de l'espace dans l'identification des structures et propriétés *(Table 4. 1)*, qui justifie alors notre recours à la *GeoComputation*, quand nous utilisons des méthodes dans lesquelles l'espace joue surtout un rôle actif.

	Rôle passif de l'espace →	***Rôle actif de l'espace***
Structures de l'espace	Statistiques de l'espace	Statistique spatiale
Propriétés de l'espace	Modélisation *dans* l'espace fixé a priori	Modélisation *de* l'espace à partir d'informations contenues dans des matrices d'interaction

D'après PUMAIN D., « Analyse spatiale », *Hypergeo* © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 4. 1. L'analyse spatiale en géographie

[21] http://edu.epfl.ch/coursebook/fr/geocomputation-ENV-546

Les principes de la *GeoComputation* peuvent être replacés dans ces distinctions, avec un primat dans le rôle actif de l'espace. Il s'agit alors d'analyser directement les informations géocodées : de nombreux logiciels SIG ont implémenté des fonctions rendant compte de la statistique spatiale. Elle est définie par la définition de variables intrinsèquement spatiales, qui permettent notamment de qualifier les distributions géographiques, venant appuyer les hypothèses de la loi de Tobler, selon laquelle l'intensité des relations est fonction de la proximité des objets analysés. Parmi elles, les indicateurs LISA (*Local Indicators of Spatial Association*) sont ceux qui permettent d'explorer des données spatiales : ils permettent de mesurer le caractère non-stationnaire et hétérogène d'une variable dans l'espace (Anselin, 1995). Ils sont globaux (Indice de Moran), mais leur décomposition locale permet d'identifier des structures de *clustering*, que l'on peut traduire en français par grappe. Par ailleurs, la modélisation de l'espace à partir d'informations contenues dans des matrices d'interaction permet également de rendre compte de l'espace, à condition qu'elles décrivent un espace hétérogène et anisotrope (Haggett, Frey et Cliff, 1977) : c'est dans cette acception que nous replaçons le concept d'accessibilité, où il s'agit de complexifier cette information dans ce chapitre. Pour autant, la prise en compte de l'information temporelle doit aussi parfois s'affranchir d'un rôle actif de l'espace dans une modélisation fixée *a priori*. Pour cela, nous nous forçons toutefois de maximiser l'information contenue dans les matrices d'interactions.

La construction des matrices d'interaction, pour rendre compte de l'évolution des accessibilités sur le temps long, se heurte à la complexité des déplacements ainsi qu'au dimensionnement de la base de données FRANcE. C'est ici que nous identifions les enjeux qui lient la géographie et l'informatique. On rejoint ici l'objectif du cours décrit plus haut qui lie l'analyse d'une information géo-référencée à des méthodes informatiques sophistiquées, où *« l'implémentation d'algorithmes est capable de rendre compte de la localisation d'un objet par rapport à ses voisins, grâce au calcul à haute performance »*. Dès lors, nous verrons dans la suite du chapitre comment nous nous emparons de la théorie des graphes, définie comme théorie informatique et mathématique. En dépit de l'utilisation détournée du graphe dans l'extension *Network Analyst* du logiciel ArcGIS, nous misons sur l'utilisation de logiciels dédiés à la structuration et l'analyse de graphe, quand il est porteur d'une information spatiale, sans nous contraindre toutefois à devoir systématiquement utiliser des informations géocodées. Ainsi, Mei Po Kwan a montré comment la croissance des performances informatiques a permis de nouvelles investigations dans les champs de l'accessibilité (Kwan, 2004). Il montre l'évolution des capacités dans la prise en compte de trajectoires individuelles à des échelles locales. Ainsi, en 1967, un réseau a été complètement implémenté manuellement avec 100 nœuds et 246 arcs. Le travail précurseur de Lenntorp est certainement le premier réseau de transport digitalisé pour l'étude de configurations dans l'espace-temps (Kwan, 2004 ; Lenntorp, 1976). Depuis, l'auteur a développé des algorithmes susceptibles d'être implémentés dans des logiciels commerciaux, mais le développement des ressources en libre accès, qui a déjà retenu notre attention dans les humanités numériques paraît ici être une piste incontournable.

Le rapprochement vers les sciences informatiques vient ici en complément de notre inscription dans les humanités numériques : la *GeoComputation* participe alors au dialogue avec d'autres disciplines (Singleton, 2014). Pour autant, s'extraire d'un environnement SIG suggère un changement de langage, qui a toutefois été anticipé lors la modélisation conceptuel des données opérée dans le troisième chapitre. Prenant au pied de la lettre la nécessaire collaboration entre chercheurs, le recours aux logiciels libres et ouverts paraît incontournable dans la *GeoComputation*

(Bivand, 2011). De manière concomitante, une organisation internationale *Open Geospatial Consortium Inc.* a été créée pour répondre à l'impératif d'interopérabilité des SIG. Mais étrangement, l'un de ses premiers protocoles est d'adapter les fichiers *shapefile*, qui est devenu *de facto* le fichier de formes standard des SIG, développés au départ par ESRI puis largement adoptés dans les logiciels libres comme propriétaires. Ces entreprises vers le monde de l'accès libre relève selon nous d'une démarche croissante dans la géographie théorique et quantitative, dans laquelle les chercheurs, pour dépasser les limites des SIG classiques, sont amenés à introduire dans leur propre recherche la pratique du codage informatique (Rey, 2009). Par la suite, le partage du code participe à l'évaluation de la reproductibilité d'une méthode de digitalisation et de collecte d'informations dans le cadre des SIG-H (Gregory et Ell, 2006), de validation et d'analyse dans le cadre de la *GeoComputation* (Bivand, 2011). Ainsi, selon le même auteur, la géographie de la santé, l'épidémiologie, l'écologie, la gestion des ressources en eau sont les domaines qui ont le plus profité du développement de logiciels en open source. La géohistoire peut y prendre toute sa place également.

Nous défendons toutefois ici l'utilisation de SIG-Serveur : l'objectif n'est pas de développer un nouveau logiciel mais d'utiliser les capacités d'un système de gestion de bases de données pour y ajouter les composants de l'information géographique *(Figure 4. 1)*, et s'adapter à des logiciels déjà existants. Son principal intérêt réside dans la capacité de stockage qui se trouve multipliée sur un serveur, plutôt que sur un ordinateur utilisateur, les données pouvant être partagées par ailleurs par plusieurs utilisateurs, reprenant les propriétés d'un système de gestion de bases de données. L'intérêt passe aussi par l'absence d'une interface utilisateur, qui participe à la diffusion sans logiciel prérequis et par des performances accrues, même si l'interfaçage est toujours possible avec des logiciels SIG, *SIG-Desktop*, pour la visualisation des données. PostGIS est aujourd'hui la librairie la plus utilisée, par l'IGN par exemple, parce qu'elle est selon les auteurs aujourd'hui la plus complète et totalement libre (Steiniger et Hunter, 2013), à l'intérieur du SGBD PostgreSQL. L'interopérabilité est permise par les librairies géospatiales : GDAL est la plus répandue. Elle permet de lire l'information spatiale dans presque tous les types d'infrastructures (Bivand, 2011 ; Steiniger et Bocher, 2009). Ces librairies permettent non seulement le transfert de fichiers vers le serveur, mais peuvent aussi permettre d'utiliser les données du serveur dans le cadre d'autres outils.

Dès lors, les premières étapes de la construction de notre instrument de recherche reprennent les enjeux de la science de l'information géographique, que l'on adapte dans le contexte de la *GeoComputation*, parce qu'elle est selon nous la scène d'une production de connaissances géohistoriques qui concilie l'analyse spatiale, le raisonnement sur le temps long et de plus importantes capacités de calcul.

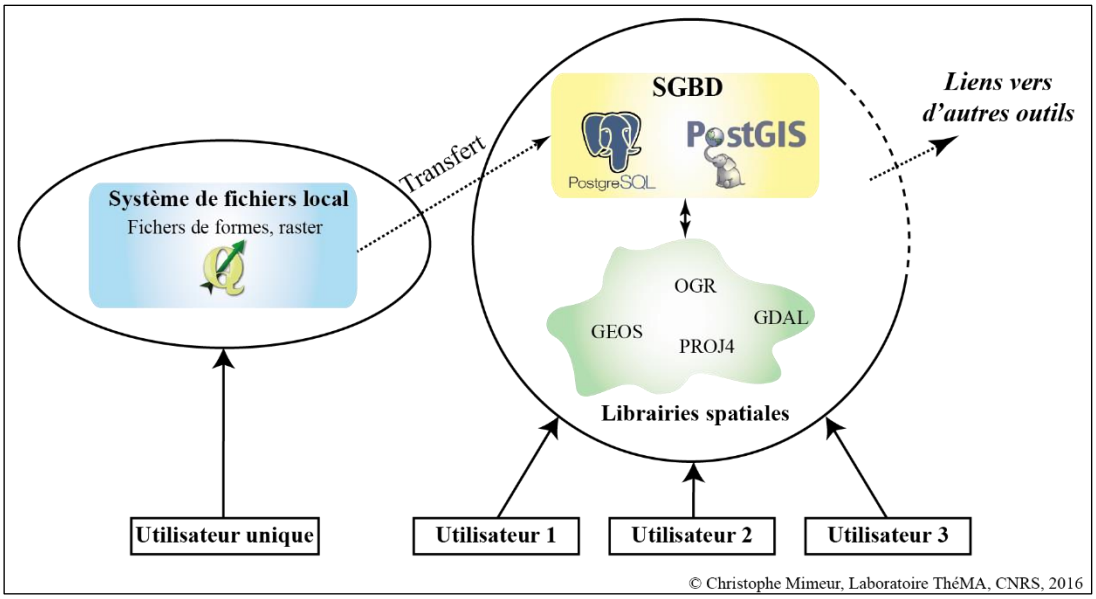


Figure 4. 1. Du SIG utilisateur au SIG serveur

4.1.3. **Une chaîne géohistorique dans un couplage modéré**

Les enjeux exprimés par un projet SIG-H ont montré comment ils n'étaient qu'en partie généraux mais aussi particuliers en fonction des objets. Ceux de la science de l'information géographique sont aussi un arbitrage entre des considérations générales et la mise en place de dispositifs particuliers. Se pose donc la question du couplage (Bivand et Neteler, 2000 ; Thévenin, 2010) entre ce qui relève strictement du SIG et ce qui n'en relève pas : dans notre questionnement géohistorique, il s'agit d'évaluer la place que l'on donne au SIG dans la construction de l'instrument de recherche *(Figure 4. 2)*.

Le couplage peut être faible (*Loose-Coupling*) : dans ce cas, on identifie des difficultés opérationnelles dues à la segmentation des opérations de traitement, qui peuvent souvent nécessiter la conversion de formats de fichier, de changer de langage informatique. A l'inverse, le couplage peut être fort (*Tight-Coupling*) : les opérations de traitement et d'analyse sont effectuées au cœur du SIG, ou inversement. Dans ce cas, les possibilités d'analyse peuvent être contraintes par les capacités du SIG, ou la prise en compte de l'information spatiale peut être contrainte par le dispositif dans lequel on implémente des fonctions SIG. En position intermédiaire, le couplage modéré (*Moderate-Coupling*) permet de tirer profit des différents dispositifs. Le lien entre les différents éléments doit alors être optimisé : on privilégie des langages et outils libres, ouverts et largement utilisés. Dans le cadre de cette thèse, ce positionnement dans la *GeoComputation* s'est d'abord concrétisé par la migration de la base de données FRANcE, jusqu'ici largement exploitée à l'aide d'ArcGIS, dans le SGBD PostgreSQL, accompagné de son extension PostGIS. Le couplage modéré permet d'opérer des choix en fonction des traitements et analyses souhaités : notre questionnement et notre approche privilégie alors ce procédé, dans une démarche abductive.

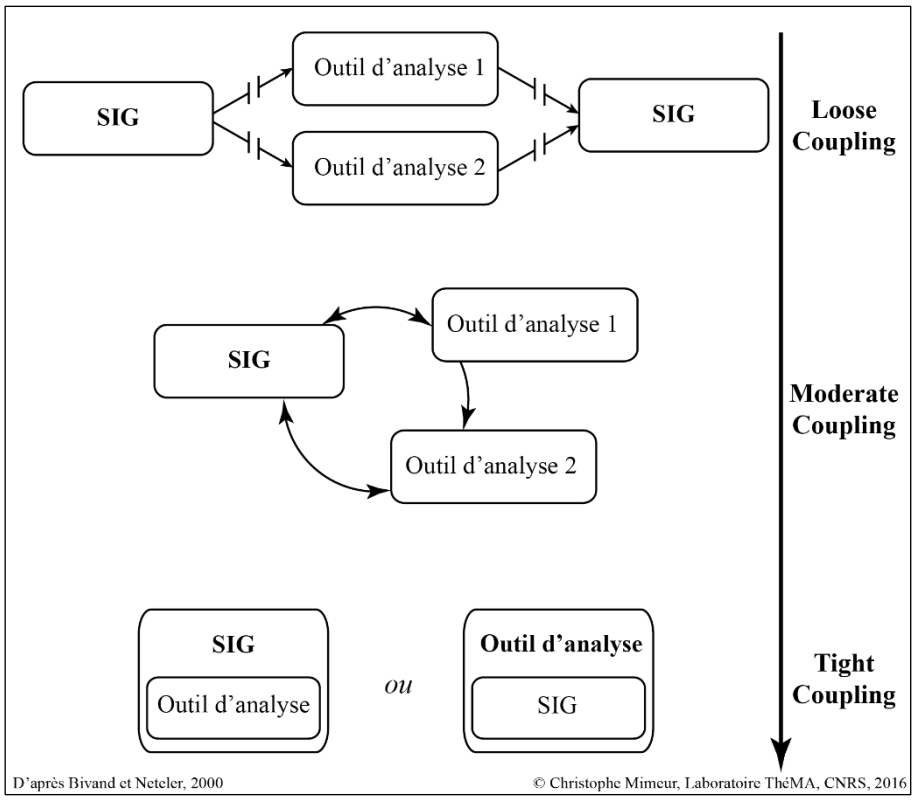


Figure 4. 2. Trois intensités de couplage entre SIG et outils d'analyse

Cette posture nous permet d'identifier ici une chaîne géo-méthodologique dans ce couplage modéré *(Figure 4. 2 et Figure 4. 3. a)*. La *Figure 4. 3* montre les quatre étapes qui associent géographie, analyse spatiale et informatique *(a)*, ainsi que le couplage modéré des différents outils qui seront utilisés dans la suite de notre travail *(b)*. Dans un premier temps, face à un volume important de données, les techniques exploratoires sont un recours ancien pour obtenir de premières informations synthétiques. L'exploration des données a été initiée dans les années 1970 par John Tukey par des méthodes statistiques (Tukey, 1977). Son application en géographie est plus récente, connue sous l'acronyme ESDA (*Exploratory Spatial Data Analysis*), mettant à contribution de nombreux indicateurs d'analyse spatiale déjà évoqués par exemple, aux côtés d'autres indicateurs plus classiques. Dans les années 1990, Stan Openshaw a contribué à de grandes avancées par le développement d'une *Geographical Explanation Machine* (GAM), illustration d'un couplage modéré qui permet un interfaçage entre SIG et capacités informatiques (Openshaw et Abrahart, 1996) qui permet de formuler des hypothèses grâce à l'identification de grappes et de distributions géographiques issues de l'analyse exploratoire des données spatiales. La **géo-exploration** est alors permise par l'interfaçage avec des logiciels SIG comme QGIS, où le dialogue entre le logiciel et le SGBD permet des représentations géographiques en fonction de requêtes spatiales et temporelles.

Dans une démarche abductive, la géo-exploration peut participer à la formulation d'hypothèses issues d'une analyse inductive des données, qui prennent place à côté d'autres hypothèses formulées par raisonnement déductif. Nous identifions alors l'étape suivante, celle d'une **géo-construction** de nouveaux objets et d'indicateurs pour se donner les moyens de vérifier ces hypothèses. Cette démarche est faite d'aller et retour avec la première, parce que les nouveaux

indicateurs peuvent être réintroduits dans la phase de géo-exploration. De plus, elle met au centre la notion d'accessibilité dans notre travail. Elle est la concrétisation de la transformation des informations contenues dans le SIG en un objet *graphe*. Parce qu'il peut être dépourvu d'une information géocodée, le graphe prend alors place dans un environnement extérieur au SIG. Dans le domaine identifié, le logiciel *R* prend en charge ce type d'objet (Beauguitte, 2012). Ce logiciel libre de traitement des données et d'analyse statistique possède un langage informatique dédié. Les possibilités de la version de base sont largement élargies par l'utilisation d'extensions, écrites dans le même langage et librement mis à disposition. Parmi elles, on ne citera pour l'instant que celles qui permettent d'exploiter un graphe et d'appliquer la théorie des graphes – *igraph*, et celles qui permet d'interfacer directement *R* avec le système de gestion de bases de données PostGreSQL – *RPostgreSQL*, celles qui permettent d'exploiter les librairies géo-spatiales – *rgdal*. Dans cette phase, les capacités de communication entre les différents éléments de notre instrument de recherche permettent de créer de nouveaux indicateurs d'accessibilité, dans les dimensionnements imposés par les échelles d'analyse.

Par la suite, le dispositif de construction laisse place à la **géo-analyse**, avec un rôle plus ou moins actif de l'information géocodée. La modélisation des relations entre transport et territoire peut se faire par la statistique multivariée ou l'économétrie : *R* dispose alors d'extensions qui permettent d'intégrer les nouvelles informations – elles peuvent être issues de matrices d'interactions décrites plus haut – dans ce type de modélisation. Ce même logiciel permet aussi de réintégrer les indicateurs et valeurs issus de ces modèles dans des analyses exploratoires, pour tester la validité des modèles et des relations. De même, la géo-analyse conduit à la formulation de modèle d'interaction. Le développement de modèles algorithmiques ou faisant appel à la simulation prend toute sa place dans cet environnement au couplage modéré (Cheng, Haworth et Manley, 2012 ; Openshaw et Abrahart, 1996 ; Singleton, 2014), dont les sorties peuvent être réintégrées dans les dispositifs décrits jusqu'à maintenant.

Enfin, placer cet instrument dans la *GeoComputation* suggère aussi des dispositifs de valorisation des données et des analyses, en résonance directe avec le positionnement géohistorique dans les humanités numériques, à travers la **géo-visualisation** (Andrienko et Andrienko, 1999). Depuis la carte figurative de Claude Minard illustrant les pertes militaires pendant la campagne de Russie sous Napoléon Ier, de nombreux développements ont permis la visualisation spatio-temporelle (Adrienko et Adrienko, 2011 ; Peuquet, 2002) par l'adéquation de données spatio-temporelles complexes tout au long de la chaîne géo-méthodologique, dans la phase d'exploration comme dans la phase de valorisation (Antoni, Klein et Moisy, 2004). L'interfaçage des données et des traitements avec le SIG peut permettre une représentation cartographique, de même que les logiciels d'analyse intègrent des représentations graphiques : elle permet alors l'identification de formes spatiales, de formes temporelles, de formes spatio-temporelles. Là encore, les passerelles entre les logiciels libres sont pensées par la base de données PostgreSQL/PostGIS. Par ailleurs, le développement de l'Internet a favorisé l'émergence du *web-mapping* (Haklay, Singleton et Parker, 2008) et participe dans ce cadre à la diffusion de données géohistoriques.

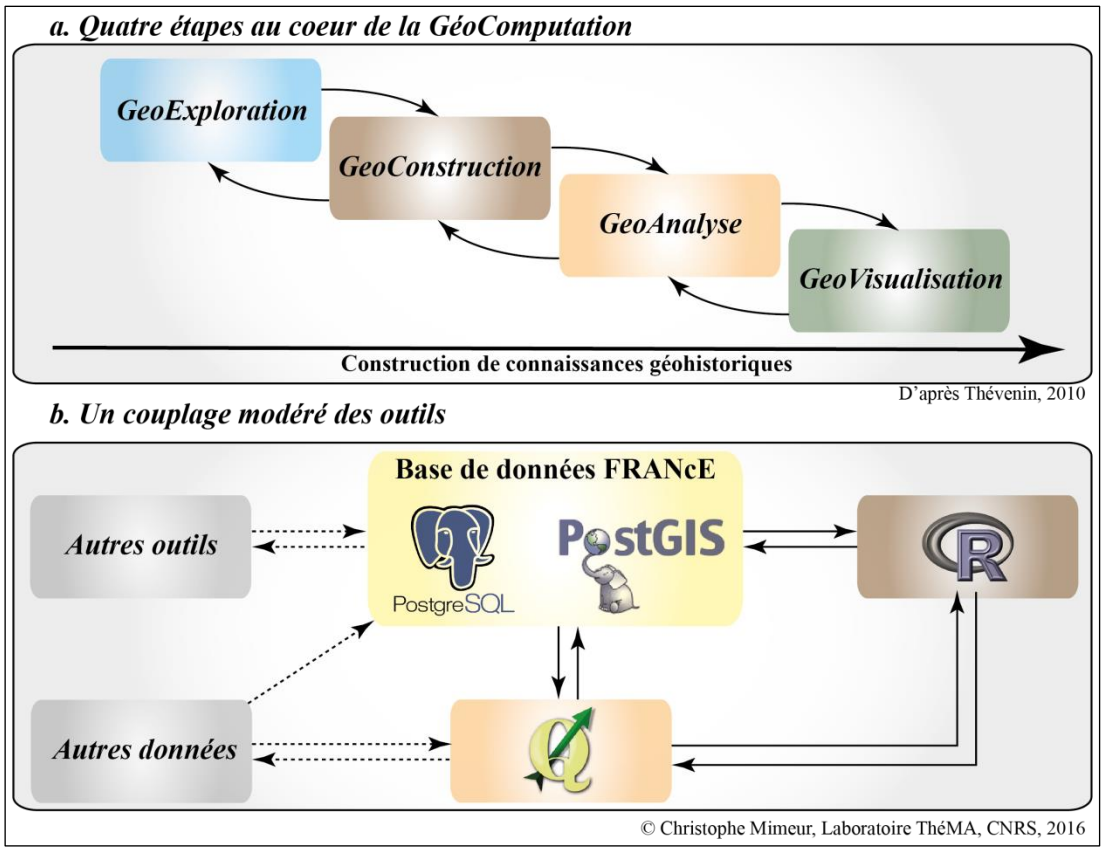


Figure 4. 3. Une chaîne géo-méthodologique dans la *GeoComputation*

Cette première section a permis de replacer la construction de l'instrument de recherche dans la *GeoComputation*, à la croisée du SIG, de l'analyse spatiale et de l'informatique *(Figure 4. 3. a)*. Elle permet de lier les enjeux de la science de l'information géographique avec les enjeux d'un questionnement géohistorique. Elle permet donc de prendre en compte l'information spatiale dans les analyses, de manière plus ou moins intense, ainsi que dans les représentations, qu'elles soient interactives ou non. Elle permet également de s'extraire, au moins provisoirement, des contraintes d'un SIG classique pour construire de nouveaux objets capables d'explorer la géohistoire du lien entre transport et territoire. Pour cela, elle s'appuie sur les croissantes capacités de calcul et de stockage pour manier des données spatio-temporelles de taille conséquente. Par ailleurs, dans notre inscription dans les humanités numériques, la *GeoComputation* offre la possibilité de s'emparer de nombreux outils libres et ouverts permettant la diversification des analyses, ainsi que de nouvelles possibilités pour le dialogue entre les disciplines et la valorisation des données *(Figure 4. 3. b)*. Dès lors, la théorie des graphes a toute sa place à la fois dans la *GeoComputation* et dans la phase de géo-construction

4.2. La théorie des graphes pour le réseau ferroviaire

Le concept de *réseau*, porteur d'une *« polysémie problématique »* (Offner, 1996), est aujourd'hui largement utilisé pour appréhender le phénomène de lien social (Lemercier, 2005). Les réseaux complexes permettent quant à eux de *« représenter les interactions entre les différents éléments d'un système »* (Queyroi, 2013). Leur analyse, leur représentation passent alors par la modélisation des réseaux en graphe, défini comme *« une structure permettant l'encodage des données relationnelles »* (Queyroi, 2013). L'analyse de réseaux est alors une démarche croissante dans les sciences humaines et sociales (Lemercier, 2005), et particulièrement en géographie (Pumain et Saint-Julien, 2010), où la théorie des graphes illustre *« les recherches sur la démultiplication des espaces géographiques, espaces de la logistiques, espaces financiers, espace des pratiques individuelles »* (Lévy et Lussault, 2003).

Cette section vise à montrer comment la théorie des graphes, comme théorie mathématique et informatique permet d'étudier la structure du réseau ferroviaire français sur le temps long, parce qu'il s'agit d'un *« remarquable outil de modélisation de situations concrètes »* (Xuong, 1992). Pour cela, elle participe à la phase de géo-construction de nouveaux indicateurs issus de la seule exploitation des données du réseau contenues de la base de données FRANcE : le graphe participe donc autant à la résolution de problèmes qu'à l'analyse de données. Pour autant, nos échelles d'analyse requièrent de s'extraire du cadre du SIG pour établir des mesures de performance du réseau (Kwan, 2004 ; Mathis, 2003) : l'objet graphe se retrouve ainsi au cœur de la démarche géo-méthodologique dans la *GeoComputation*, quand *« la durée de calcul de certains algorithmes croît très vite avec la taille du graphe, c'est un problème d'efficacité d'algorithme et de complexité de calcul »* (Mathis, 2003). Nous voyons ici comment la théorie des graphes donne de premières clés d'entrée pour l'étude de la dynamique d'un réseau pour la suite de notre travail.

4.2.1. **Les principes de la modélisation des réseaux de transport**

Pour modéliser des relations, le recours au graphe est devenu presque systématique, tant *« les graphes constituent l'outil théorique le plus utilisé pour la modélisation et la recherche des propriétés des ensembles structurés. Ils interviennent à chaque fois que l'on veut représenter et étudier un ensemble de liaisons entre les éléments d'un ensemble fini d'objets »* (Beauquier, Berstel et Chretienne, 1992). Le graphe est alors défini comme une *« figure géométrique formée d'un ensemble de points appelés sommets ou nœuds et de lignes joignants ces sommets arêtes ou liens »*[22] (Pumain et Saint-Julien, 2010). L'adoption d'un tel objet dans notre question de recherche est justifiée par les proximités entre les propriétés du graphe et les propriétés d'un réseau technique, telles qu'elles ont été définies dans la première partie de ce travail : une structure, une dynamique, une logique (Dupuy et Offner, 2005). Le graphe reprend aussi les concepts de connexité, de connectivité (Pumain et Saint-Julien, 2010) que nous avons évoqués dans la construction du

[22] Par convention, dans la suite de ce manuscrit, nous privilégierons l'utilisation des vocables nœuds et liens, ainsi que leur traduction anglaise employée de manière conventionnelle dans les articles scientifiques anglo-saxons : *vertices* pour les nœuds, *edges* pour les liens.

« mythe » de l'effet structurant. Pascale Dancoisne a été précurseur en France par l'utilisation de la théorie des graphes pour l'analyse des réseaux ferroviaires (Dancoisne, 1984). Pour autant, la modélisation des réseaux pose les questions de la transformation du réseau en un graphe, de son analyse, ainsi que de la représentation des réseaux par les graphes (L'Hostis, 2003), rejoignant les considérations plus générales de la *GeoComputation*.

Le statut de la théorie des graphes reste toutefois ambigu tant son utilisation est aujourd'hui pluridisciplinaire (Commenges, 2013) : elle est utilisée par les mathématiques, l'informatique, la physique, la sociologie ou la géographie. La théorie des graphes permet de *« résoudre un certain nombre de problèmes classiques comme les chemins les plus rapides entre une ou plusieurs origines et destinations, la capacité d'un réseau »* (Mathis, 2003). Ces problèmes paraissent alors aussi concrets que vastes. La théorie des graphes peut être considérée dans notre cas comme un formalisme, toutefois indispensable pour la construction de nouvelles connaissances géohistoriques. Nous reprenons ici deux exemples emblématiques de problèmes récurrents posés à l'objet *graphe (Figure 4. 1)*. On cite volontiers, aux prémices de la théorie des graphes et surtout parce qu'il s'agit d'un problème de transport, le célèbre problème, très opérationnel, des ponts de Königsberg (L'Hostis, 2003), édicté par Leonhard Euler au milieu du XVIIIème siècle. Ce problème mathématique consiste à créer une promenade qui travers une et une seule fois chacun des sept ponts de la ville aujourd'hui nommée Kaliningrad. Le second problème classique est celui de la coloration de la carte de Guthrie en 1852 : l'utilisation du graphe lui permet de colorier la carte des Etats Européens de quatre couleurs différentes sans que deux régions adjacentes aient la même couleur (Queyroi, 2013).

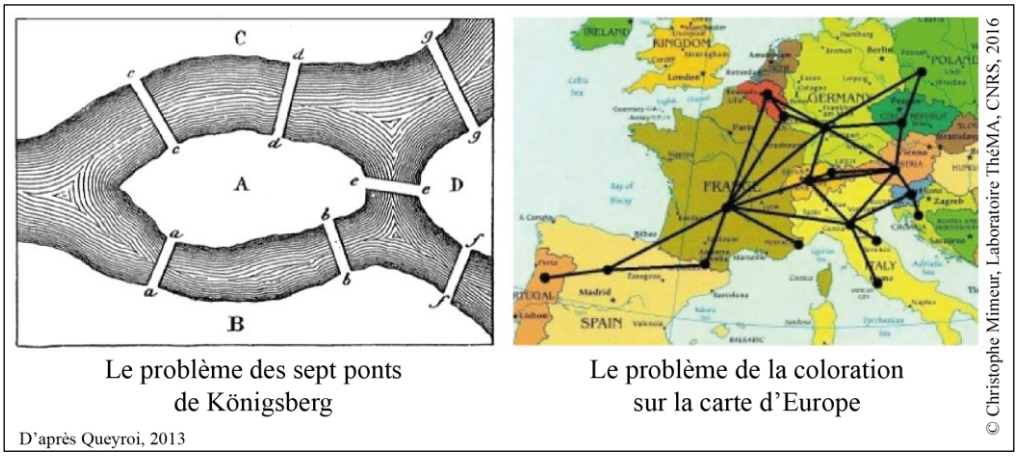


Figure 4. 4. A l'origine de la théorie des graphes …

Le graphe est un objet mathématique. L'étude de leur lien fait alors partie d'une branche des mathématiques appelée *topologie*. Le cadre théorique des graphes permet ainsi la réalisation d'algorithmes : le plus célèbre d'entre eux est sans doute l'algorithme de Dijkstra, qui résout le problème du plus court chemin (Dijkstra, 1959). Il trouve écho dans notre questionnement parce qu'il permet d'appréhender d'abord l'accessibilité comme un simple indicateur de la distance (Mercier, 2008). Malgré son caractère simple, nous verrons comment sophistiquer cette mesure en ajoutant d'autres variables et caractéristiques des réseaux. La détermination des plus courts chemins

entre chaque nœud (ou lieu) d'un réseau permet d'appréhender la séparation spatiale des activités humaines d'une part (Morris, Dumble et Wigan, 1979), de s'emparer de la loi de Tobler d'autre part. Dans la théorie des graphes, cette distance peut revêtir différentes formes : la valeur des liens traduit la rugosité, c'est-à-dire l'effort consenti pour traverser ces liens. La distance kilométrique paraît naturelle, mais l'information sur les vitesses contenue dans FRANcE permet d'envisager des distances temporelles, où l'algorithme de Dijkstra recherche le temps de parcours le plus rapide sur le réseau ferroviaire pour rejoindre chaque paire de lieu. En termes économiques, cette distance peut enfin être un coût généralisé d'un déplacement, qui allie contraintes monétaires et temporelles. La solution au problème du plus court chemin se concrétise par une matrice associée de distances entre chaque paire de nœud. Dans le domaine de la géographie des transports, elle est souvent appelée matrice origine-destination, à partir de laquelle on peut définir une distance moyenne d'un nœud à tous les autres. On parle alors d'accessibilité relative. Cette mesure peut participer à la construction de matrices d'interaction, nécessaires à la production de connaissances géohistoriques. Le dimensionnement de telles matrices permet de diversifier les indicateurs mais requiert d'autant plus des capacités de calcul.

Derrière ces problèmes très opérationnels, le formalisme des graphes participe aussi à l'analyse des données que l'on y intègre car la *« théorie des graphes vise la formalisation des réseaux et la mesure de leurs propriétés »* (Lévy et Lussault, 2003). Indépendamment du contexte dans lequel il s'inscrit, s'émancipant ainsi de toute considération géométrique, la théorie des graphes est en capacité de mesurer la structure des relations entre les nœuds qui le composent et donc d'extraire des connaissances à partir d'observations contenues dans le graphe. C'est dans ce cadre que Pascale Dancoisne a revisité la chronologie du réseau ferroviaire français par les indicateurs issus de la théorie des graphes (Dancoisne, 1984). Pour autant, il faut garder à l'esprit quelques limites du formalisme des graphes, qui est *« incapables de représenter et de décrire précisément et sans hypothèses implicites un réseau de routes à partir des éléments nécessaires au calcul, par exemple, des chemins ou du flux maximum »* (Mathis, 2003). La recherche fondamentale, en mathématique surtout, s'est surtout attachée à la relation entre les sommets plutôt qu'à la description des sommets eux-mêmes (Mathis, 2003). Ainsi, nous verrons comment les sorties du graphe vont davantage venir enrichir les informations des liens entre les objets plutôt que sur les objets eux-mêmes, même si nous nous efforcerons de créer de nouvelles connaissances sur les nœuds du réseau à partir de l'analyse des graphes, parce que désormais, *« les systèmes de traitement de l'information autorisent des approches […] prenant en charge le rôle de l'élément, qui lui-même peut être éventuellement décrit en tant que graphe, ce qui ouvre la voie à de nouvelles études de la relation réseau/territoire »* (Lévy et Lussault, 2003).

Enfin, le troisième challenge que l'on identifie pour le formalisme des graphes est leur représentation. Dans les domaines de recherche le plus souvent cité, la représentation n'est que secondaire (Queyroi, 2013) : par le formalisme des graphes, *« seul importe de savoir comment les sommets sont reliés »* (L'Hostis, 2003). Le géographe y voit souvent la perte d'une partie de l'information, par rapport au résultat cartographique. Le graphe se retrouve alors confronté aux *« contraintes de la ressemblance »* (Mathis, 2003) : parce que la carte Michelin donne l'impression d'un tracé détaillé et exact d'une route, la largeur des voies n'est pourtant qu'une similitude, et n'est pas une représentation à l'échelle de la largeur réelle des routes.

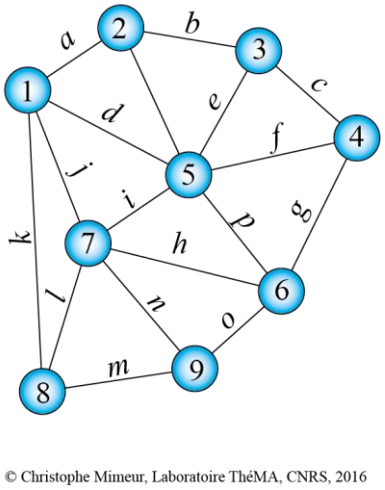
Nous retenons des principaux principes de la modélisation des réseaux de transport la représentation symbolique permise par la théorie des graphes afin de résoudre des problèmes simples au sein desquels nous plaçons l'accessibilité par les temps de parcours. Par ailleurs, la théorie des graphes participe à notre posture épistémologique quand elle permet d'explorer les données contenues dans le SIG-H pour enrichir l'information géohistorique. Elle est aussi au cœur de notre chaîne géo-méthodologique par les enjeux de représentation qu'elle met en exergue. La prochaine étape réside alors dans la transformation de l'information géohistorique en un graphe.

4.2.2. **Du SIG-H au graphe : processus de simplification et d'abstraction**

D'après Philippe Mathis, la modélisation des réseaux par les graphes est porteuse d'un paradoxe. Outre les nombreux avantages déjà décrits plus hauts, couplés avec le *« développement fulgurant des ordinateurs et micro-ordinateurs »*, la théorie des graphes a beaucoup de difficultés à *« satisfaire les deux critères essentiels de tout travail scientifique : la reproductibilité et la comparabilité en ce qui concerne particulièrement la modélisation des réseaux et la production de cartes et/ou d'images de synthèse »* (Mathis, 2003). Pour autant, nous retenons dans la perspective de notre travail l'efficacité de ce formalisme dans le domaine du calcul, permettant de répondre aux exigences du dimensionnement de notre question de recherche. La suite de la section montre aussi comment la modélisation du réseau ferroviaire français, puis de l'ensemble du territoire français obéit à une anticipation : *« pragmatisme et efficacité sur la théorisation »* (Mathis, 2003), laissant tout de même entrevoir des capacités de généralisation de la méthode proposée. Dans cette section, nous reprenons les conclusions du chapitre 3 *(Figure 3. 10)*, qui montrent une mise en connexité du réseau ferroviaire français à partir de 1860. C'est la date à partir de laquelle nous envisageons ici la transformation du SIG-H en graphe : l'objectif est alors d'obtenir **un graphe dont les nœuds sont des gares et dont les liens sont des tronçons ferroviaires**. Le graphe obtenu est alors un graphe planaire, avec la formulation mathématique :

$$G = (V, E, f) \tag{5}$$

Le graphe G est composé d'un ensemble V de sommets, d'un ensemble E de liens et où f est une application surjective $f : E \rightarrow V \times V$ selon laquelle chaque lien est associé à un couple de sommets. Le graphe obtenu est alors planaire dans la mesure où deux liens *« ne se croisent pas en dehors de leurs extrémités, c'est-à-dire un sommet du graphe »* (Mathis, 2003), qui peut être représenté sur un plan. C'est aussi un graphe simple dans la mesure où il ne possède pas d'arcs multiples ni de boucle. Ce n'est pas tant la représentation géométrique que la table décrivant f qui est utilisée pour appliquer la théorie des graphes *(Figure 4. 5)*.



E	V1	V2
a	1	2
b	2	3
c	3	4
d	1	5
e	3	5
f	5	4
g	4	6
h	6	7
i	5	7
j	1	7
k	1	8
l	7	8
m	8	9
n	7	9
o	6	9
p	5	6

Figure 4. 5. Graphe planaire

Cependant, l'information contenue dans la base de données FRANcE ne permet pas une simple transformation des nœuds ferroviaires et des tronçons en un graphe planaire. Nous identifions plusieurs difficultés dans le processus de transformation :

- Notre approche sur la longue durée suppose une évolution de la structure du réseau, donc nécessairement du graphe, à partir de trois complexifications :
  - Tous les nœuds sont des intersections sur le réseau ferroviaire mais tous les nœuds ne sont pas des gares ;
  - Un nœud peut être créé ou supprimé au cours de la période, matérialisant l'ouverture ou la fermeture d'une gare ;
  - Un lien peut être créé ou supprimé au cours de la période, matérialisant l'expansion ou la rétraction du réseau.

Ainsi, les nœuds représentant les gares sont des sommets particuliers du graphe G. Pour l'instant, nous isolons ces nœuds dans une sous-table. De plus, la constitution de la base de données FRANcE, les données relatives aux vitesses et à la population ont été calibrées à la décennie. Nous conservons cette granularité temporelle par décennie en ajoutant une variable binaire par décennie, attestant de la présence ou l'absence du tronçon dans le réseau. On peut donc réécrire l'équation (5) en ajoutant l'information d_{1860}, d_{1870}, …, d_n où $d_n: E \rightarrow \{0;1\}$ et $d_n: V \rightarrow \{0;1\}$ selon laquelle chaque lien est associé à un codage 1 ou 0 pour chaque décennie :

$$G = (V, E, f, d(E,V)_n) \quad (6)$$

- L'une des principales contributions de la base de données FRANcE dans l'étude du réseau ferroviaire est l'information sur les vitesses de déplacement sur les tronçons ferroviaires. Elle apporte une complexification supplémentaire dans le graphe : un

lien peut être parcouru de manière plus ou moins rapide selon l'évolution de la vitesse moyenne.

Le besoin de généralisation de cette information se concrétise ici par la pondération du graphe, en conservant la granularité temporelle par décennie. La fonction de pondération, dans la recherche du plus court chemin, traduit souvent la distance kilométrique d'un tronçon. Mais dans notre approche par les temps de parcours, cette fonction de pondération doit prendre en compte la distance temporelle : l'effort nécessaire pour traverser le tronçon est alors nommé *impédance*. Elle est le simple rapport entre la distance métrique et la vitesse moyenne de chaque tronçon. De cette sorte, nous pouvons nous abstenir de l'information $d(E)_n$ portant sur les liens et présente dans l'équation (6), et la retranscrire dans la fonction $\omega_{1860}, \omega_{1870}, \dots \omega_n$, qui associe à chaque lien $e \in E$ un réel $\omega_n(e)$ qui représente l'impédance. L'absence du tronçon pour une décennie se traduit donc par une valeur nulle. On obtient désormais un graphe pondéré :

$$G = (V, E, f, d(V)_n, \omega_n) \quad (7)$$

A l'aide de l'équation (7), nous sommes donc en mesure de mobiliser la base de données FRANcE pour construire le graphe G, modélisant le réseau ferroviaire français depuis 1860. La *Figure 4. 6* décompose le graphe G en un ensemble de tables qui définissent les deux composantes V et E du graphe, assorties des informations inédites issues de FRANcE. Le passage de la table des intersections à la table des nœuds passe par la définition de $d(V)_n$. La construction de la table des liens nécessite d'identifier les sommets qui constituent leur extrémité, pour obtenir *f*. Le passage des vitesses à l'impédance permet d'obtenir la table des liens avec l'information ω_n. Pour cela, un outil est directement implémenté dans l'extension PostGIS : il a l'avantage de s'accommoder directement d'informations géométriques pour générer un graphe à référence géographique (Mathis, 2003 ; Singh et al., 2015). Le module *pgRouting* permet alors de créer la topologie du réseau à partir d'une table issue de la base de données FRANcE : elle est utilisée pour construire la table relationnelle, en conservant les attributs et la géométrie du réseau, alors que la table *reseau_vertices_pgr* est créée par la fonction de *pgRouting*, permettant d'identifier les nœuds et de déterminer les extrémités de chaque lien.

La structure qu'on lui donne permet de définir des sous-graphes, dans deux dimensions : le sous-graphe peut traduire la modélisation du réseau à une décennie donnée, comme il peut traduire la modélisation d'un réseau régional ou départemental. Pour autant, le graphe *G* répond bien aux exigences de ressemblance (Mathis, 2003), du fait de la digitalisation du réseau dans un SIG-H, dans une perspective géohistorique et en partie patrimoniale. Il reproduit la géométrie exacte du réseau, créant des sommets permettant de dessiner la sinuosité du réseau ferroviaire. Mais par contre, il ne rentre pas directement dans les critères de précision et de pertinence de notre positionnement : *« pour la plupart des réseaux et notamment ceux de transport collectif, l'accès au réseau ne peut être que discontinu […]. On ne saute pas, en courant, dans un bus ou dans un train en marche, ni d'un véhicule à l'autre »* (Stathopoulos, 1997).

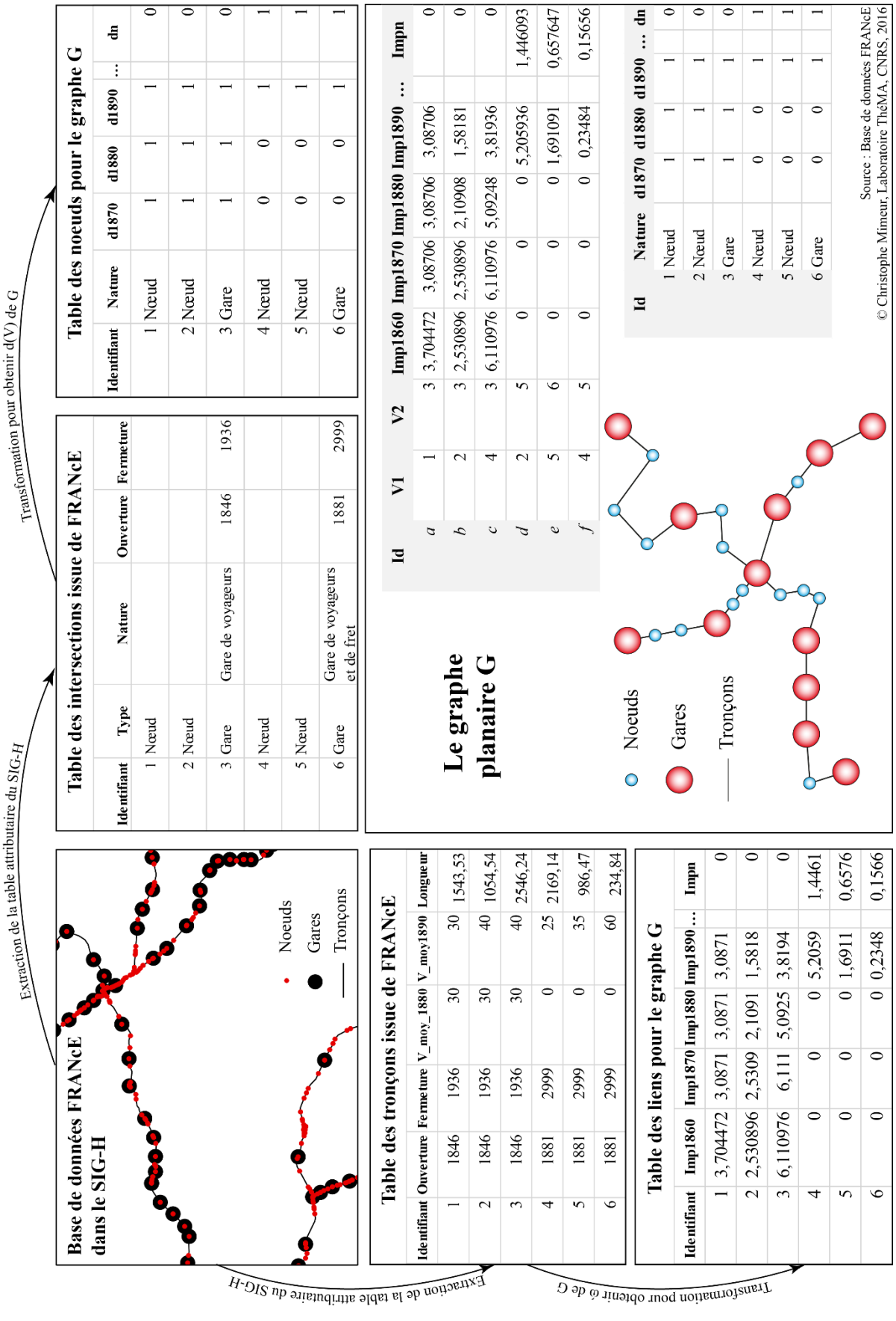


Figure 4. 6. Du SIG-H au graphe G

Selon nous, le graphe G construit à partir du SIG-H ne peut construire de connaissances géohistoriques pertinentes à partir du moment où il ne peut faire la distinction entre les points d'entrée sur le réseau et les autres, car *« à partir d'une définition matricielle du graphe, toutes les réalisations se valent, sont équivalentes en théorie des graphes »* (Mathis, 2003). Ainsi, nous proposons ici une méthode qui permet de passer de la *« ressemblance »* d'un réseau dans un graphe à une *« représentation fonctionnelle »* par le graphe (Dupuy et Stransky, 1996 ; Mathis, 2003). L'une des plus anciennes représentations fonctionnelle est la *« table de Peutinger »*, qui date du IIème ou IVème siècle : sans offrir une représentation fidèle de la réalité, la table permet une représentation symbolique qui permet de saisir la structure du réseau ainsi que les distances entre les différentes étapes *(Figure 4. 7 a.)*. La référence géographique est ici conservée pour les nœuds du graphe, mais les liens deviennent abstraits. De manière analogue, nous proposons ici de simplifier le graphe G permettant de poursuivre l'objectif évoqué en début de section : il faut pour cela supprimer les nœuds qui ne sont pas des gares, tout en gardant les propriétés d'un graphe simple et planaire *(Figure 4. 7 b.)*.

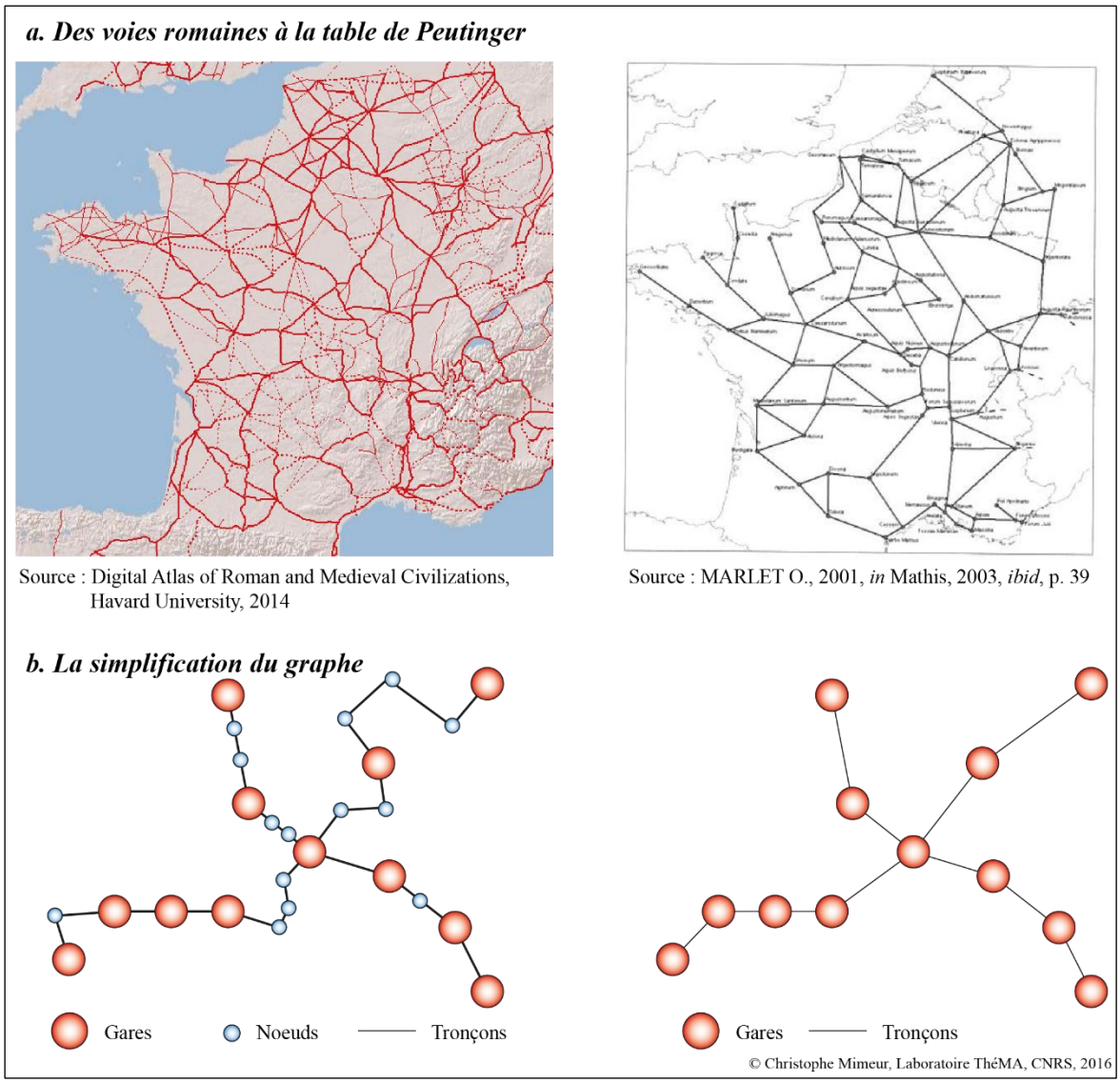


Figure 4. 7. Du graphe ressemblant au graphe fonctionnel

Cette généralisation ne se fait toutefois pas sans perte d'informations sémantiques. Pour autant, l'impédance pour parcourir le réseau de gare à gare reste une information indispensable. Elle est obtenue à partir de la construction d'une matrice de gare à gare, pour autant qu'elles soient effectivement contigües : nous supprimons de manière itérative les nœuds qui ne sont pas des gares tout en maintenant la connectivité. On a donc désormais un graphe G' qui relie les gares deux à deux *(Figure 4. 8)*. Cette hypothèse est acceptable dans la mesure où les intersections entre plusieurs lignes sont dans l'extrême majorité des gares ou alors à proximité immédiate d'une gare (Caron, 1997). Il faut toutefois noter que ce réseau peut contenir parfois des groupes de plusieurs gares toutes connectées deux à deux. C'est notamment le cas dans la périphérie de Paris : dans ces cas-là, nous rejoignons les nécessaires adaptations à faire dans un projet géohistorique, où l'intervention manuelle du géomaticien est requise pour garantir la vraisemblance du réseau, par comparaison avec le réseau géométrique contenue initialement dans FRANcE. Pour autant, l'information sur les temps de parcours reste exacte et ne compromet pas la « vraisemblance » des temps de parcours sur le réseau. Notons par ailleurs que le cas parisien requiert la création de liens *ad-hoc*, qui permettent de lier les différentes gares parisiennes : cette approximation se justifie dans l'étude du réseau à l'échelle nationale, alors que le centre de l'Etoile de Legrand est précisément Paris (Caron, 1997). La *Table 4. 2* montre comment le passage du graphe G ressemblant au graphe G' fonctionnel permet au graphe de s'insérer dans les critères de pertinence, nécessaires à la construction de nouveaux critères de précision. Le nœud parisien explique la différence de 1 entre le nombre de gares et le nombre de nœuds dans le graphe simplifié. L'opération menée montre que la simplification a diminué le nombre de nœuds et de liens d'environ 85 % en 1860. Par ailleurs, l'écart entre les décennies justifie le recours à cette méthode parce que les perturbations liées à la sinuosité du réseau ne sont pas stables entre les décennies, et viendraient donc fausser les analyses sur la structure du réseau, qui prennent précisément en compte ces caractéristiques primaires du réseau.

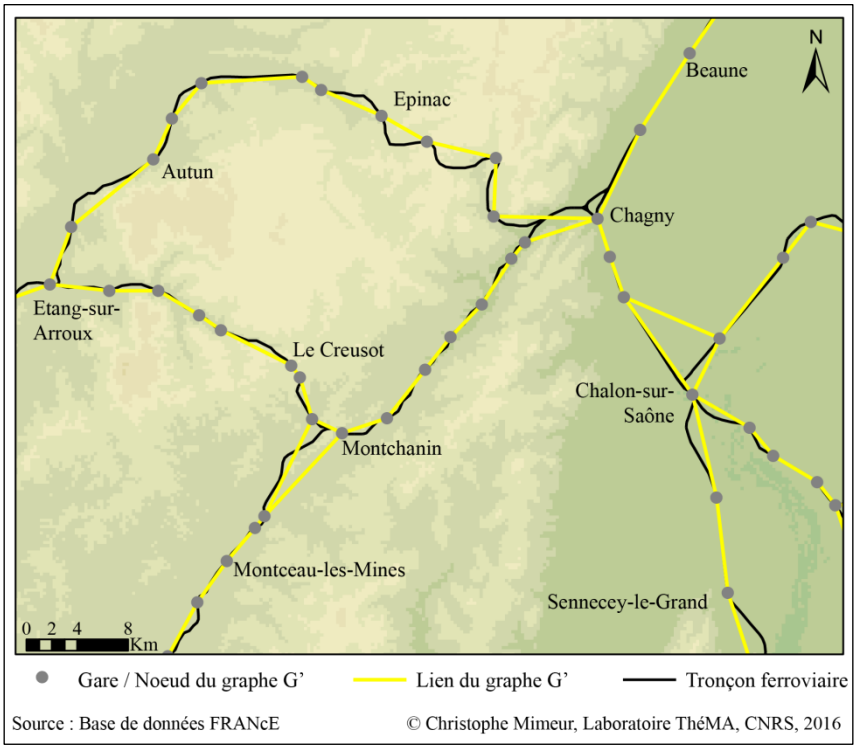


Figure 4. 8. Un sous-graphe de G' de 1880

Date	Nombre de gares	Graphe issu de FRANcE		Graphe simplifié			
		Nombre de nœuds	Nombre de liens	Nombre de nœuds	*Ecart*	Nombre de liens	*Ecart*
1860	1 713	11 467	11 515	1 714	*-85,05%*	1 807	*-84,31%*
1870	3 092	19 545	19 656	3 093	*-84,17%*	3 321	*-83,10%*
1880	4 716	26 356	26 599	4 717	*-82,10%*	5 202	*-80,44%*
1890	6 738	32 380	32 868	6 739	*-79,19%*	7 572	*-76,96%*
1900	8 266	35 422	36 076	9 267	*-73,84%*	9 308	*-74,20%*
1910	9 808	37 861	38 681	9 809	*-74,09%*	11 051	*-71,43%*
1920	10 806	39 304	40 262	10 807	*-72,50%*	12 184	*-69,74%*
1930	10 903	39 757	40 738	10 904	*-72,57%*	12 316	*-69,77%*

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 4. 2. La simplification du graphe en chiffres

A partir de l'abstraction du réseau en objet graphe, nous avons désormais les moyens de conduire des analyses propres à la théorie des graphes, participant à la création de nouvelles connaissances géohistoriques, s'attachant à la structure du réseau, ainsi qu'à son évolution sur le temps long, en exposant les apports mais aussi les limites de telles analyses.

4.2.3. **Des indices globaux aux indices locaux dans une démarche géohistorique**

Si l'objet réseau peut être étudié en lui-même, son analyse peut aussi participer à la réflexion sur le fonctionnement et l'organisation des espaces. Structure, dynamique, logique (Dupuy et Offner, 2005) : dans quelle mesure la théorie mathématique des graphes permet-elle d'appréhender ces trois propriétés du réseau ? Et surtout, sont-elles pertinentes dans le cadre de réseaux territoriaux ? Même si *« la densité reste la seule véritable mesure globale et spatiale pour décrire un réseau »*, elle fait cependant *« implicitement référence, en tant que moyenne, à la notion d'homogénéité »* (Genre-Grandpierre et Frankhauser, 1998). La *Figure 3. 13 (p.125)* montre ce processus d'homogénéisation, mais seulement d'un point de vue métrique, à une échelle largement agrégée. Jusqu'à maintenant, les analyses menées dans le précédent chapitre ont permis de préciser les rythmes de croissance du réseau ferroviaire à partir des différentes classes de vitesses entre 1860 et 1930 : ils laissent donc penser qu'à l'instar de nombreux réseaux de transport et indépendamment des échelles, une logique hiérarchique, par la recherche d'économies d'échelle, participe à une accentuation de la hiérarchisation des voies de chemin de fer. Ainsi, nous explorons ici la capacité de la théorie des graphes à produire des indicateurs capables de caractériser *« une desserte non-homogène, une hiérarchie d'espaces non desservis […] à toutes les échelles »* (Genre-Grandpierre et Frankhauser, 1998). Nous cherchons donc à produire des indicateurs structurels, morphologiques

et hiérarchiques du réseau ferroviaire sur le temps long. Ils participent à *« l'enrichissement sémantique de ces données vectorielles »* (Hamaina, Leduc et Moreau, 2012).

Dans ce travail, nous faisons le choix de ne pas utiliser les indicateurs globaux de graphe, dans la mesure où des indicateurs aux valeurs similaires peuvent décrire deux réseaux à la forme très différente (Pumain et Saint-Julien, 2010). Gabriel Dupuy a proposé d'expliquer ces limites en rappelant la philosophie dans laquelle s'inscrit la construction de ces indicateurs (Dupuy et Crews, 1991) : selon lui, elle répond à une optique de développement territorial, dans laquelle Kansky souhaitait étudier la conception du système des *highways* américaines depuis les années 1960. Accompagné de William Garrison, il leur fallait montrer que le développement des autoroutes, aux dépens du réseau ferré, allait être bénéfique pour toutes les villes américaines (Garrison, 1990). Plutôt que de parler d'objectivation mathématique, Gabriel Dupuy estime que ces indicateurs participent à une théorie de l'accessibilité territoriale par les réseaux. Selon nous, ces indicateurs servent à décrire une structure et une morphologie d'un réseau en termes de circuité, attestant des potentialités de parcours le long d'un réseau : il s'agit alors de la définition d'une morphologie *fonctionnelle* (Genre-Grandpierre et Frankhauser, 1998), à travers ces capacités connectives, dont l'objectivation spatiale reste en revanche imparfaite, parce qu'elle n'est pas observable à l'échelle des nœuds et des liens du graphes.

Au-delà du simple mécanisme de croissance du réseau, par sa structure et sa morphologie, ce sont les mécanismes de diffusion et de rétractation dans l'espace qui ressortent de notre problématique. La complexification de l'information géohistorique requiert la construction d'indicateurs à des niveaux d'agrégations successifs, compte tenu des échelles spatiales et temporelles de ce travail. Les indices locaux s'annoncent alors pertinents pour qualifier un élément du graphe en rapport avec les autres, pouvant participer à l'explication de *« caractéristiques socio-économiques changeantes des territoires connectés »* (Ducruet, 2010a). Par ailleurs, si l'analyse des morphologies fonctionnelle et géométrique des réseaux peine à étudier la hiérarchie du réseau, la position des nœuds et des liens dans le graphe peut en revanche participer à la *« vérification de certaines lois d'organisation des réseaux »* (Ducruet, 2010b). C'est dans ces indicateurs locaux qu'il faut chercher des apports pour l'appréhension de l'accessibilité, qui permettent d'exploiter pleinement l'information inédite sur la vitesse dans la base de données FRANcE. Les indicateurs locaux de graphe permettent d'explorer la hiérarchie du réseau par les capacités circulatoires qu'il induit, à partir du graphe valué, dans deux grandes familles d'indicateurs.

Les mesures locales de voisinage visent à étudier un nœud du graphe par rapport à ces voisins immédiats et sont surtout utilisés pour l'analyse de réseaux sociaux (Ducruet, 2010b). Par ailleurs, les mesures locales d'ensemble sont en revanche souvent appliquées à des graphes planaires décrivant des réseaux de transport. Ce sont eux qui répondent classiquement au problème de localisation, par le calcul de plus courts chemins entre les différents nœuds du réseau. Ils seront donc mobilisés dans la construction des temps de parcours et des matrices d'interaction. Parmi eux, la centralité intermédiarité – *Betweenness* – retient notre attention. Elle est simplement définie comme le nombre de plus courts chemins du graphe passant par chaque sommet du graphe :

$$Betw_i = \frac{Nombre\ de\ plus\ courts\ chemins\ passant\ par\ un\ sommet\ i}{Nombre\ total\ de\ plus\ courts\ chemins} \quad (8)$$

La centralité d'intermédiarité permet d'étudier la force de polarisation d'un nœud, notamment pour mettre en évidence les stratégies de localisation des multinationales dans les villes européennes (Rozenblat, 2010). Ainsi, cette mesure permet d'appréhender quels sont les points « incontournables » du réseau, parce qu'ils sont ceux qui permettent de limiter la distance à parcourir dans le réseau, de manière globale. Appliquée au système ferroviaire de Hong-Kong, la centralité d'intermédiarité apparaît comme celle qui porte le meilleur potentiel pour caractériser l'importance d'un nœud dans le réseau, pour répondre à des problématiques d'adaptation de l'offre et de management des risques (To, 2016). Elle entre en résonance alors avec notre questionnement sur l'objectivation de la construction du réseau ferroviaire, et surtout de son évolution : est-il régi par les principes d'équité territoriale ou alors par les principes d'efficacité territoriale appliqués aux réseaux, qui chercheraient alors à minimiser la longueur du réseau par la hiérarchisation des voies.

Cette même mesure se calcule également sur les liens *v* (Girvan et Newman, 2002). Elle permet alors la définition de corridors, ceux qui canalisent les plus courts chemins :

$$Betw_v = \frac{Nombre\ de\ plus\ courts\ chemins\ passant\ par\ un\ lien\ v}{Nombre\ total\ de\ plus\ courts\ chemins} \quad (9)$$

La *Figure 4. 9* montre comment, à partir d'un graphe non-valué *(1.a.)*, l'étude de la répartition des plus courts chemins décrit une logique centre/périphérie *(1.b.)*. Dans ce cas, c'est la structure globale du réseau qui influe sur les configurations locales : dans un réseau aléatoire, d'une forme régulière, les axes qui concentrent les plus courts chemins sont situés au centre et la mesure d'intermédiarité diminue avec l'éloignement. En revanche, l'exemple donné dans *2.a.* est un graphe valué, où l'épaisseur des liens traduit la vitesse à laquelle on peut la traverser, distribuée de manière aléatoire, suivant une répartition normale sur les 112 liens de ce réseau géométrique. L'étude de la centralité d'intermédiarité des liens qui prend en compte la pondération des liens suit une autre logique que l'exemple *1.b.* La structure hiérarchique des vitesses influe directement, en plus de la structure globale du réseau, sur la répartition des plus courts chemins *(2.b.)* : il en ressort une organisation d'axes structurants, dans laquelle des corridors de passage sont identifiables. De la même manière que pour les nœuds, l'étude de la position des liens dans la structure du graphe répond à des problématiques de vulnérabilité des réseaux, mais possède selon nous un pouvoir descriptif fort, en plus d'un potentiel explicatif dans les inégalités d'accès au réseau.

L'information sur les vitesses de déplacement peut ainsi nous permettre, pour un même réseau, de mesurer les écarts entre un réseau dont la vitesse est homogène sur tous les tronçons et un réseau avec les vitesses effectives. Cette mesure peut naturellement être appréhendée sur le temps long : elle participe alors à évaluer les processus de concentration ou de diffusion des plus courts chemins, à partir de la double évolution quantitative et qualitative du réseau. Elle soulève également à ce stade de nos investigations un enjeu de représentation. Au-delà d'un semis de points complexes de 11 000 gares à l'expansion maximale du réseau, la représentation de l'intermédiarité des liens participe aux enjeux de géovisualisation. Il s'agit ici d' *« appuyer l'interprétation des*

indicateurs bruts pour lesquels le travail de décomposition des effets spatiaux et des effets réseau s'effectuait jusqu'à présent de manière exclusivement qualitative » (Gleyze, 2007). Dans une perspective géohistorique, nous verrons dans notre troisième partie les apports de cette mesure dans l'appréhension de la structure du réseau, à partir de la comparaison entre un réseau homogène et un réseau hiérarchisé d'une part, et à partir de la comparaison du réseau sur l'ensemble de notre période d'étude d'autre part.

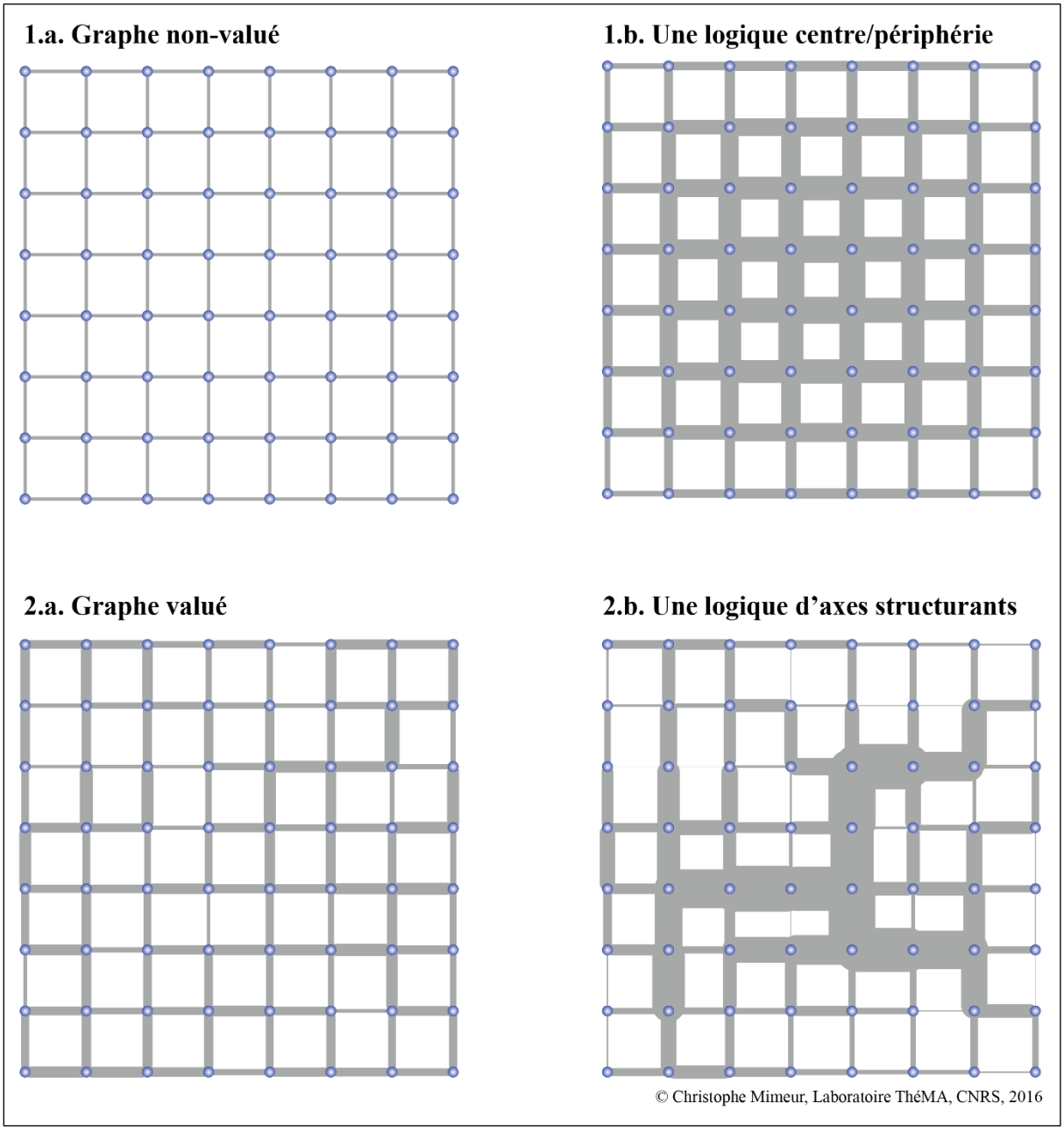


Figure 4. 9. La centralité d'intermédiarité des liens dans la théorie des graphes

Ainsi, la mobilisation de la théorie des graphes montre comment elle est efficace pour modéliser les réseaux de transport. Elle permet de formaliser les relations entretenues par les gares,

à partir de l'information sur les vitesses de la base de données FRANcE, à travers un processus de simplification qui transpose le graphe : il n'a plus des propriétés ressemblantes mais des propriétés fonctionnelles. Cette étape est nécessaire à la construction des indicateurs de graphe, permettant de décrire la structure et les propriétés du graphe. Si les mesures globales ont souvent montré leurs limites, en ne se cantonnant qu'aux aspects morphologiques, alors les mesures locales revêtent des enjeux de croisement avec d'autres variables pour analyser le rapport réseau et territoire ainsi que des enjeux de visualisation dans notre cadre géohistorique pour rendre compte de caractéristiques hiérarchiques. Cependant, le graphe G' ne permet pas pour l'instant de poursuivre nos investigations à l'échelle des 36 000 communes françaises.

4.3. L'accessibilité des 36 000 communes dans FRANcE : entre précision et généralisation

Le caractère nécessairement discontinu de l'accès au réseau ferroviaire (Stathopoulos, 1997) a été le prétexte de la construction d'un graphe fonctionnel pour le seul réseau ferré. Il garantit aussi la vraisemblance des temps de parcours et des comportements de déplacements, en lien avec les critères de pertinence et de précision. Pour autant, ces mêmes critères suggèrent l'exhaustivité de l'information spatio-temporelle pour les quelques 36 000 communes françaises, alors même que *« les liaisons rurales se heurtent longtemps encore à l'obstacle des lieux et des distances »* (Studeny, 1995). La prise en compte d'une chaîne totale de déplacement est alors un gage pour répondre à l'hypothèse de Jean Ollivro, qui détermine la maîtrise des déplacements comme un *« vecteur des inégalités »* (Ollivro, 2009).

Ainsi, jusqu'à l'essor des innovations mécaniques qui révolutionnent la vitesse, *« les possibilités tangibles de se déplacer rapidement au loin sont inexistantes et l'éventail des déplacements reste extrêmement réduit »* (Ollivro, 2009). Nous défendons donc ici une approche par les *« moyennes usuelles »* plutôt que sur les records de vitesses, qui suggère ici la construction de *« réseaux multiples »* (Flonneau et Guigueno, 2009) : l'objectif est de définir la suite de la chaîne de déplacement, et comment la prendre en compte avec le graphe G'.

4.3.1. **Cheminer vers la gare**

La constitution d'un réseau multiple, nécessaire à une bonne appréhension des déplacements de l'époque, entre en résonance avec le bornage spatio-temporel de notre étude. En nous intéressant à la phase de croissance du réseau, depuis le milieu du XIXème siècle jusqu'en 1930, le cheminement vers la gare est à concevoir dans une *« intense mobilité locale »* où *« l'homme place alors la totalité de ses contacts sensoriels sur un espace restreint »* et où *« les allers et retours à pied multiplient les passages sur des lieux connus, épuisés, identiques, sans surprise »* (Ollivro, 2000). Ainsi, même si le déplacement en diligence, notamment par les postes et messageries royales

(Verdier et Bretagnolle, 2007) se développe sur des moyennes et longues distances à partir de la fin du XVIIème siècle, nous posons l'hypothèse que sur les courtes distances, les déplacements à pied restent encore importants à cette époque. Jean-Jacques Rousseau s'en défend également quand il affirme ne dépendre *« ni des chevaux, ni du postillon. […] Et, ne dépendant que de [lui]-même, [il] joui[t] de toute la liberté dont un homme peut jouir »* (Rousseau, 1762).

Si rendre compte de l'accessibilité à une ressource par la force pédestre à l'échelle intra-urbaine a souvent été traitée dans la littérature scientifique (Geurs et Ritsema von Eck, 2001 ; Jiménez, Palau et Henneberg, 2015), des extensions à de larges échelles ont requis d'autres types de modélisation de l'accessibilité, *a fortiori* dans la longue durée. En 2007, Ian Gregory et Paul Ell montrent que dans une perspective historique, l'accessibilité à un réseau est souvent appréhendée de manière directe (Gregory et Ell, 2007 ; Thévenin, 2010). Dans cette acception, le chemin initial jusque vers la gare la plus proche n'est décrit que de manière binaire *(Figure 4. 10, a. et b.)*. Dans le premier cas, huit communes sont considérées comme ayant un accès direct au réseau alors que le second cas n'identifie que trois portes d'entrée au réseau. La description de ce chemin peut en revanche être enrichie par une information en termes de distance euclidienne *(Figure 4. 10, c. et d.)*, comme le laissent suggérer ces approches par buffer. En fonction de seuil de distances, on peut identifier des communes non desservies directement par le réseau. En revanche, ces seuils ne permettent pas d'appréhender la rugosité de l'espace.

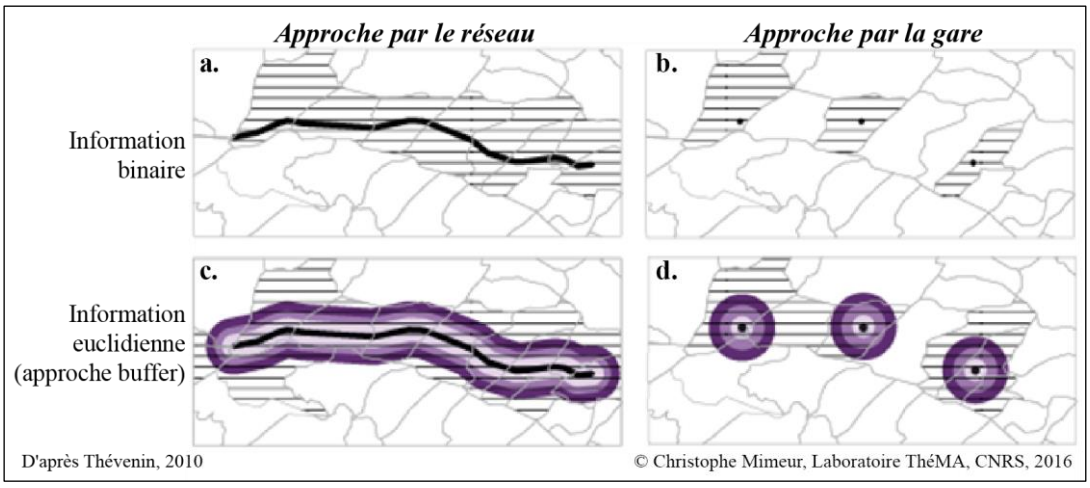


Figure 4. 10. Une approche classique par l'accessibilité directe

Or, ici, notre objectif est bien d'évaluer un temps de parcours pour l'ensemble de la chaîne de déplacement et pour l'ensemble du territoire. Il s'agit de mesurer la distance temporelle entre les centres fonctionnels des 36 000 communes vers la gare plus proche dans un premier temps, sur le temps long, à mesure que le réseau se développe et les gares se multiplient. Par ailleurs, rejoindre une commune sans gare nécessite aussi le recours à un autre segment pédestre depuis la gare de destination jusqu'à la commune finale.

La mesure de l'accessibilité physique est un compromis utilisé en géo-archéologie pour mesurer des distances temporelles entre sites (Garmy et al., 2005 ; Kaddouri, 2007). Alors que

l'accessibilité directe ne prend pas en compte l'hétérogénéité et l'anisotropie de l'espace, l'accessibilité physique permet en revanche d'intégrer les contraintes de l'environnement géographique : elles sont de premier ordre quand la force musculaire est fortement contrainte par le relief. Ces travaux consistent à pondérer les distances euclidiennes par les contraintes de relief : alors que la vitesse moyenne d'un piéton est le plus souvent fixée à 5 km/h, elle est largement réduite à mesure que la pente se dresse selon la fonction suivante *(Figure 4. 11)*.

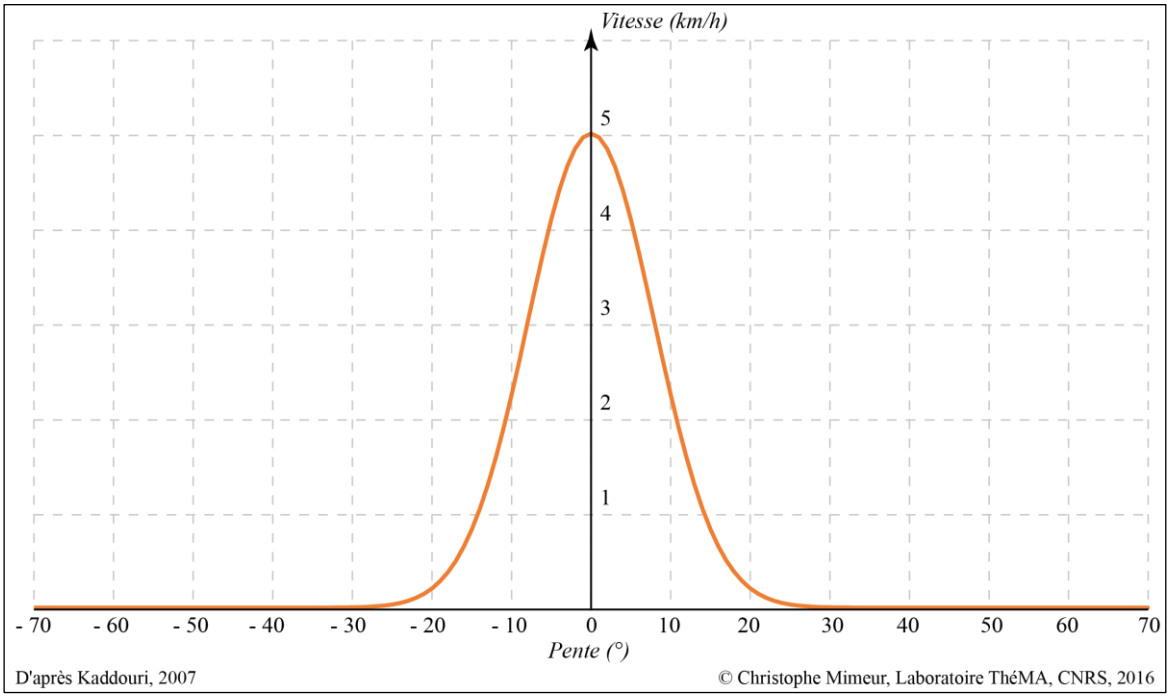


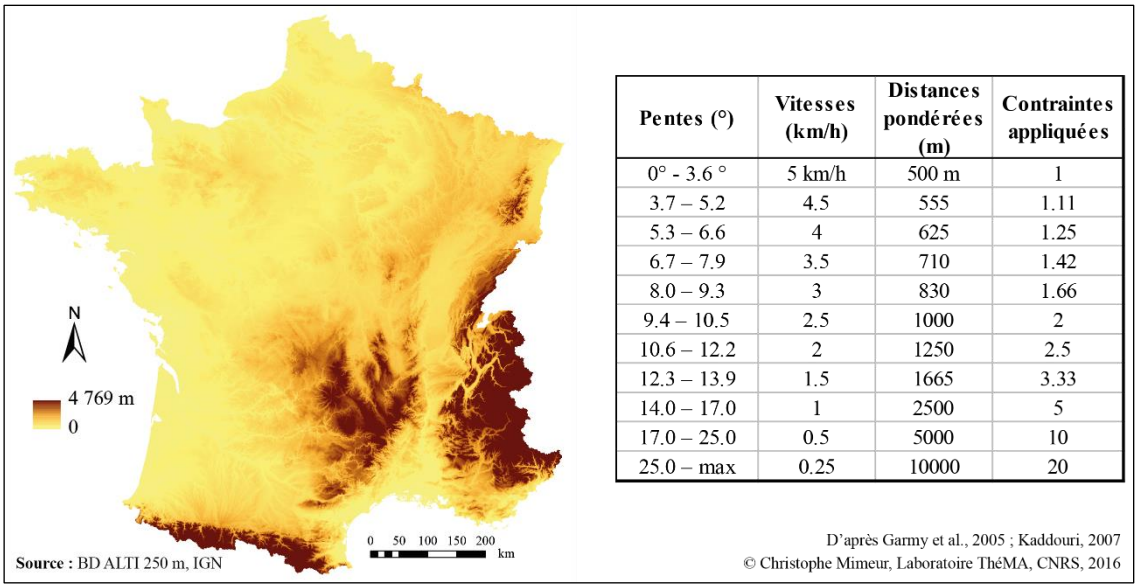
Figure 4. 11. La fonction de déplacement pour mesurer l'accessibilité piétonne aux gares ferroviaires

La suite de notre démarche consiste à exploiter cette fonction de vitesse piétonne pour caractériser le cheminement vers la gare, à partir et vers les 36 000 communes. De plus l'objectif est de complexifier davantage la chaîne complète à partir d'un arbitrage entre ces deux modes de déplacement que sont la route à pied et le chemin de fer, à partir d'une approche classiquement utilisée en archéo-géographie. Enfin, il s'agit de questionner l'arbitrage entre le recours à la source historique et la généralisation dans une perspective géohistorique à partir de d'approches plus théoriques. Les larges échelles de notre questionnement posent ici encore la question du dimensionnement des traitements géomatiques qui découlent de ces approches.

4.3.2. **Modéliser le cheminement pédestre : quelle généralisation du réseau ?**

Dans notre questionnement géohistorique, compléter la chaîne de déplacement est contraint par le manque de données. Notre positionnement a déjà proposé une généralisation de l'information sur les vitesses le long du réseau ferroviaire français. Nous proposons ici une méthodologie afin de couvrir de manière exhaustive les cheminements pédestres vers les gares ferroviaires, sur le temps long : il s'agit de créer et de valoriser des *« données spatiales pauvres sémantiquement »* (Hamaina, Leduc et Moreau, 2012) en se donnant les moyens de vérifier la vraisemblance de ces informations, à partir d'informations historiques ponctuelles, venant participer à l'évaluation d'une démarche généralisée par l'archive. Deux étapes successives permettent de conduire le raisonnement jusqu'à son but : s'assurer de la pertinence de l'utilisation de la fonction de déplacement, s'assurer de la précision de cheminements plus ou moins théoriques, à partir de terrains d'études de superficies différentes, relevant d'une situation orographique diversifiée.

La première étape propose d'appliquer la fonction de déplacement à un Modèle Numérique de Terrain (MNT). A partir du MNT *BD_ALTI* mise à disposition par l'IGN, il s'agit de calculer les coûts de chaque pixel du fichier raster, avec une résolution de 250 mètres, à l'échelle nationale. Cela revient à appliquer une contrainte de distance en fonction de la pente, selon la *Figure 4. 12*. L'objectif est alors de tracer le chemin optimal, passant par les pixels les moins pénalisés, résultant d'un compromis entre la distance euclidienne et les coûts du raster. La *Figure 4. 13* montre le tracé des chemins des communes situées dans un rayon de 20 kilomètres autour de Dijon. Leur simple observation montre un réseau en arbre ou en espalier (Pumain et Saint-Julien, 2010) : la forme des chemins semble se rapprocher de problématiques propres au bassin-versant. Pour autant, cette méthode pointe rapidement ses limites à notre échelle. Elle est utilisée en archéo-géographie à des échelles réduites, mais elle s'avère très coûteuse en capacité de calcul quand on l'étend à la France entière. Aussi, même si elle nous permet de définir les cheminements vers les gares les plus proches sur la longue durée, elle ne permet pas d'appréhender le chemin optimal combiné avec le réseau ferroviaire.



Pentes (°)	Vitesses (km/h)	Distances pondérées (m)	Contraintes appliquées
0° - 3.6 °	5 km/h	500 m	1
3.7 – 5.2	4.5	555	1.11
5.3 – 6.6	4	625	1.25
6.7 – 7.9	3.5	710	1.42
8.0 – 9.3	3	830	1.66
9.4 – 10.5	2.5	1000	2
10.6 – 12.2	2	1250	2.5
12.3 – 13.9	1.5	1665	3.33
14.0 – 17.0	1	2500	5
17.0 – 25.0	0.5	5000	10
25.0 – max	0.25	10000	20

Figure 4. 12. L'application des contraintes en fonction de la pente définie par le MNT

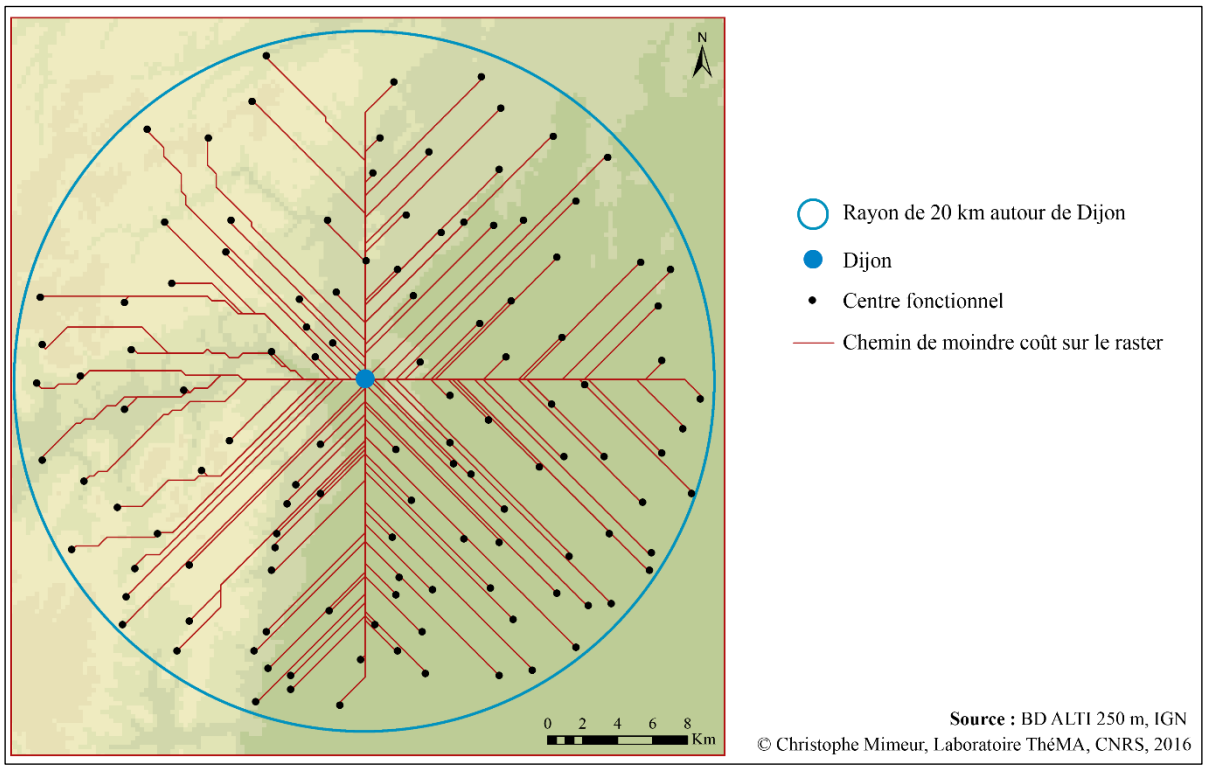


Figure 4. 13. Un réseau en espalier par les chemins de moindre coût sur le MNT

Prendre en compte un chemin vers la gare la plus optimale nécessite la généralisation des chemins de chaque centre fonctionnel vers toutes les gares. A ce stade de nos investigations, nous faisons l'hypothèse que les chemins sont porteurs d'une forte inertie, que les réseaux actuels étaient *a priori* franchissables à pied, et qu'ils sont eux aussi contraints par le relief. De manière très prudente, nous explorons la capacité d'une base de données actuelle pour tracer les cheminements entre les communes. En utilisant la base de données *BD_CARTO* de l'IGN, nous ôtons

nécessairement des tronçons étudiés, ceux qui sont de type autoroutier. Ainsi, afin de combiner les contraintes des tracés et les contraintes de relief, nous affectons les mêmes contraintes que dans la méthode précédente : une pondération est appliquée à la longueur des tronçons en fonction de leur pente moyenne, estimée à partir du même MNT. Le coût devient une impédance, comme c'est le cas pour le réseau ferroviaire. La même routine informatique permet de transformer ce réseau en un graphe, dans lequel l'algorithme de Dijkstra peut être appliqué. Deux comparaisons sont effectuées sur deux terrains différents : l'une sur la zone dijonnaise, caractérisée par un relief de côte à l'ouest et par la plaine de Saône à l'est ; l'autre aux environs de Thann, à l'entrée de la vallée de la Thur au pied des Vosges. Pour les mêmes couples d'origine et de destination (Dijon et Thann), la *Table 4. 3* compare les valeurs des plus courts chemins, en termes de distance métrique et de distance temporelle. Dans les deux zones, on constate une forte corrélation, significative à plus de 99 %, entre les chemins calculés sur le MNT et ceux calculés depuis la base *BD_CARTO*. Aussi, la variation entre les deux métriques et les deux méthodes montre des amplitudes similaires dans tous les cas.

Comparaison des cheminements pédestres sur le MNT et à partir de BD CARTO			**Coefficient de variation**	**Coefficient de corrélation *(Pearson)***
Zone de Dijon	*Distance métrique*	Chemin géométrique depuis le MNT	.38	.99
		Chemin depuis BD_CARTO	.38	
	Distance temporelle	Chemin géométrique depuis le MNT	.40	.98
		Chemin depuis BD_CARTO	.39	
Zone de Thann	*Distance métrique*	Chemin géométrique depuis le MNT	.36	.99
		Chemin depuis BD_CARTO	.36	
	Distance temporelle	Chemin géométrique depuis le MNT	.36	.89
		Chemin depuis BD_CARTO	.38	

Source : BD ALTI 250 m, IGN ; BD CARTO, IGN © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 4. 3. Comparaison des cheminements pédestres sur le MNT et à partir de BD CARTO

Pour autant, le décalage temporel entre notre période d'étude et la base actuelle nous conduit à l'approfondissement d'une telle entreprise. Il s'agit d'établir dans quelle mesure le chemin peut-être réaliste d'une situation au XIXème siècle. Pour cela, nous prenons l'exemple d'un chemin entre Etaules, au nord-ouest de Dijon et le centre de la capitale des Ducs de Bourgogne. A partir de la carte d'Etat-Major[23], levée en 1866, nous superposons les chemins tracés à l'aide du MNT d'une part, à l'aide de BD_CARTO d'autre part *(Figure 4. 14)*. Comme attendu, le premier est très géométrique, ignorant le réseau représenté sur la carte d'Etat-Major. D'ailleurs, on constate que le chemin tracé à l'aide de BD_CARTO suit en revanche une route qui existe sur la carte. A cet instant, nous postulons donc que l'appréciation des temps de parcours pédestre à l'aide d'une base de données *« réseau »* peut répondre à nos besoins pour caractériser les temps de parcours des 36 000 communes.

[23] Davantage de détails concernant cette carte seront donnés dans le Chapitre 9.

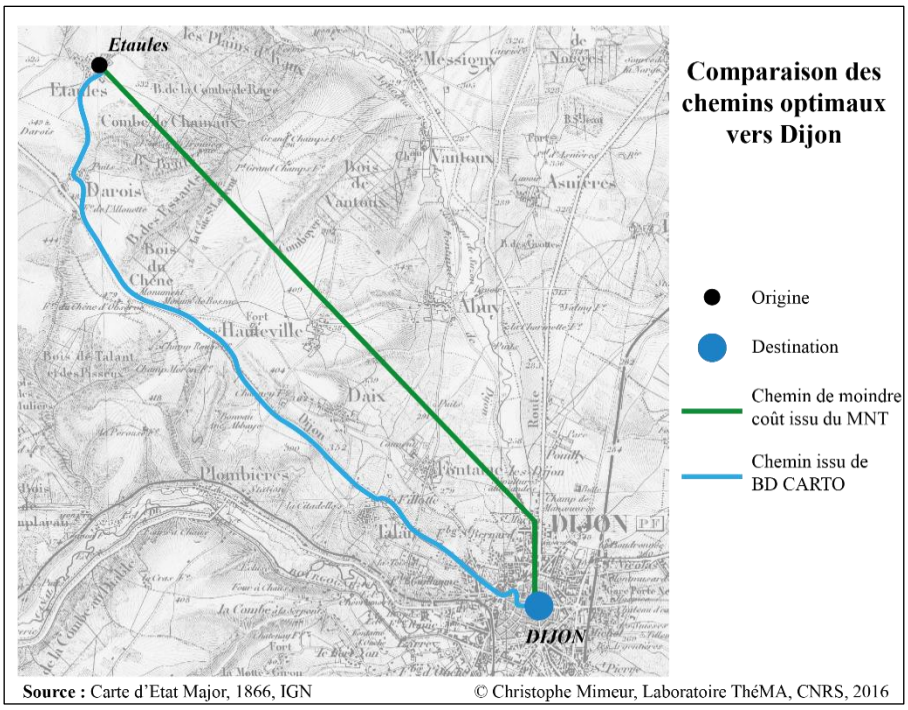


Figure 4. 14. Comparaison des chemins optimaux vers Dijon

Mais une autre complexification de notre approche réside dans la possibilité de cheminer à pied, sans emprunter le train. Ainsi, dans la période où le réseau ferroviaire est relativement lâche, il peut être bien plus intéressant de ne pas utiliser le train pour se rendre d'une commune à une autre plutôt que de contraindre notre raisonnement à rejoindre la gare la plus proche de manière indispensable. Il ne s'agit plus alors de définir les chemins les plus courts vers la gare la plus proche ou même la plus optimale pour emprunter le réseau ferré mais de définir les chemins de et vers chaque commune française. Nous posons alors l'hypothèse qu'il serait plus opportun de tracer des cheminements entre toutes les communes françaises afin de pouvoir arbitrer entre des cheminements bimodaux ou piétons. Pour autant, avec près de 1 641 000 tronçons routiers, la base issue de l'IGN nécessite des temps de calcul importants, en plus d'une intervention manuelle pour corriger les imperfections topologiques d'un réseau à une telle échelle.

A ce moment, il s'agit d'imaginer une autre méthode, plus théorique, mais qui peut difficilement être taxée d'anachronisme. Elle doit être capable de maximiser l'information entre deux communes mais aussi minimiser l'information à l'échelle nationale, de manière à calculer des chemins sur le territoire français : les études mathématiques portant sur les télécommunications sont d'un apport non-négligeable (Gleyze, 2007), quand elles visent à *« ancrer au mieux une trajectoire en ligne droite sur un semis de sommets »*. Ces travaux sont aussi connus pour l'élaboration de graphes théoriques (Chapelon, 1997). Ainsi, si l'arbre de recouvrement minimal vise à relier toutes les ressources avec une densité minimale de liens, alors la triangulation de Delaunay vise au contraire une densité maximale de liens pour ces mêmes ressources, tout en optimisant les trajectoires (Hamaina, Leduc et Moreau, 2012). Elle nous paraît la plus vraisemblable dans la mesure où les cheminements à pied bénéficient d'un champ des possibles important. Cette méthode consiste à *« découper l'espace selon le diagramme de Voronoï construit sur le semis des sommets du réseau et à identifier les sommets dont les cellules de Voronoï sont traversées par la trajectoire en question. »* (Gleyze, 2007). Alors, la définition des chemins correspond à la succession de

sommet, définissant une structure organique *(Figure 4. 15)*. Nous trouvons dans la définition de ce réseau plusieurs avantages et identifions quelques réserves. Le réseau théorique créé ne compte plus que 107 000 liens entre les 36 000 communes et permet alors des calculs bien plus aisés. Pour autant, il s'agit de s'assurer de la vraisemblance des temps de parcours calculés à l'aide de ce réseau. Pour cela, la même méthode a été utilisée pour valuer les liens : la définition des distances pondérées de chaque lien à partir du MNT, la transformation en un graphe planaire. La comparaison s'appuie sur les temps de parcours calculés sur deux zones plus larges qu'auparavant. La première est le département de la Côte d'Or (8 700 km²), avec un relief modéré et un plateau fragmenté ; la seconde est la région Pays-de-Loire (32 000 km²), à dominante plane. La comparaison s'établit sur les distances temporelles, à partir d'un calcul unipolaire vers la préfecture et d'un calcul multipolaire entre chaque paire de communes, pour le réseau théorique issu de la triangulation et le réseau issu de *BD_CARTO*. La *Table 4. 4* montre ici encore la forte corrélation entre les temps de parcours calculés sur les deux réseaux. Par ailleurs, la distribution des valeurs est similaire et les valeurs extrêmes sont reproduites dans les deux cas. Il est connu que la différence entre les plus courts chemins calculés par la triangulation de Delaunay n'excède pas 2,4 fois la distance euclidienne entre les sommets. Ici, le coefficient de corrélation de Spearman montre par ailleurs que le réseau théorique reproduit le rang des temps de parcours constatés sur une base de données actuelle, au moins pour le cas français.

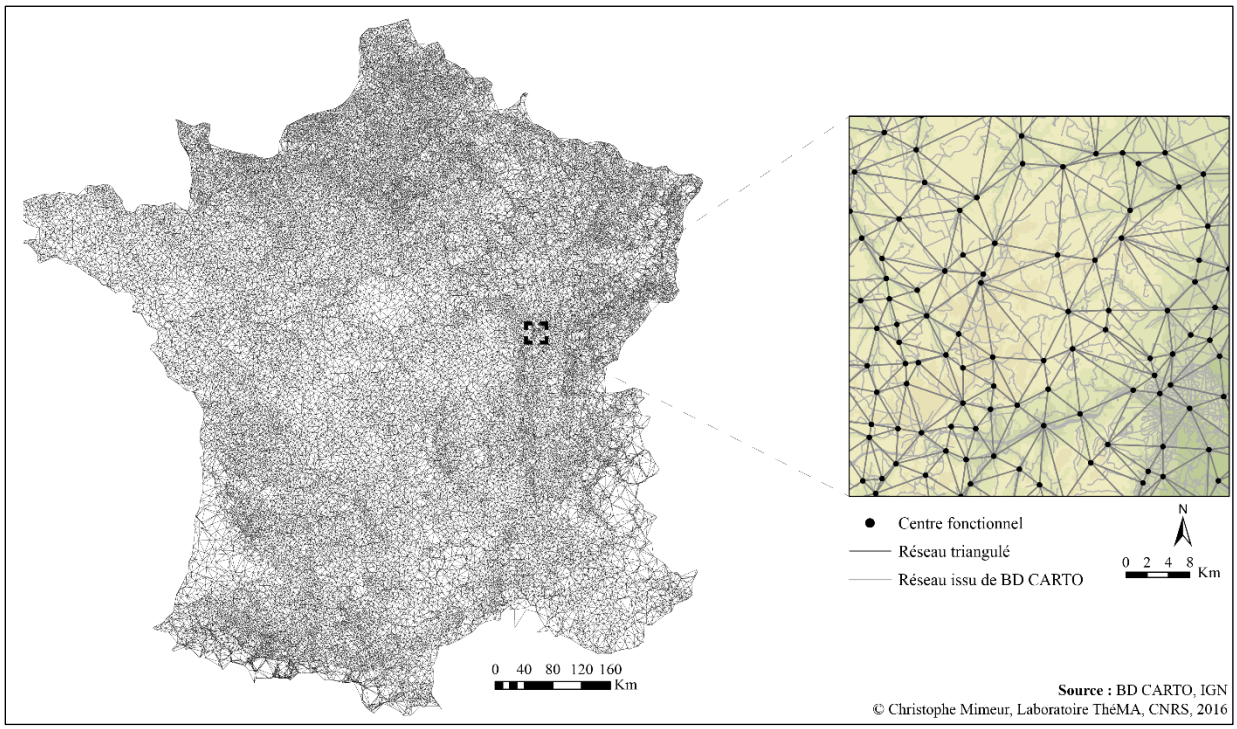


Figure 4. 15. La triangulation des 36 000 communes françaises : approche comparée avec BD CARTO

Comparaison des cheminements pédestres à partir de BD CARTO et de la triangulation			Superficie (milliers de km²)	Coefficient de variation	Coefficient de corrélation *(Pearson)*	Coefficient de corrélation *(Spearman)*
Temps de parcours unipolaire	*Nantes*	BD_CARTO	32	.48	.99	.99
		Triangulation		.49		
	Dijon	BD_CARTO	8,7	.50	.98	.98
		Triangulation		.49		
Temps de parcours multipolaire	*Pays de la Loire*	BD_CARTO	32	.18	.99	.99
		Triangulation		.18		
	Côte d'Or	BD_CARTO	8,7	.15	.96	.95
		Triangulation		.13		

Source : BD CARTO, IGN © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 4. 4. Comparaison des cheminements pédestres à partir de BD CARTO et de la triangulation

Désormais, compte tenu de la perte très limitée d'information dans le recours à un graphe théorique et de l'éloignement d'un procès d'anachronisme, nous décidons d'adopter le réseau issu de la triangulation de Delaunay pour modéliser les cheminements pédestres entre les 36 000 communes. Nous identifions ici un apport quant à la modélisation de réseaux simples à partir d'une information largement lacunaire, voire inexistante. Cette méthode pourrait être identifiée comme un moyen de pallier l'incertitude, l'incomplétude auxquelles le modélisateur géohistorien est souvent confronté quand il travaille à des échelles larges. De cette sorte, le nouveau graphe P est défini comme $P = (V, E, f)$, de la même manière que l'équation (5), où la fonction f définie une impédance piétonne qui prend en compte la dénivellation issue du MNT. Il permet de calculer des plus courts chemins exclusivement à pied. Il permet tout autant de calculer des plus courts chemins vers toutes les gares françaises. Alors, l'ultime étape consiste à modéliser les déplacements de sorte que le plus court chemin choisisse la gare la plus optimale pour se déplacer le plus rapidement possible sur le réseau ferroviaire.

4.3.3. **Les différentiels de réseaux dans un multigraphe**

Modéliser l'ensemble de la chaîne de déplacement revient à modéliser un déplacement bimodal, entre marche à pied et train *(Figure 4. 16. a.)*. En reprenant l'architecture d'un SIG, on imagine ainsi un réseau à plusieurs couches d'information *(Figure 4. 16. c.)*. La littérature traitant de la théorie des graphes s'est emparée des difficultés de prendre en compte des relations de différents niveaux (Kivelä et al., 2014). Les auteurs soulignent que pendant longtemps, les graphes se sont contentés de décrire les liens entre deux nœuds par un *« lien simple, statique, non-valué »* (Kivelä et al., 2014). La recherche sur les systèmes complexes s'est alors concentrée sur la complexification des liens, qui sont souvent très hétérogènes, à travers la direction des liens, leur longueur, leur permanence. D'autres apports ont traités des différents types de connexion entre les graphes. Dans la littérature de la sociologie, la prise en compte de différents types de liens a connu un vif succès qui a multiplié les dénominations (Kivelä et al., 2014), parmi lesquelles graphe multi-

couche, graphe multi-varié, graphe multi-niveau, graphe multi-relationnel. Dans la modélisation des réseaux de transport, le géographe est familier d'un réseau multi-modal, dont les théoriciens des graphes se sont eux aussi emparés (Lozano et Storchi, 2002). Dans une perspective réaliste, de nombreuses contributions se sont attachées à mesurer la viabilité des chemins dans un graphe modélisant un réseau multi-modal à des échelles urbaines et périurbaines (Fernández et al., 1994).

En considérant la large échelle du territoire national, nous proposons ici la construction d'un graphe unique, capable de traduire les temps de parcours en utilisant les deux modes précédemment définis dans les graphes G' et P, en s'assurant de la viabilité d'une telle combinaison. La première étape consiste à considérer le centre fonctionnel et la gare d'une même commune comme un seul et un même point du graphe multi-modal : à nos échelles, nous postulons que le temps d'accès d'une commune à une gare qu'elle possède dans son périmètre est nul. Pour autant, cette accommodation ne remet pas en cause la richesse de l'information dans le graphe G', de gare à gare, et dans le P, de commune à commune.

La seconde étape consiste à prendre en compte des liens dans le graphe combiné qui peuvent à la fois relever de la marche à pied ou d'un tronçon ferroviaire. Etant donné le caractère dynamique du réseau, nous devons conserver ces deux tronçons. Sans catégoriser les liens, les deux types de liens sont fortement discriminés par leur pondération, dans la mesure où la vitesse sur les tronçons ferroviaires est au moins trois fois supérieure pour les sections les plus lentes, mais au moins sept fois en moyenne pour l'ensemble du réseau ferré.

On peut donc construire un graphe, qui combine les deux arbitrages – rejoindre la gare la plus optimale, arbitrer avec un cheminement entièrement à pied –, comme la combinaison des deux précédents *(Figure 4. 16. c.)* :

$$F = G' + P \tag{10}$$

$$F = (V, E, f, \omega_n) \tag{11}$$

Le graphe F est défini par les nœuds V, qui représentent les centres fonctionnels des 36 000 communes comme les gares ferroviaires, par les liens N issus de la triangulation et du réseau ferroviaire, où la fonction *f* associe les sommets, et où ω_n est l'impédance pour le traverser. On comprend alors qu'il peut exister un ou deux liens entre deux mêmes sommets, selon qu'ils décrivent un cheminement pédestre ou un tronçon ferroviaire. Ainsi, le graphe F est de la famille des multi-graphes, définis comme un graphe qui permet l'existence de liens multiples, qui possèdent les mêmes nœuds de début et de fin *(Figure 4. 16. b.)*. Toutes ces opérations sont possibles grâce au package *igraph* de R (Csardi et Nepusz, 2006), qui s'abstient de la géométrie des nœuds et des liens, permettant de grandes capacités de calcul à partir de graphes de grande taille *(Figure 4. 16. d.)*.

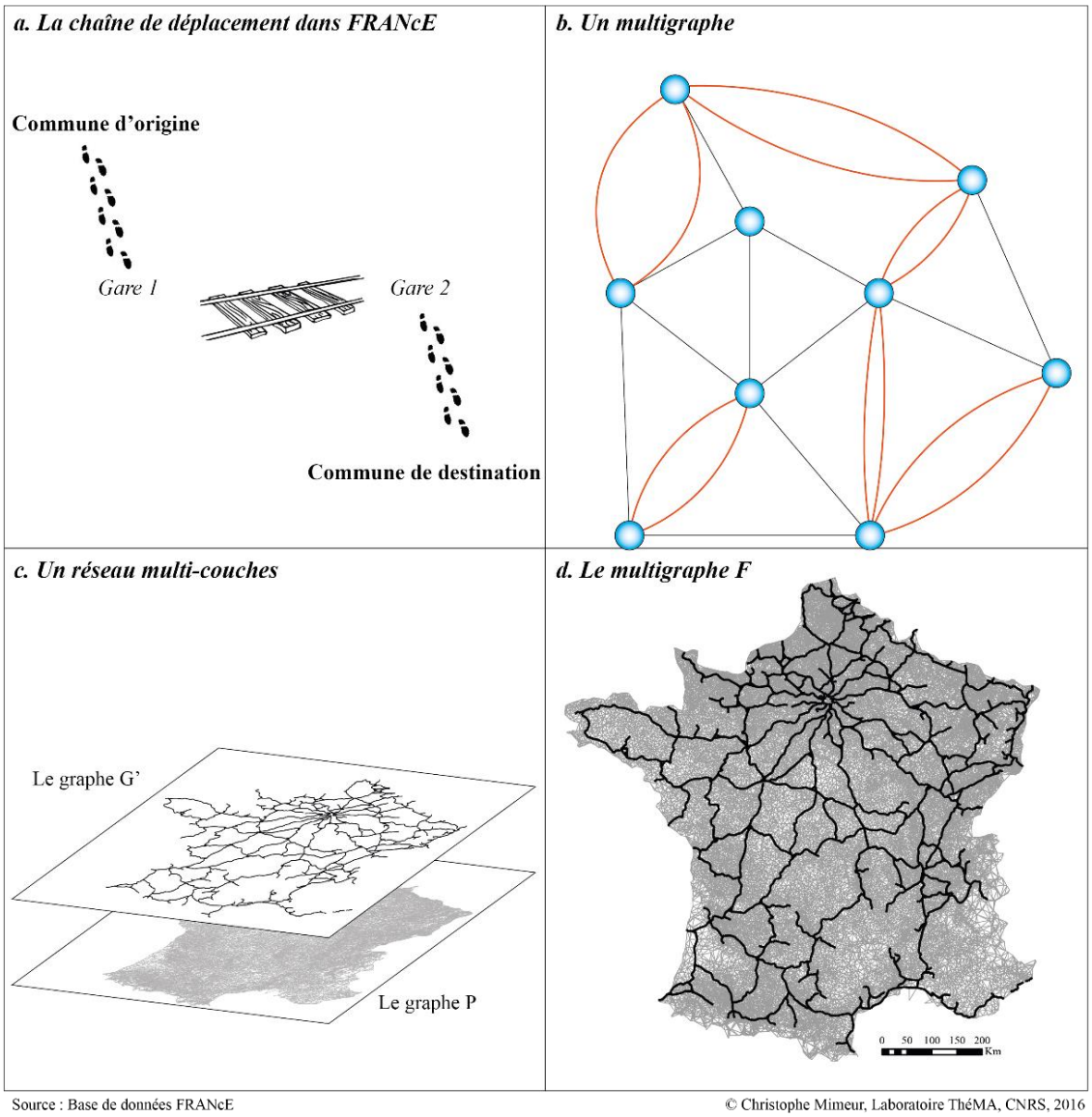


Figure 4. 16. Le multigraphe F pour les déplacements dans FRANcE

Nous proposons enfin un exemple de l'évolution des temps de parcours depuis Perros-Guirec, dont l'évolution de la desserte est très probante, jusqu'à Paris entre 1860 et 1930 *(Figure 4. 17)* depuis le multigraphe F. La détermination des plus courts chemins par l'algorithme de Dijkstra est ici bien capable de reproduire la vraisemblance des déplacements, parce qu'il traduit la forte discrimination des vitesses entre le cheminement pédestre et le voyage en train mais aussi parce qu'il n'y a pas de ruptures de charge intempestives. La signature des temps de parcours depuis Perros-Guirec traduit ici une triple évolution dans le temps d'accès vers la capitale :

- Une forte diminution du temps de parcours entre les deux villes : près de 45 heures pour se rendre à Paris en 1860 alors qu'il ne faut plus que 9 heures pour s'y rendre en 1930 ;

- Un gain de temps largement dû à l'arrivée du réseau à proximité de Perros-Guirec : le réseau ferroviaire pénètre la Bretagne à partir de 1870 et permet ainsi de voyager en train de Plouaret à Paris dès cette décennie. Les ouvertures respectives des gares de Lannion puis Louannec réduisent le temps d'accès pédestre au réseau ;
- Un gain de temps influencé par l'évolution des vitesses sur le réseau ferroviaire, fruit des innovations techniques et des investissements consentis (Caron, 1997) à toutes les échelles du réseau.

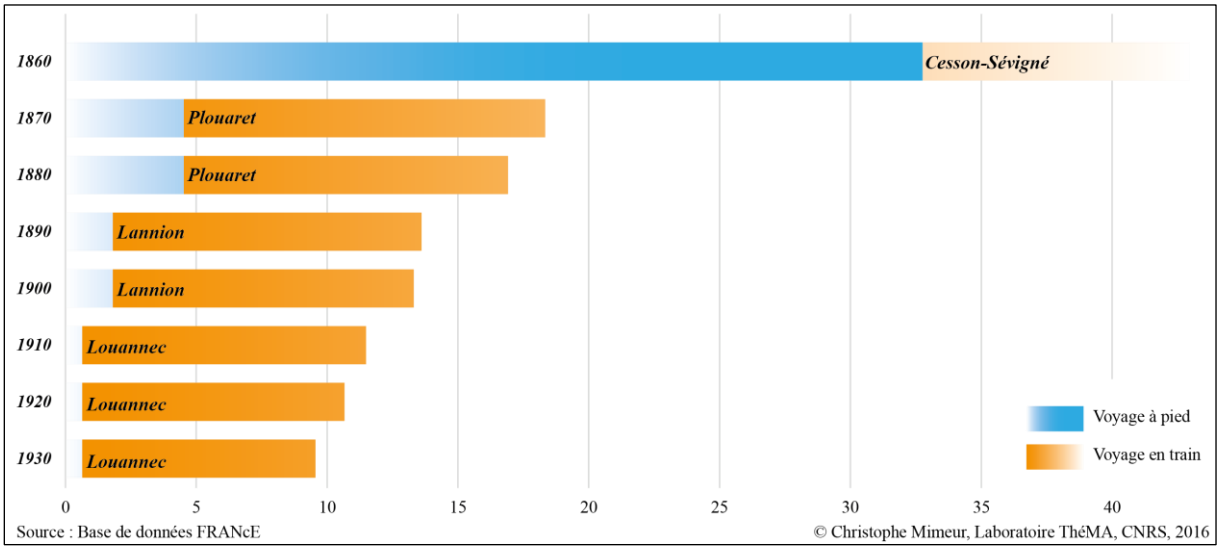


Figure 4. 17. Temps de parcours depuis Perros-Guirec jusqu'à Paris de 1860 à 1930

On constate à ce point de notre raisonnement que le multigraphe F est en capacité de calculer pour chaque paire de ville un temps de parcours par le réseau ferroviaire et/ou par la marche à pied, suivant les critères de pertinence et de précision qui sont à la base de notre positionnement pour appréhender les relations entre transport et territoire.

Conclusion

L'objectif de ce quatrième chapitre était de donner un cadre pour le développement d'une chaîne géo-méthodologique pour la construction de connaissances géohistoriques du lien entre réseau et territoire sur la longue durée. Le dimensionnement de nos échelles et de la base de données FRANcE participe au rapprochement avec les sciences informatiques et mathématiques de manière à s'émanciper, de manière au moins temporaire, du SIG : l'inscription sur la scène de la *GeoComputation* permet de replacer le graphe au centre de notre instrument de recherche *(Figure 4. 18)*. Souvent appréhendée par les outils *(Figure 4. 18. a.)*, nous préférons une approche par l'objet et la question de recherche, qui permet tout aussi bien de lier géographie, analyse spatiale et science informatique *(Figure 4. 18. b.)*.

Dans ce triangle, le formalisme des graphes répond ainsi aux enjeux de capitalisation de l'information contenue dans FRANcE de trois manières différentes. La première est la capacité de l'objet « graphe » à répondre aux critères de construction des indicateurs d'accessibilité à travers la recherche des plus courts chemins. La deuxième est la capacité de dialogue entre graphe et SIG pour la transformation de l'information brute en un graphe fonctionnel, de manière à produire de nouvelles informations, susceptibles de devenir des connaissances géohistoriques, sur la structure et la morphologie du réseau ferroviaire français sur le temps long. La troisième est la capacité d'un tel objet à participer aux enjeux géohistoriques oscillant entre la quête d'une source historique précise et les besoins d'une généralisation de l'information à de larges échelles. De ce point de vue, nous proposons une méthode permettant de généraliser un graphe de cheminements entre l'ensemble des communes françaises, prenant en compte les contraintes topographiques.

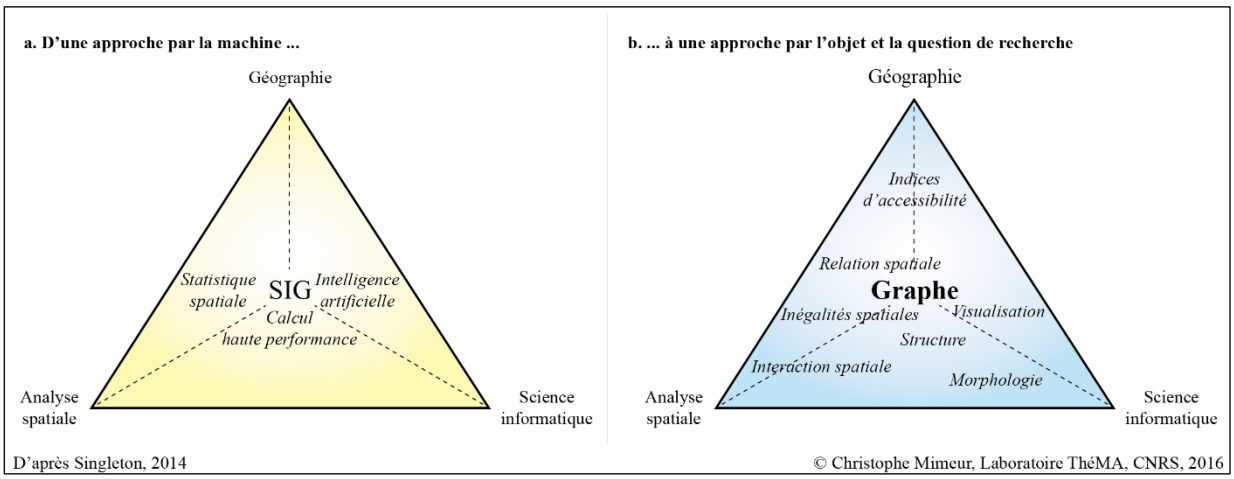


Figure 4. 18. Les contours de l'instrument de recherche pour la base de données FRANcE

Ainsi, nous obtenons à ce stade un multigraphe *F* capable d'obtenir les temps de parcours par les modes piétons et ferroviaires sur l'ensemble du territoire. Face aux enjeux de la construction d'indicateurs d'accessibilité, ils participent à la création de matrices d'interaction, permettant désormais d'envisager la modélisation des relations entre accessibilité et population.

Chapitre 5. MESURER DES RELATIONS DE LONG TERME ENTRE INFRASTRUCTURE ET TERRITOIRE

Introduction

Les études géohistoriques se sont souvent trouvées confrontées à l'incomplétude de la donnée, devant alors se borner à des analyses à une micro-échelle et faisant de la montée en généralité un objectif presque inatteignable. Cette règle a souvent prévalu dans les monographies régionales de la géographie française. Pour autant, *« les progrès dans les SIG, la disponibilité croissante de données numériques à haute résolution et des techniques économétriques sophistiquées ont tous contribué à l'intérêt académique et l'étude de long terme des impacts des infrastructures de transport sur les territoires »* (Kasraian et al., 2016). Dans leur revue de la littérature, ces auteurs montrent que les principales conclusions portent sur une règle qui semble pencher en faveur d'un effet structurant d'une infrastructure : l'accès au réseau, par une simple distance euclidienne, est déterminant dans la conversion fonctionnelle des territoires, avec des résultats plus significatifs sur le réseau routier que sur le réseau ferroviaire (Kasraian et al., 2016). Dans ce travail, le matériau empirique que nous avons collecté dans la base de données FRANcE, la population d'une part, le multigraphe *F* d'autre part, de même que le questionnement que nous avons soulevé dans notre première partie, nous permettent désormais d'envisager une modélisation des relations venant contribuer aux nombreuses études empiriques sur le lien entre croissance et développement des réseaux. Nous proposons de sophistiquer l'approche simple de l'accès au réseau par l'accessibilité, en lien avec les critères édictés en introduction de seconde partie. Nous avons alors pour le cas français des moyens de mesures précis et diversifiés de la position des lieux par rapport au réseau, venant limiter une importance trop grande donnée aux estimations et souvent décriée par les détracteurs du mythe de l'effet structurant.

Ce chapitre montre les rapprochements entre les questionnements de l'économie historique et la posture épistémologique que nous avons définie, dans un objectif de capitalisation d'une information géohistorique par ailleurs déjà collectée. Dans ce cadre, nous nous inspirons alors de concepts de la Nouvelle Economie Géographique pour enrichir les indicateurs d'accessibilité, posant la question des échelles, et de méthodes économétriques permettant de prendre en compte la dimension temporelle. Dès lors, nous posons l'hypothèse que le recours à des méthodes qui dépassent celle des Moindres Carrés Ordinaires peut permettre d'améliorer les estimations de l'effet réseau sur la croissance de la population.

Ce chapitre vise à exposer notre démarche pour analyser les relations entre croissance de la population et accessibilité entre 1860 et 1930 dans une perspective géohistorique de long terme. Tout d'abord, il s'agit d'éprouver les capacités du nouveau graphe à enrichir les indicateurs décrivant la dotation des communes françaises en réseau (5.1.). Pour autant, la confrontation du réseau et de la population doit prendre en compte un biais important d'endogénéité des relations : en mobilisant les apports de l'économie historique, nous proposons une modélisation par l'économétrie de panel (5.2.). Cette spécification justifie par ailleurs le recours à la diversification géohistorique des données pour affiner l'analyse des effets de réseaux (5.3.), jusqu'à l'utilisation de variables instrumentales.

5.1. Caractériser la dotation en réseau par l'accessibilité

La mesure de l'accessibilité sur le long terme a fait l'objet de plusieurs investigations (Axhausen, 2008 ; Bruinsma et Rietveld, 1998 ; Gutiérrez, Condeço-Melhorado et Martín, 2010 ; Vandenbulcke, Steenberghen et Thomas, 2009). Pourtant, des auteurs mentionnent que l'usage du terme d'accessibilité n'est pas souvent suivi d'une définition et de mesures précises. Dans leur revue de littérature, Dena Kasraian et ses co-auteurs relèvent de nombreuses approches qui utilisent de simples distances métriques, souvent en ligne droite, alors que des mesures le long des réseaux restent encore rares dans les études empiriques de long terme (Kasraian et al., 2016) : le choix des infrastructures à étudier l'a souvent emporté sur le choix des indicateurs à adopter (Bruinsma et Rietveld, 1998). Après avoir défini des indicateurs d'accessibilité fonctionnelle, nous définissons des indicateurs d'accessibilité multipolaire, dans une perspective abductive et dans le cadre des critères de pertinence et de précision évoqués dans la première partie de la thèse, en mobilisant immédiatement l'instrument que nous venons de construire.

5.1.1. **Des indicateurs fonctionnels …**

A partir des quatre propriétés des indicateurs de l'accessibilité évoquées en introduction de cette seconde partie (Morris, Dumble et Wigan, 1979), nous développons ici notre réflexion dans la construction des indicateurs dans la base de données FRANcE, en mettant l'accent que les *« éléments de séparation spatiale qui traduit les changements en termes de performance du système de transport »* ainsi que sur l'interprétation et la traduction *« en termes de politiques urbaine ou de transport »*. De notre point de vue, l'approche par les temps de parcours participe une complexification de l'approche. Nous optons à ce stade pour une complexification progressive des indicateurs, afin d'évaluer plus tard l'apport de chaque indicateur dans la modélisation des relations avec la croissance démographique : notre approche combine alors une information *vers* et *le long du* réseau de transport, à partir du multigraphe *F*. Une comparaison des variabilités des indicateurs créés nous montre comment la complexification participe à une plus forte variabilité, et participe aussi à la visualisation de la diffusion du réseau et des impacts spatiaux : pour cela nous optons pour des visualisations à l'échelle nationale mais aussi régionale.

Pour commencer, on reprend l'indicateur le plus souvent utilisé dans la littérature empirique : l'accessibilité n'est souvent qu'une variable binaire, selon que l'unité d'observation possède une gare ou non (Atack et al., 2010 ; Hornung, 2012b). Elle est mesurée dans la *Table 5. 1. a.*. Une première mesure permet alors de simplement enrichir l'information vers le réseau, où le temps de parcours à pied entre 1860 et 1930 remplace la simple variable binaire *(Table 5. 1. b.)*. Seul le graphe *P* est alors pris en compte, elle est définie par :

$$Acc_ga_{igt} = \min(D_{igt}) \quad (12)$$

où *i* désigne la commune d'origine et *g* l'ensemble des gares pour une période *t*.

	a. Accès direct à la gare					*b. Temps de parcours vers la gare (en minutes)*				
	Min	*Max*	*Moyenne*	*Ecart-type*	*Coefficient de variation*	*Min*	*Max*	*Moyenne*	*Ecart-type*	*Coefficient de variation*
1860	0	1	0,0425	0,2018	4,74423	0	4837,49	364,95	422,98	1,15901
1870	0	1	0,0776	0,2675	3,4476	0	3495,12	205,35	252,57	1,22994
1880	0	1	0,1183	0,32301	2,72951	0	2893,92	144,14	180,25	1,25051
1890	0	1	0,1689	0,3747	2,2177	0	2893,92	106,57	147,69	1,38578
1900	0	1	0,2073	0,4053	1,9555	0	2502,54	88,96	118,99	1,33753
1910	0	1	0,2456	0,4304	1,7526	0	2502,54	76,56	108,67	1,41953
1920	0	1	0,27105	0,4445	1,6399	0	2502,54	70,69	99,29	1,40455
1930	0	1	0,27335	0,44568	1,6304	0	2188,2	69,72	94,87	1,63082

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 5. 1. La complexification de l'indicateur d'accessibilité au réseau

Appliquée à la base de données FRANcE, la construction d'une variable binaire montre que sur le temps long, la dispersion de la distribution est croissante, alors que la diffusion du réseau montre une dispersion décroissante des temps de parcours à pied vers la gare la plus proche. Les grandes valeurs des coefficients de variation montrent une grande variabilité des valeurs autour de la moyenne, parce qu'il y existe des valeurs extrêmes. De plus, cette première complexification montre la prise en compte d'effets indirects de la croissance du réseau : même si une gare n'est pas ouverte dans une commune – décrit par un 0 dans la variable binaire – alors l'ouverture d'une gare à proximité de cette même commune peut avoir un impact dans la mesure où elle permet une amélioration de l'accessibilité au réseau ferroviaire. Ainsi, une représentation cartographique des temps de parcours vers la gare la plus proche entre 1860 et 1870 montre comment la diffusion du réseau a largement homogénéisé les conditions d'accès au réseau *(Figure 5. 1)*. Si en 1860, de nombreuses contrées sont encore éloignées de plus de 24 heures de la gare la plus proche, seules quelques zones dans un relief de moyenne ou haute montagne sont encore concernées en 1930. Alors qu'il fallait en moyenne six heures, avec de grands écarts, pour rejoindre une gare en 1860, il ne faut qu'un peu plus d'une heure pour marcher jusqu'au train en 1930, avec des écarts qui se sont considérablement réduits durant notre période d'étude.

Seuls les reliefs des Pyrénées et du Massif Central montrent encore des valeurs extrêmes, qui combinent à la fois une densité moindre du réseau, et des difficultés physiques pour cheminer vers la gare, étant données les pentes importantes. Par ailleurs, la croissance de la densité du réseau a pour conséquence un effacement de l'effet linéaire pour rejoindre une gare, tant le semis des gares tend à se régulariser, caractérisant l'effet d'une pieuvre. Ce constat vient confirmer l'approche du multigraphe *F*, qui privilégie non pas la gare la plus proche, mais la gare la plus efficiente pour réduire le temps de parcours global, dans la mesure où faire quelques minutes de marche en plus peut avoir un effet fort sur le temps de parcours en train par la suite. Pour autant, nous ne pouvons pas limiter ces investigations à cette première complexification parce qu'aucune mesure le long du réseau ferroviaire n'a été encore prise en compte : pour l'instant, cet indicateur ne rend pas compte des propriétés qualitatives et quantitatives du réseau lui-même, indépendamment de sa structure et de sa géométrie.

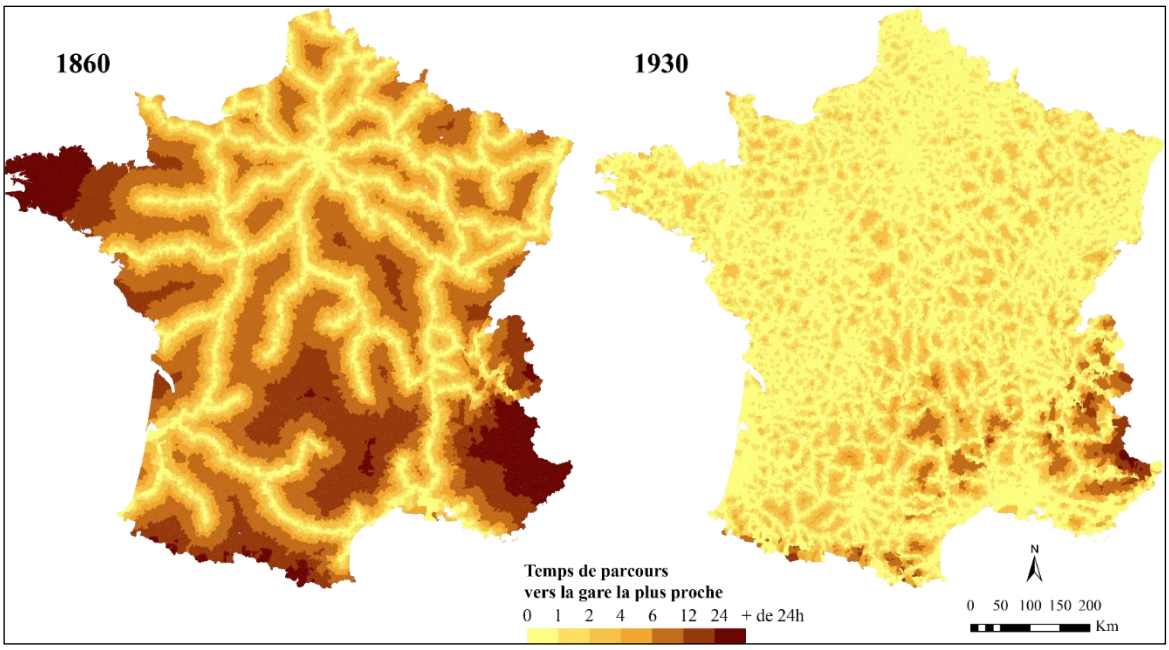


Figure 5. 1. Temps de parcours vers la gare la plus proche en 1860 et en 1930

Une première complexification géohistorique consiste à modéliser des accessibilités unipolaires, définies comme la capacité d'une commune à rejoindre un centre fonctionnel, à partir de la structure administrative française. De manière hybride, à partir du multigraphe *F*, cette fois, cet indicateur est noté :

$$Acc_fonc_{it} = \min(D_{ift}) \quad (13)$$

où *i* désigne la commune d'origine et *f* le centre fonctionnel considéré – capitale, préfecture régionale ou préfecture départementale – pour une période *t*. Il permet par exemple de modéliser le temps de parcours nécessaire à chaque habitant pour rejoindre la capitale Paris, alors même que l'histoire de la constitution du réseau a montré comment il est imaginé dès le départ comme un réseau en étoile partant de la capitale (Caron, 1997). On répète cette opération pour les capitales des anciennes régions administratives métropolitaines et les préfectures départementales. Le choix de ces variables rentre dans l'établissement des critères de précision et de pertinence, quand *« dès avant 1870, on parle, dans l'histoire ferroviaire française, des chemins de fer régionaux »* (Gonjo, 1972), et que les voies ferrées d'intérêt local devaient rejoindre chaque chef-lieu de canton, suivant les grandes orientations du Plan Freycinet de 1878. A ce stade, nous posons l'hypothèse que ces trois indicateurs, confrontés à la croissance de la population, sont une opportunité afin d'apporter de nouveaux indices quant au niveau d'échelle à partir duquel nous pouvons approfondir notre raisonnement géohistorique : si l'accès à la capitale, dans un Etat centralisé est une variable qui joue sur la croissance des populations, est-ce que la capacité d'une commune à rejoindre un centre de décision à une méso- ou à une micro-échelle a plus d'impact et de pouvoir explicatif dans la trajectoire démographique des communes françaises ? Ici, un simple aperçu de ces trois indicateurs pour l'année 1900 montre comment les disparités spatiales diffèrent ou se rapprochent d'un centre fonctionnel à l'autre *(Figure 5. 2)*.

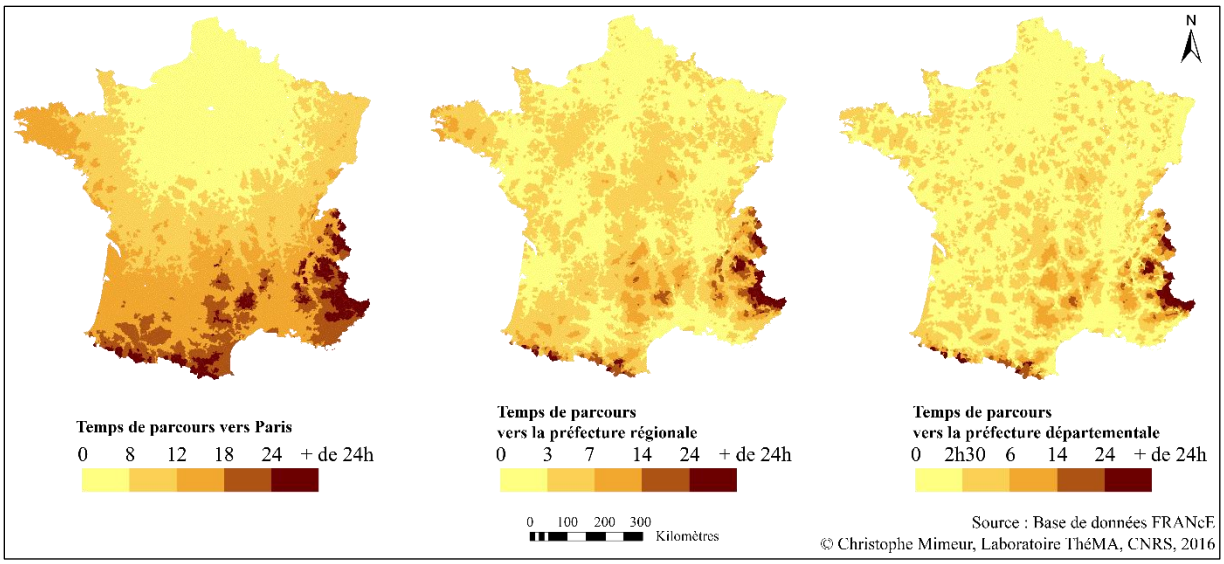


Figure 5. 2. Accessibilité fonctionnelle en 1900

Ainsi, l'accessibilité à Paris montre, de manière naturelle, une disparité entre Nord et Sud, suivant un processus concentrique. L'éloignement par la distance kilométrique est un facteur important, même si des disparités plus fortes se font sentir au sud d'une ligne entre Bordeaux et Lyon. Dans ce cas, le prolongement des lignes principales vers Bordeaux et Toulouse d'une part, vers Marseille et Grenoble d'autre part montre une meilleure accessibilité à la capitale tandis que le Massif Central a un rôle de barrière vers le Languedoc-Roussillon, qui n'est pas atténué par la vitesse sur l'arc ouest méditerranéen.

A l'échelle régionale, les disparités ne sont plus les mêmes, elles mettent en exergue des territoires en marge des centralités régionales en 1900. Si la densité du réseau ferroviaire semble influencer les résultats, alors l'effet des vitesses se fait ici davantage sentir : pour des densités qui sont proches, des différentiels d'accessibilité ressortent toutefois, et nous posons ici l'hypothèse que les vitesses de déplacement influent directement sur ces résultats. Ainsi, autour de Rennes, la rupture est prononcée et rapide à mesure que l'on se déplace vers l'ouest alors qu'une ligne rapide entre Rennes et Le Mans participe à un phénomène plus diffus et à une meilleure accessibilité à la capitale bretonne depuis l'est de la région. On peut faire la même remarque quant à l'axe entre Bordeaux et l'Espagne qui renforce l'accessibilité régionale du Pays Basque. En revanche, le Morvan semble largement pénalisé, par un relief modéré d'une part, et par l'absence de lignes transversales d'autre part, alors même que cette structure ne ressortait pas dans l'accessibilité à la capitale.

Enfin, à l'échelle départementale, l'accessibilité à la préfecture décrit des différentiels très locaux en 1900, dans lesquels l'impact du relief semble jouer un rôle majeur dans le Massif Central, les Alpes et les Pyrénées. De même, en dépit d'une homogénéisation progressive du réseau ferroviaire, encore en 1900, de fortes disparités demeurent dans le centre de la France. Un examen approfondi sur la hiérarchie du réseau doit nous permettre encore davantage de mettre en exergue ces inégalités dans la dotation en réseau.

Ainsi, la mise en œuvre d'accessibilité unipolaire montre ici ses avantages dans la

familiarité que le lecteur a avec le territoire français. La disparité entre Nord et Sud ressort très fortement pour l'accessibilité à Paris, où la dimension stellaire du réseau se poursuit jusqu'en bas de la Vallée du Rhône et vers Bordeaux, ne montrant pas un rôle déterminant des liaisons transversales. Elles ressortent en revanche dans les accessibilités vers les préfectures régionales et départementales. Si la structure du réseau semble déterminante dans la construction de tels indicateurs, alors nous verrons comment la prise en compte des vitesses revêt un enjeu important dans une perspective cette fois-ci diachronique. Pour autant, cette démarche procède du choix arbitraire de villes centres, prises pour référence dans la construction unipolaire. Dans la démarche abductive que nous avons définie, nous postulons aussi que la construction d'indicateurs d'accessibilité, indépendamment de la sélection de centres définie *a priori*, peut participer à la complexification de variables décrivant la dotation en réseau des 36 000 communes françaises.

5.1.2. **… aux indicateurs multipolaires**

La construction d'indicateurs d'accessibilité multipolaire permet de poser une question plus globale qu'avec les indicateurs d'accessibilité fonctionnelle : globalement, comment une commune profite-t-elle du réseau de transport et de ses performances intrinsèques ? Ainsi, ce type d'indicateurs semble être le plus à même de décrire une interdépendance vertueuse entre la croissance des revenus – examinée à travers la croissance démographique dans notre cas –, la croissance d'économies d'échelles et l'agrandissement des marchés par la baisse généralisée des coûts de transport (Axhausen, 2008). Ce cercle vertueux est celui décrit par la Nouvelle Economie Geographique (Krugman, 1991). Alors, la construction de ces indicateurs propose un raffinement de l'approche pour diversifier les mesures, dans une perspective de long terme.

Indépendamment de la position géographique des villes et sur le réseau, ces indicateurs ont pour vocation à mesurer la performance des cheminements sur le réseau. Dans l'esprit d'une situation contrefactuelle[24] souvent décrite dans les expériences en économie historique (Diebolt, 2001), alors nous misons aussi sur la confrontation des indices faisant appel soit à une vitesse homogène, soit à la vitesse hétérogène effectivement constatée dans la base de données FRANcE. Ces indicateurs entrent alors dans la famille de ceux des potentiels d'accessibilité (Geurs et Ritsema von Eck, 2001), dont les efforts en termes d'intelligibilité sont à démontrer ici.

L'indicateur multipolaire caractérise le temps de parcours nécessaire pour voyager d'une ville à toutes les autres. A partir de la théorie des graphes, la mesure de l'excentricité traduit la distance maximum nécessaire pour rejoindre tous les sommets. Pour autant, nous adoptons ici des mesures basées sur des valeurs moyennes plutôt que sur des valeurs maximum, en complément de nos propos sur l'usage de vitesses moyennes plutôt que les records de vitesse. Construites à partir du multigraphe F, elles ne doivent pas être confondues avec les mesures de centralité évoquées dans le *Chapitre 4* à partir de la théorie des graphes : une ville peut être alors très accessible parce qu'elle est à proximité d'un nœud central du graphe, sans que cette même ville soit centrale. Par ailleurs,

[24] Nous reviendrons sur l'importance de la prise en compte de telles situations dans la section suivante.

la prise en compte de valeurs moyennes nous permet de pondérer les indicateurs obtenus à partir d'autres informations géohistoriques.

Nous définissons ici un indicateur multipolaire, défini comme le temps de parcours moyen pour chaque commune *i* du multigraphe F vers toutes les autres communes *j*, pour chaque décennie *t* :

$$MULTI_{it} = \frac{1}{N_j - 1} \sum_{N} \Delta_F(i,j) \tag{14}$$

Dans cette équation et celles qui suivent, l'indicateur $\triangle$ caractérise le plus court chemin entre *i* et *j*, selon que l'on prenne en compte le graphe *F* ou le graphe *P*. *N* caractérise le nombre de liens dans le graphe. Si la formule mathématique propose un exposant sur la distance entre *u* et *v*, alors la majorité des études d'accessibilité lui donne une valeur de 1 (Gutiérrez, Condeço-Melhorado et Martín, 2010), c'est donc la valeur que nous retenons aussi ici. On construit en parallèle un indicateur d'accessibilité multipolaire, à partir des données de population contenues dans la base de données FRANcE :

$$MULTIPOND_{it} = \frac{1}{\sum_N POP_{jt}} \sum_{N} POP_{jt}.\Delta_F(i,j) \tag{15}$$

La *Figure 5. 3* montre l'évolution de l'accessibilité multipolaire entre 1860 et 1920. En 1860, le temps de parcours moyen est largement impacté par la structure et la géométrie du réseau ferroviaire, démontrant ainsi le rôle centralisateur de l'Etoile de Legrand. Au contraire des mesures de centralité et des mesures d'accessibilité fonctionnelle, les effets de bord ne ressortent pas ici, à moins que la structure du réseau n'y joue un rôle, comme c'est le cas en Bretagne en 1860, mais aussi dans les reliefs alpins et pyrénéens. La poche dessinée par le Massif Central ressort tout autant. Vingt ans plus tard, alors que la valeur moyenne chute, une dichotomie entre Nord et Sud semble se dégager, à la manière dont elle ressortait déjà pour l'accessibilité fonctionnelle Pourtant, la structure stellaire ne semble jouer un rôle que dans la moitié nord du pays. En 1900, alors que la « fièvre ferroviaire » s'est opérée dans le pays, sous l'impulsion de l'Etat d'une part, des initiatives privées d'autre part, le maillage plus complexe du réseau joue davantage que la dimension stellaire autour de Paris. Confrontées pourtant aux différentiels dans la densité des réseaux *(cf. Figure 3. 13, p. 125)*, les régions du centre de la France sont bien situées dans la structure du réseau. On note aussi la persistance de la barrière du Massif Central. Ainsi, en 1920, l'apogée du réseau ferroviaire montre une homogénéisation dans les temps de parcours moyen, alors qu'il faut encore compter près de 24 heures pour rejoindre les zones les plus enclavées du Massif Central.

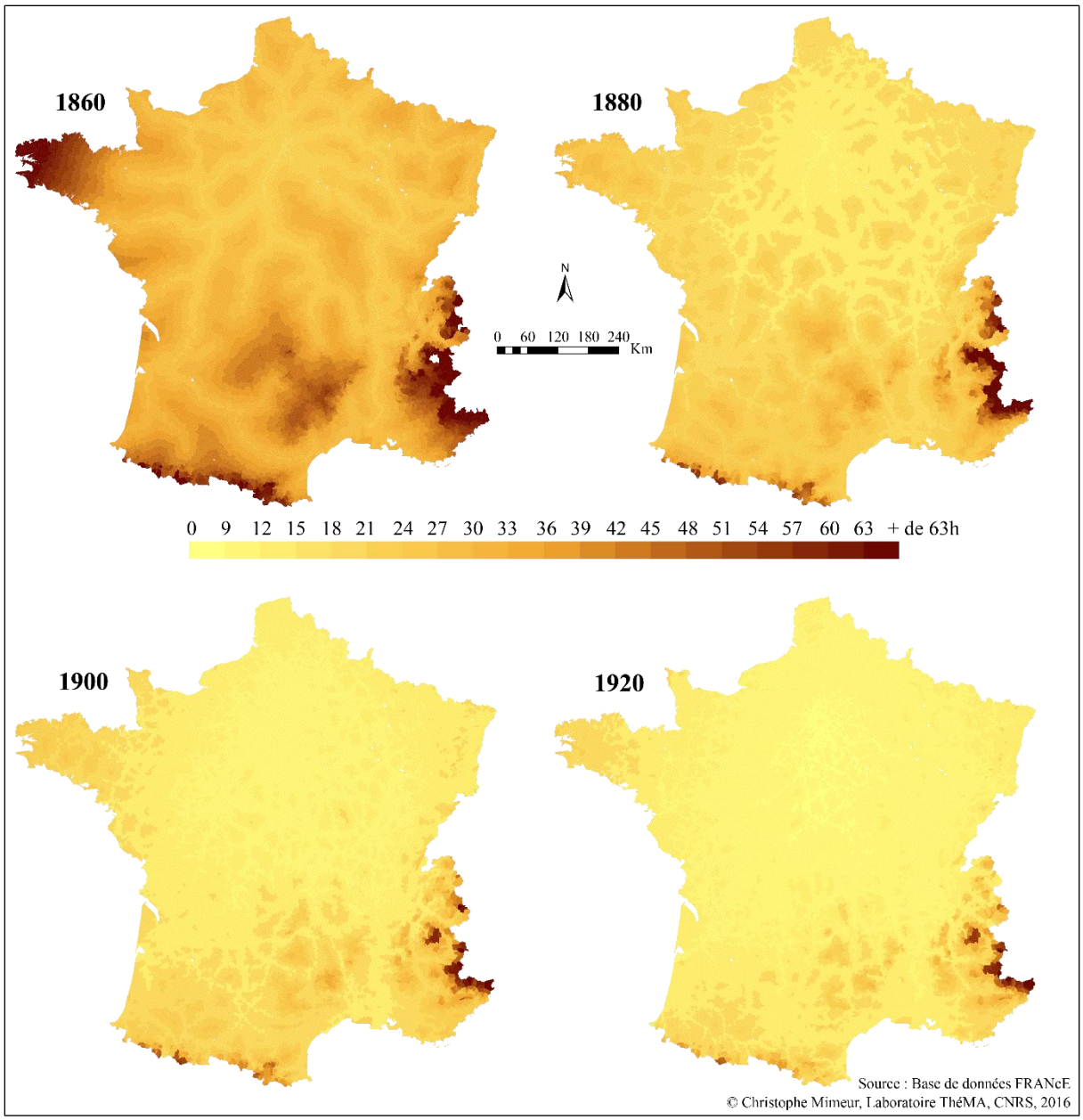


Figure 5. 3. Accessibilité multipolaire en 1860, 1880, 1900 et 1920

Quand on confronte la distribution des deux indicateurs moyens et pondérés *(Figure 5. 4)*, la première remarque réside dans la diminution croissante des temps de parcours moyens entre 1860 et 1930. Si en 1860, le temps de parcours moyen est de 28 heures, alors il chute en-dessous des 24 heures dès la décennie suivante, pour atteindre un peu plus de 11 heures en 1930. La tendance est bien entendu conservée quand on se penche sur l'indicateur pondéré par la population. Pourtant, le temps de parcours moyen augmente quand il est pondéré en 1860, en raison de la faible diffusion du réseau sur le territoire français. Par la suite, les deux moyennes sont quasiment identiques, venant démontrer la propriété universaliste du réseau ferroviaire.

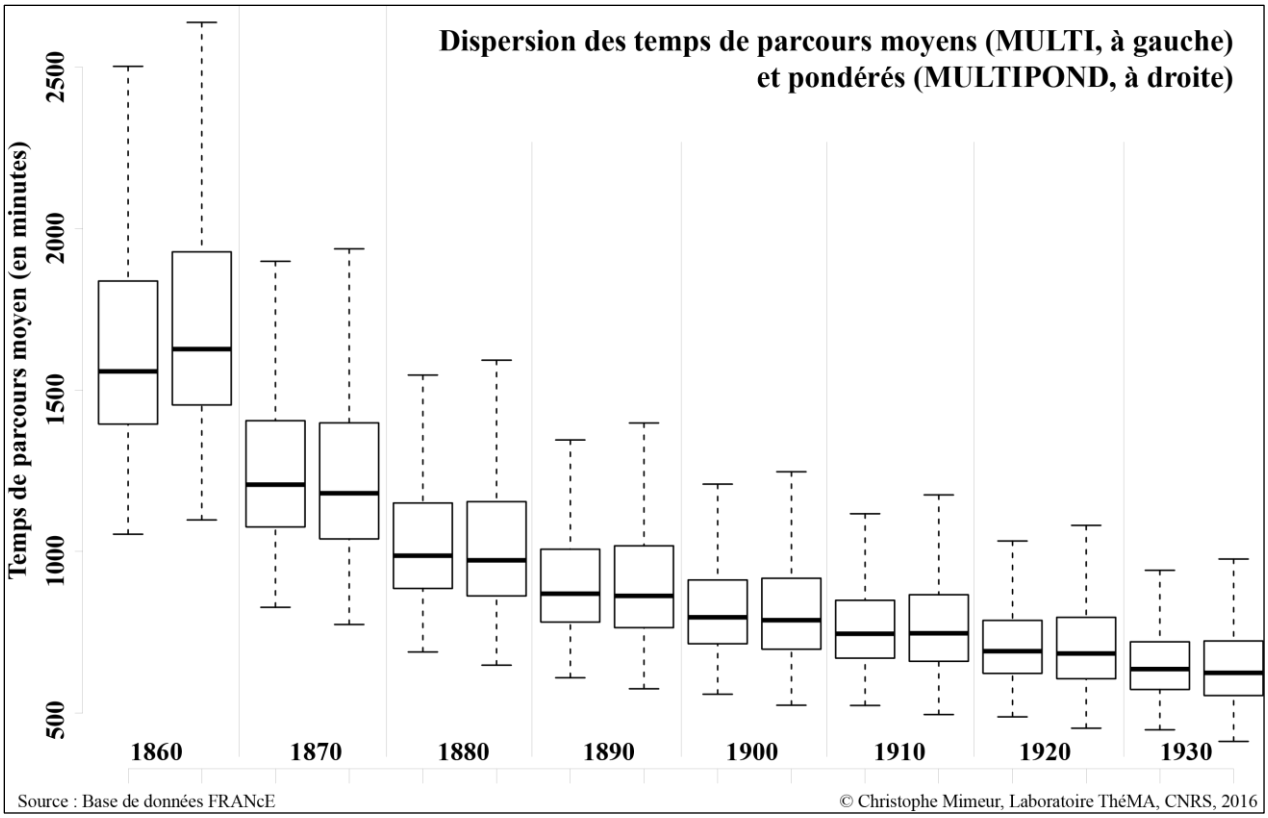


Figure 5. 4. Dispersion des temps de parcours moyens et pondérés

Pour autant, cette propriété doit être remise en cause quand on analyse cette fois la dispersion des deux indicateurs, perceptible à travers des écarts interquartiles plus forts pour le second indicateur. Alors sur le temps long, l'indicateur pondéré par la population offre une plus grande variabilité : en dépit d'une extension du réseau qui participe au rattrapage en termes de performances d'accessibilité moyenne, l'indicateur pondéré montre des écarts plus grands entre les communes. Si la dispersion absolue décrit les performances et la diffusion du réseau, le maintien d'une dispersion relative, plus marquée quand on prend l'indicateur pondéré, montre de son côté une forte hétérogénéité, que nous estimons pouvoir imputer aux différentiels de vitesse sur le réseau.

Afin d'étudier plus en profondeur l'impact des vitesses dans le processus de diminution des temps de parcours moyen d'une part, et dans le processus de variabilité d'autre part, nous proposons ici la construction d'indicateurs qui combinent alors deux effets : la comparaison d'une situation avec le train et d'une situation sans le train, ainsi que l'effet des différentiels de vitesse le long du réseau ferroviaire. Pour chaque paire de communes *u* et *v*, il s'agit de calculer le temps de parcours sur le graphe P et sur le multigraphe F pour obtenir le ratio entre ces deux situations. De la même manière que précédemment, on peut obtenir un indicateur pondéré par la population :

$$RATIO_{it} = \frac{1}{N_j - 1} \sum_{N} \frac{\Delta_F(i,j)}{\Delta_P(i,j)} \quad (16)$$

$$RATIOPOND_{it} = \frac{1}{\sum_N POP_{jt}} \sum_{N} POP_{jt} . \frac{\Delta_F(i,j)}{\Delta_P(i,j)} \quad (17)$$

A ce stade, nous estimons que ce type de confrontation peut être une démarche transposable dans le cadre de l'étude d'autres réseaux historiques, voire dans l'étude des concurrences entre modes de transport dans nos réseaux de transport contemporains, avec l'arrivée de la grande vitesse ferroviaire notamment. Dans la *Figure 5. 5*, on mesure l'effet de la concurrence entre le cheminement à partir d'une vitesse homogène, et le cheminement possible avec le chemin de fer à différentes dates de notre période d'étude. L'intensité de la concurrence entre les deux modes traduit les obstacles topographiques d'une part et l'effet de la vitesse ferroviaire d'autre part. Cet indicateur permet alors d'explorer dans quelles mesures le chemin de fer participe au désenclavement de certaines contrées du territoire français, à mesure que le réseau s'étend sur le territoire français. La représentation de la distribution des valeurs se fait ici par quartile. Dans la mesure où la marche à pied est largement contrainte par le relief, une approche intuitive mène à penser que les régions les plus accidentées bénéficient d'abord de la vitesse sur le chemin de fer, mais la construction des voies est elle-même contrainte par le relief. Alors, en 1860, on retrouve une distribution des valeurs entre Nord et Sud de la France, faisant toutefois apparaître des divergences dans chaque moitié. Au Nord, on lit une structure en axe au départ de Paris, vers Lille, Caen, Le Mans, Dijon. Ailleurs, on identifie des poches dans le Nord-Est du pays, de même qu'une organisation concentrique autour de Tours, ainsi que Clermont-Ferrand. Au sud de la France, la vallée du Rhône ressort alors très fortement, de même qu'un axe se poursuivant depuis Tours jusqu'à Bordeaux. Le rôle transversal de l'axe Atlantique/Méditerranée n'a dans cet indicateur que peu d'impact.

Au cours du temps, une logique concentrique semble se supplanter à une logique radiale : la vallée du Rhône semble dessiner un ensemble distinct, sur un axe Mâcon-Avignon, s'étendant à l'est jusqu'à Grenoble. Au nord de la France, un ensemble semble moins répondre à une logique d'axe qu'en 1860 : il se détache nettement de la côte normande et de la Côte d'Opale. Enfin, un ensemble émerge au cours du temps le long de la vallée de la Loire, ainsi qu'un autre dans la vallée du Lot. Nous misons alors sur cet indicateur, porteur d'une forte différenciation spatiale, qui tend à faire ressortir des structures régionales au cours du temps, nous menant à la construction d'autres indicateurs dans la section suivante. Cette lecture est aussi possible quand on pondère cet indicateur par la population. Nous ne présentons pas ici la confrontation des dispersions de ces deux nouveaux indicateurs dans la mesure où les remarques sont du même ordre que pour les indicateurs multipolaires simples évoqués plus haut.

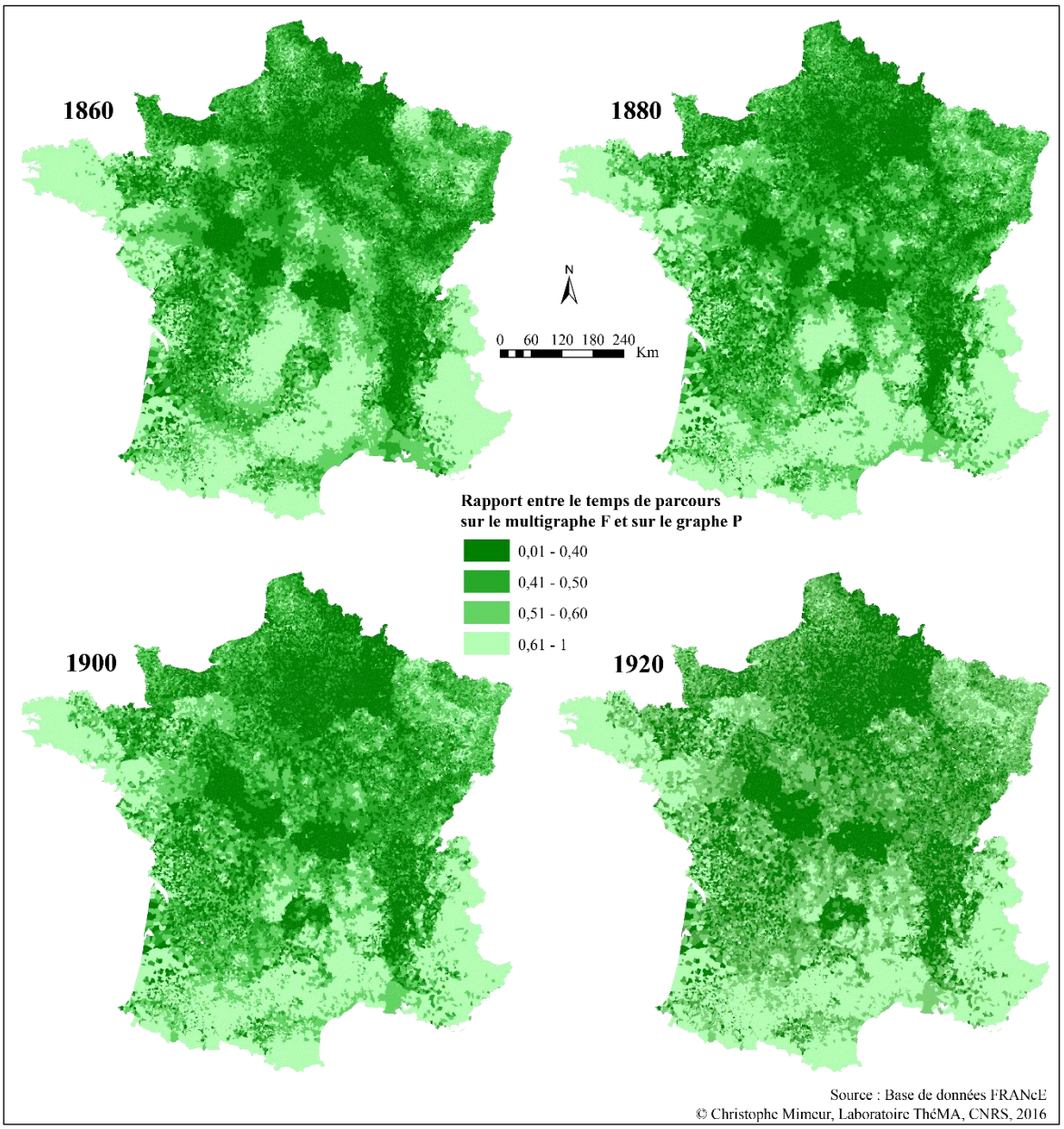


Figure 5. 5. L'effet vitesse en 1860, 1880, 1900 et 1920

La construction d'indicateurs multipolaires participe à la démarche abductive de notre travail, où ils définissent des temps de parcours moyens vers toutes les autres observations ou vers la population française. Dans cette même démarche, la confrontation avec un temps de parcours sans le train participe à l'étude des performances du réseau et du désenclavement de certaines régions, à l'aide d'une cartographie intelligible. Nous verrons plus tard en quoi son pouvoir est plus descriptif qu'explicatif. Pour autant, nous voyons également dans ce type d'indicateurs de nouvelles opportunités afin d'alimenter notre raisonnement sur les échelles de compréhension et de potentielles explications des relations entre croissance démographique locale et infrastructures de transport.

Le concept de *« market potential »* est à attribuer à Chauncy D. Harris en 1954, dans la revue de l'Association des Géographes Américains (Harris, 1954). Il le définit comme une simple adaptation du potentiel d'accessibilité à la population, défini par l'approche pondérée des indicateurs multipolaires définis plus haut. Il pose l'hypothèse que la localisation des industries américaines est fonction de l'accès au marché qu'elles se procurent (Harris, 1954). De cette manière, il propose de distinguer un accès au marché régional d'un accès au marché global. Cette approche s'est largement diffusée ces dernières décennies dans le courant de la Nouvelle Economie Géographique (NEG), dans le développement de la théorie du commerce international (Crozet et Lafourcade, 2009 ; Liu et Meissner, 2015), posant l'hypothèse théorique que la capacité d'échange d'un pays est un déterminant de sa croissance économique. Pour cela, la littérature utilise une mesure de l'accessibilité globale entre les pays, par l'application d'un simple modèle gravitaire, à l'image des indicateurs multipolaires déjà produits dans notre travail, avec dans notre cas une approche par le temps de parcours. L'une des premières applications a consisté à alimenter le débat sur le leadership de la productivité américaine dans le monde au XIXème siècle, attribué au très large marché intérieur, qualifié de domestique dans la démonstration. Par ailleurs, le *« market potential »* a été utilisé sur des analyses historiques portant sur les déterminants de localisation d'industries en Espagne (Combes et al., 2011 ; Tirado, Paluzie et Pons, 2002).

Ce concept et cet arrière-plan théorique prennent leur sens dans la modélisation analytique des relations entre croissance démographique locale et dotation en réseau, dans la mesure où nous posons l'hypothèse que les performances du réseau participent à la concentration des populations dans les communes les mieux dotées. Rappelons ici que la taille démographique des communes, sur le temps long, apparaît comme un *« excellent résumé »* et est souvent corrélée avec *« le nombre d'emplois, d'établissements ou de logements … mais aussi qualitatifs, comme la diversité ou la rareté des activités et la variété des populations présentes »* (Pumain, 1997). Dans ce cadre, la NEG parle d'économies d'agglomération par la croissance des marchés accessibles.

Pour autant, nous misons ici une adaptation conceptuelle du *« market potential »* dans les critères de précision et de pertinence posés dans ce travail. Nous proposons de transposer le concept de *« market potential »* à l'ouverture de l'horizon spatio-temporel permis par le chemin de fer. Nous voyons dans l'utilisation de cet indicateur un moyen de traduire les transformations dans la perception de la vitesse, créées par l'essor de la rapidité différenciée (Ollivro, 2000), en faisant intervenir deux niveaux d'échelle. La littérature de la NEG utilise les termes de marché régional/national (Harris, 1954), domestique/étranger (Head et Mayer, 2004), de composantes intérieures/extérieures (Montout et Robin, 2013). Appliqué à l'accessibilité sur le long terme par le réseau ferroviaire, nous choisissons la terminologie d'un potentiel d'accessibilité multiscalaire, décrivant l'impact du réseau ferroviaire pour élargir l'horizon spatio-temporel d'une commune à un niveau local d'une part, et l'impact du chemin de fer pour s'ouvrir au reste du territoire national, en conservant une approche multipolaire. Le choix des deux niveaux d'échelle à considérer, régional et national est justifié par la littérature portant sur le développement du chemin de fer parle très tôt du développement de chemins de fer régionaux (Bouneau, 1990 ; Caralp-Landon, 1959).

La définition du potentiel multiscalaire s'obtient par le calcul des temps de parcours aux deux niveaux d'échelle définis, pondérés par les données de population de la base de données FRANcE. La *Figure 5. 6* montre le principe de construction de l'indicateur. On définit alors PP_{LOC} d'une commune *i* comme le potentiel d'accessibilité locale par la population vers toutes les autres communes *j* de la région *r*. La distance est représentée dans la *Figure 5. 6* par les traits les plus

épais. Parallèlement, on construit PP_{NAT} comme le potentiel d'accessibilité nationale par la population vers toutes les autres communes *j* extérieures à la région *r*. La distance est alors représentée par les traits les plus fins dans la *Figure 5. 6*.

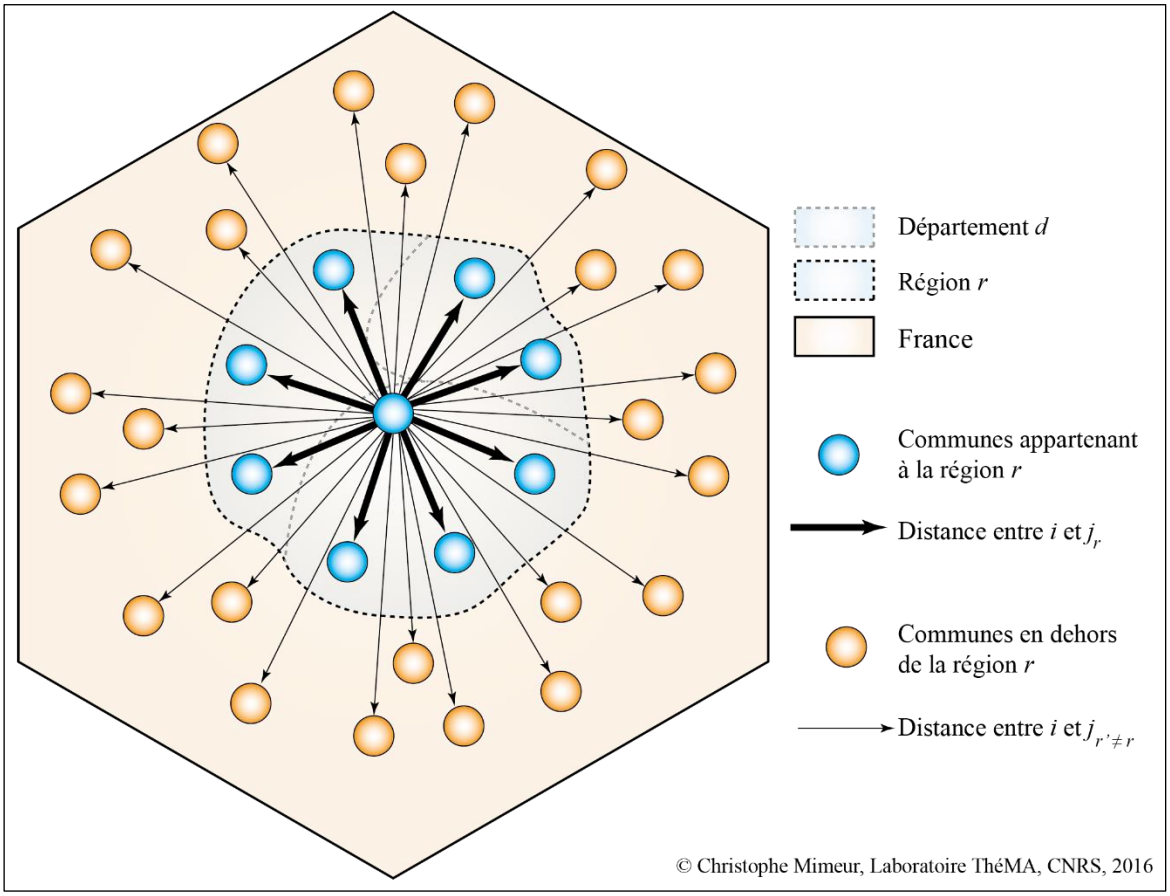


Figure 5. 6. La construction du *market potential* pour la base de données FRANcE

La construction de l'indicateur répond aux équations suivantes :

$$PPLOC_{i\in r} = \sum_{j\in r} \frac{POP_j}{\Delta_F(i,j)} \quad \text{et} \quad PPNAT_{i\in r} = \sum_{j\in r'\neq r} \frac{POP_j}{\Delta_F(i,j)} \tag{18}$$

Pour l'heure, nous avons jeté les bases d'une complexification de la dotation en réseau par des indicateurs d'accessibilité qui définissent arbitrairement des destinations à rejoindre ou alors qui sont construits de manière indépendante de toutes destinations prédéfinies. Si nous posons l'hypothèse d'un raffinement des analyses à partir de cette complexification, notre réflexion se trouve vite heurtée par les biais inhérents à toute démarche qui lie réseau et territoire et qui ont souvent été décriés dans la littérature de la controverse des effets structurants (Offner, 1993a). Il s'agit alors d'envisager une modélisation de l'effet réseau qui prend en compte l'endogénéité des relations entre accessibilité et population : un regard vers l'économie historique nous permet de définir la spécification que nous allons donner à l'étude de ces relations.

5.2. Un rapprochement vers l'économie historique

Des effets de long terme ont déjà été étudiés quant à l'impact des infrastructures de transport : on a cité dans notre première partie les travaux pionniers de Robert Fogel et Albert Fishlow. Les apports de l'économie historique (Fishlow, 1965 ; Fogel, 1962) nous ont d'abord permis d'identifier l'importance du recul temporel quant aux effets structurants des infrastructures de transport. Nous identifions aussi des pistes méthodologiques bien que leur approche se soit souvent limitée à un niveau macroéconomique. Les débats disciplinaires suscités par l'économie historique viennent alimenter notre réflexion épistémologique et nous conduisent à choisir une méthodologie économétrique, qui associe l'information géohistorique collectée et capitalisée jusqu'à ce stade, en plus de donner des pistes d'investigation de la dynamique temporelle.

Si de telles études sur les relations entre infrastructure et utilisation du sol sur le temps long ont déjà été conduites aux Etats-Unis surtout, et dans d'autres pays d'Europe (Kasraian et al., 2016), il n'existe pas à notre connaissance de telles entreprises à l'échelle de la France entière sur notre période d'étude. L'objectif est alors d'identifier les apports de la *cliométrie* pour utiliser notre propre matériau empirique, en venant alimenter les nombreuses études empiriques, tout en visant la construction de nouvelles connaissances géohistoriques du lien entre croissance de la population et accessibilité des territoires. Pour cela, nous identifions dans le champ cliométrique un cadre d'étude pertinent qui permet aussi de limiter les nombreux biais inhérents à l'étude des liens entre transport et territoire.

5.2.1. **L'économie historique : entre cliométrie et empirisme**

Notre premier chapitre sur la littérature traitant de l'effet structurant des infrastructures de transport a montré comment les études menées sur le court terme étaient nombreuses, à des échelles réduites, mais souffraient d'un manque de recul quant aux nombreux éléments à prendre en considération. C'est finalement aussi le point de départ identifié par Robert Fogel, quand il cherche à répondre à l'axiome d'*« indispensabilité »* des infrastructures. Sa démarche consiste à une *« projection quantitative des sciences sociales dans le passé »* (Diebolt, 2004). Il mobilise alors la théorie économique pour éclairer des problèmes jusqu'alors non-résolus. Dans le monde anglo-saxon, cette *« new historic history »* apparaît au début de la seconde moitié du XXème siècle. Outre les travaux portant sur le rôle du chemin de fer aux Etats-Unis, Robert Fogel a également montré que les conditions de vie des esclaves du Sud des Etats-Unis étaient meilleures que celles des travailleurs libres dans le secteur agricole, dans le Nord des Etats-Unis. L'économie historique a finalement trouvé sa consécration avec l'obtention du Nobel d'Economie par Robert Fogel et Douglass North en 1993, qui emploient des méthodes quantitatives pour étudier les changements économiques et institutionnels. Elle est souvent considérée comme une science de synthèse entre les sciences historiques et les sciences économiques.

Du côté des historiens français, les rapprochements avec la science économique obéissent surtout à deux autres écoles. La première est marxiste, et la seconde est incarnée par Le Roy Ladurie

avec des modèles *« écologico-démographique »*, mettant au centre des déséquilibres démographiques des variables environnementales et la fourniture d'aliments (Berque, 1981). Il participe au courant de la *« Nouvelle Histoire »* au début des années 1970, par des analyses micro-historiques.

La troisième école identifiée dans l'historiographie française est celle qui se rapproche de l'école américaine d'histoire quantitative. Elle est nommée *cliométrie*, définie comme la *« combinaison d'explications causales dans des modèles, avec ou sans situation contrefactuelle, qui peut être opérante dans une situation historique donnée »* (Diebolt, 2016). L'emploi du contrefactuel consiste à imaginer un monde sans chemins de fer, comme le suggère Fogel pour la situation américaine en 1890. Il s'agit d'une démarche différente des études sur le lien entre transport et territoire à l'échelle des projets, qui font l'objet d'analyses *ex-ante* et *ex-post* (Delaplace, 2014). Mais, dans notre cas, le multigraphe *F* peut proposer une situation contrefactuelle théorique, celle d'une marche à pied généralisée. Toutefois, nous l'envisageons davantage comme une alternative et comme un élément permettant de rendre compte de l'effet des vitesses sur le réseau ferroviaire : à partir d'une situation homogène, comment la diffusion du réseau ferroviaire et sa hiérarchisation progressive participe-t-elle à la différenciation des territoires. Ainsi, si l'histoire des mobilités de l'époque montre que sur des moyennes et courtes distances, la marche à pied domine, alors sur les longues distances, le développement du réseau postal est une autre alternative (Verdier et Bretagnolle, 2007), qui s'est développée sous l'Ancien Régime et qui tend à se stabiliser au lendemain de la Révolution Française. L'exigence comportementale exigée dans la construction des indicateurs d'accessibilité d'une part, dans notre positionnement constructiviste d'autre part, vient fragiliser l'emploi du contrefactuel à partir des données dont nous disposons. Au demeurant, la définition de Claude Diebolt démontre que le recours au contrefactuel n'est pas systématique.

D'ailleurs, son emploi est presque toujours critiqué par les historiens. Si les conclusions de Fogel ont été parfois reçues par un sentiment de provocation, alors ses détracteurs se sont dit préférer un *« niveau d'apparence faiblement scientifique qu'est la relation détaillée d'un exemple choisi, […] [en espérant] saisir la réalité passée mieux que ne l'ont permis les généralisations de ses prédécesseurs »* (Berque, 1981). En France, ce courant de l'économie et/ou de l'histoire a souvent été qualifié de matérialiste, en laissant de côté des manifestations superstructurelles, quand le courant dominant de l'historiographie se réclamerait de s'occuper *« davantage des humbles que des grands »* et voient dans la complexification mathématique un manque de prise en considération de la qualité de la donnée (Berque, 1981). Parmi les plus fortes critiques adressées, *« on fait aussi remarquer que les conclusions tirées sont plus souvent négatives que positives : elles rejettent des interprétations jusqu'alors acceptées, plus souvent qu'elles n'affirment des certitudes au-delà de tout doute raisonnable »* (Heffer, s. d.).

Nous reconnaissons toutefois dans ces vifs débats épistémologiques une scène sur laquelle la controverse portant sur les effets structurants peut trouver sa place : loin de chercher de telles conclusions provocatrices, nous retrouvons toutefois les enjeux quantitatifs de la géohistoire dans ceux de la cliométrie et de l'histoire quantitative. Elle participe, comme la géohistoire, *« à reconstituer l'évolution des sociétés passées non plus à partir de faits anecdotiques et individuels, d'évènements politiques, diplomatiques ou militaires exclusivement, mais en tenant compte du rôle historique des structures matérielles, des faits de masse et des forces inconscientes »* (Diebolt, 2001). Dans les humanités numériques, le rapprochement entre les *Spatial Humanities* et la *New Economic History* peut alors s'appuyer sur les nombreuses avancées en SIG, qui viennent enrichir

la diversité, la validité et la qualité de la donnée (Kasraian et al., 2016). Le terme de *cliométrie* fait non seulement référence à la mesure, mais à la déesse grecque Clio, qui est la muse de l'Histoire. Mais n'oublions pas que si Clio est traditionnellement représentée avec une clepsydre, représentant le temps passé ; elle est souvent accompagnée du globe terrestre, rendant légitime un rapprochement de la géographie. Ainsi, mis en perspective dans notre démarche géo-méthodologique, ce rapprochement participe au débat entre théorie et empirisme. Si les tenants de la cliométrie penchent davantage du côté de l'utilisation de la théorie économique pour investiguer des problèmes non-résolus dans le passé, d'autres y voient en revanche principalement une démarche empirique. On peut y lire un balancier chronologique où *« jadis la discipline se définissait d'abord par une utilisation explicite de la théorie et, en second lieu seulement, par l'accumulation de nouvelles preuves »* (Heffer, s. d.), alors qu'aujourd'hui les avancées permettent une large collection de données.

De ce point de vue, la démarche empirique en cliométrie consiste à délaisser la modélisation structurelle pour tester des mécanismes de force afin d'établir des relations causales à partir de larges données diversifiées, collectées ou construites de manière inductive (Diebolt, 2005), en restant attaché au fait que des facteurs externes participent aussi à ce processus (Kasraian et al., 2016). Elle trouve alors un succès croissant avec le développement de séries chronologiques et temporelles longues, faisant appel de plus en plus aux sciences informatiques, elle aussi[25], dans une démarche exploratoire. On peut citer alors les travaux français d'Henri Gilles, qui s'attache à montrer dans quelle mesure les disparités régionales des morts de la Grande Guerre s'expliquent par des facteurs démographiques. Ainsi, à partir d'une analyse *« qui a mêlé facteurs démographiques, économiques et géographiques, la thèse de régions plus sacrifiées que d'autres semble devoir être nuancée »* (Gilles, Guironnet et Parent, 2014). C'est dans cette lignée que nous voyons dans les principes de la nouvelle économie historique l'opportunité de mesurer la force de l'effet structurant sur le temps long, en dépit des interprétations déjà existantes et souvent tranchées. Dans notre questionnement, cette opportunité est renforcée par la capacité à diversifier les indicateurs d'accessibilité, de même qu'à mobiliser des informations géohistoriques capables d'être intégrées dans notre modèle conceptuel de données. Nous retenons enfin des avancées de l'économétrie sa capacité à prendre en compte la dynamique temporelle : dans ce penchant empirique de l'économie historique, *« reste à estimer les équations du modèle à l'aide des techniques statistiques adéquates ; le calcul de régressions permet de distinguer quelles sont les variables explicatives pertinentes »* (Heffer, s. d.).

Finalement, une des propriétés de la cliométrie nous semble centrale dans notre positionnement géohistorique quantifié : l'utilisation des techniques et les résultats qui en ressortent, de toute évidence, participent au débat par l'introduction de nouvelles perspectives qui n'existaient pas jusqu'alors (Diebolt, 2016). Ainsi, devant la diversité des techniques économétriques, nous allons nous attacher à montrer les apports et les limites des techniques les plus couramment utilisées dans l'étude de long terme des liens entre transport et territoire, de même qu'à identifier quelques grands principes économétriques qui vont influencer le choix du type de modélisation ainsi que le choix des variables à intégrer dans le modèle.

[25] Parmi les revues spécialisées de la cliométrie, on peut alors citer *Journal of the Association for History and Computing*, qui a existé entre 1998 et 2010 ; de même que la revue *Explorations of Economic History*.

5.2.2. **Précautions et méthodes pour modéliser la croissance par le réseau**

Nous visons à définir une modélisation des données issues de FRANcE pour évaluer l'effet structurant de l'infrastructure de transport sur la croissance de la population. Est-ce que la *« dotation en réseau »* favorise la croissance démographique ? De premières avancées ont été effectuées par la diversification des indicateurs d'accessibilité en mobilisant les graphes construits dans le chapitre précédent.

Plus largement que pour les seules infrastructures de transport, de nombreuses approches économétriques se sont attachées à étudier de manière empirique le lien entre l'investissement public et la croissance économique (Banister et Berechman, 2004), s'inscrivant dans la lignée de la théorie de la croissance endogène (Barro, 1990). Nous proposons ici de transposer les limites méthodologiques de ces approches soulevées par la littérature à notre question, parce qu'elles identifient les précautions que le chercheur doit prendre en compte quand il entreprend une modélisation économétrique, rendant compte des liens entre croissance et infrastructure. Ensuite, nous montrerons quelles approches ont été privilégiées dans l'étude du couple réseau/territoire, et laquelle nous retenons dans notre perspective de long terme.

L'objet de la controverse scientifique portant sur les effets structurants est un problème auquel l'économètre est couramment confronté : il s'agit de l'endogénéité, issue de la dynamique d'un système. Définie comme la corrélation entre le paramètre et le résidu, elle caractérise une surestimation des capacités explicatives d'une variable, faisant écho aux difficultés méthodologiques pointées par Jean-Marc Offner (1993a). Pour autant, il existe des techniques statistiques qui permettent de prendre en compte ces imperfections pour rendre compte d'estimations plus justes. Nous ne traitons pas pour le moment du biais de simultanéité, qui sera l'objet d'un développement particulier à la fin de ce chapitre : il est caractéristique des problèmes du sens de la causalité de la relation, car des méthodes économétriques se sont développées pour limiter ce biais, qui trouve un écho important dans notre positionnement géohistorique.

Parmi les autres biais d'endogénéité, des études nationales venant compromettre (Gramlich, 1994) la relation positive (Aschauer, 1989) entre l'investissement public et la croissance économique ont montré qu'elle ne prenait souvent pas en compte des effets exogènes d'éviction. Il s'agit alors d'un biais d'hétérogénéité inobservée. L'une des principales raisons d'un tel biais est l'omission de variables, où la *« crainte de la fameuse variable cachée responsable d'associations statistiques fallacieuses hante le chercheur en sciences sociales »* (Robin, 2000). Le risque est bien que l'estimation obtenue tienne en effet compte de l'effet propre lié à l'infrastructure, mais aussi d'autres qui sont exogènes. Appliquées aux infrastructures de transport et à la croissance de la population, nous postulons que des variables de l'environnement géographique, constantes dans le temps, peuvent participer à limiter ce biais dans la relation causale. De nombreuses études introduisent alors des variables portant sur une offre locale d'aménités de même que sur la situation géographique[26] des unités d'observation (H. Hanson, 2005). Dans notre période d'étude, nous

[26] Dans ces estimations, l'auteur prend ainsi en compte, entre autres, le taux d'ensoleillement, la vitesse du vent, la quantité de précipitations, la présence d'un point d'eau.

posons l'hypothèse que ces variables de situation permettent de mieux estimer le rôle du réseau, indépendamment d'un accès à d'autres réseaux et à certaines aménités par exemple. Une autre source d'endogénéité est l'erreur de mesure : pour autant, notre démarche consiste à définir des critères de pertinence et de précision dans la construction des paramètres à intégrer dans le modèle. De même, *« une erreur de mesure sur la variable dépendante diminue la précision des estimateurs mais ne les biaise pas asymptotiquement »* (Robin, 2000) : cela signifie que dans une proportion raisonnable, l'erreur de mesure ne conduit pas nécessairement au rejet de la modélisation.

Au-delà des précautions à prendre dans la modélisation des relations entre croissance et développement du réseau ferroviaire, il s'agit d'évoquer maintenant la pluralité des méthodes utilisées pour identifier des effets de long terme. Pourtant, la comparabilité des effets entre les différents pays ou les différents types de réseaux se heurte à la grande diversité méthodologique d'une part, mais aussi aux échelles spatiales et temporelles retenues, ainsi qu'à la granularité mise en œuvre d'autre part (Kasraian et al., 2016).

La méthode la plus classique est celle des Moindres Carrés Ordinaires. Il s'agit d'une régression linéaire qui est souvent utilisée comme un préalable à des modélisations plus spécifiques (Atack et al., 2010 ; Hornung, 2012b ; Kasraian et al., 2016 ; Silveira et al., 2013). Même si elle peut être associée à de simples coefficients de corrélation, l'utilisation de cette méthode ne peut qu'en revanche se limiter qu'à des *« snapshot »*, en fonction des intervalles disponibles dans le corpus de données mobilisé, décrivant des coupes transversales. Par ailleurs, la mise en exergue d'évolution ne peut également se faire que par paire de dates retenues : elle permet néanmoins de saisir des tendances, des comportements qui permettent d'affiner le choix des variables ou des périodes plus spécifiques à retenir. En revanche, elle atteint rapidement ses limites dans notre cas quand elle ne prend pas en compte des effets individuels sur le temps long, susceptibles de faire émerger des trajectoires de population différenciées.

Une seconde méthode, plus exploratoire, a été utilisée par Jeremy Atack à propos du développement du réseau ferroviaire dans le MidWest américain (Atack et al., 2010) : il s'agit de la méthode des doubles différences. Couramment utilisée dans l'évaluation des politiques publiques (Bertrand, Duflo et Mullainathan, 2004), elle trouve sa place dans les cadres définis des bilans prévus par la Loi d'Orientation des Transports Intérieurs (1982) portant sur les infrastructures de transport. Elle consiste à identifier deux groupes : un groupe de traitement, qui dans ce cas est impacté par le réseau, et un groupe de contrôle, dont la dotation en réseau est restée nulle. Pour autant, cette méthode a été utilisée par l'auteur sur une seule décennie, entre 1850 et 1860. En dépit d'une démonstration d'un impact positif du réseau sur la croissance de l'urbanisation plutôt que sur la croissance de la densité de population, ce type de modélisation ne permet pas non plus de prendre en compte des effets individuels sur le temps long, capables de différencier des effets de court terme ainsi que des effets de long terme, issus d'une tendance globale.

Une troisième voie porte vers les tests de causalité, le plus connu étant celui de Granger. Ils permettent de mesurer le sens de la causalité, et même souvent d'établir une bi-directionnalité des relations (Beyzatlar, Karacal et Yetkiner, 2014). Pour autant, leur principale limite réside dans l'utilisation de données bivariées, qui montre rapidement le risque d'une grande hétérogénéité inobservée (Tranos, 2012). Ainsi, appliqué à la relation entre croissance économique et linéaire

d'infrastructures dans l'Europe des 15, le test montre seulement une causalité dans les deux sens, mais aussi une forte endogénéité entre les deux variables étudiées (Beyzatlar, Karacal et Yetkiner, 2014). Appliqué à la relation entre croissance économique régionale et cyber-infrastructure, il penche plutôt sur un impact positif de l'infrastructure, en dépit d'une grande hétérogénéité des régions qui n'est pas pris en compte dans le modèle (Tranos, 2012).

A partir de l'ensemble de ces méthodes, nous identifions des enjeux propres à notre questionnement. L'utilisation de deux variables décrivant d'une part le système de peuplement et d'autre part le système de transport semble se heurter de manière évidente au biais d'endogénéité : une part du pouvoir explicatif est alors déjà contenue dans la variable à expliquer. Dans ce chapitre, le sens de la causalité que nous souhaitons tester est bien l'évolution de la population en lien avec la dotation en réseau. Alors, il faut d'abord combler l'hétérogénéité inobservée. Dès lors, nous devons envisager d'autres variables, pouvant expliquer la croissance de la population, mais qui soient indépendantes, le plus possible, de la dotation en réseau. Dans notre raisonnement géohistorique, nous nous emparons à nouveau de notre positionnement constructiviste, en faisant appel à de nouveaux critères de pertinence et de précision qui peuvent participer à la croissance de la population. Pour autant, nous devons également faire face aux difficultés inhérentes à une étude dans le passé d'une part, de long terme d'autre part : il faut disposer des données qui couvrent l'ensemble de l'espace étudié et de la période retenue. Cette réflexion est l'objet de la section suivante.

De même, les modélisations économétriques décrites jusqu'à maintenant donnent une large place à l'étude transversale, ou, dans le meilleur des cas, à une étude par intervalle. Par ailleurs, à propos des travaux de Pascale Dancoisne, Gabriel Dupuy souligne comment la dimension temporelle doit être investiguée de manière plus approfondie, étant donné les connaissances déjà établies à partir d'indicateurs globaux issus de la théorie des graphes (Dupuy et Crews, 1991). En effet, sur le long terme, les caractéristiques du réseau, les conditions d'exploitation, et l'évolution de leur différentiel peuvent avoir, sinon un impact indirect, un impact direct sur la relation entre réseau et territoire. Ainsi, jusqu'à maintenant, les études que nous avons mentionnées ne font pas appel à l'évolution des vitesses le long des réseaux, mais s'intéressent surtout à sa structure et sa diffusion dans l'espace (Kasraian et al., 2016).

Dans l'hypothèse de l'identification d'effets de court terme d'une part, de long terme d'autre part, alors nous devons faire appel à un autre type de modélisation. Nous parions à ce stade de nos investigations sur le potentiel apport de la dimension diachronique dans la modélisation. Cela nous amène à délaisser dans notre raisonnement les modèles strictement spatiaux. Alors, l'économétrie par panel semble détenir un potentiel dans l'étude du lien entre croissance et réseau.

5.2.3. **Les grands principes de l'économétrie de panel**

Dans sa *« Petite apologie des données de panel »*, Brigitte Dormont voit dans cette méthode économétrique *« la source d'une réflexion souvent fructueuse permettant d'approfondir le phénomène étudié »* (Dormont, 1989). Les données de panel se définissent comme des données relatives à des statistiques observées à plusieurs reprises dans le temps : au final, c'est une répétition de coupes. Il s'agit ici de définir les avantages et inconvénients d'une telle démarche, et d'évaluer son apport dans notre questionnement.

Ce type de modélisation rencontre un vif succès dans les approches managériales et financières, s'appuyant sur un suivi systématique des performances des entreprises, de leurs filiales, de leurs clients ou de leurs commerciaux. Dans notre cas, nous disposons à la fois de la croissance de la population et de la dotation en réseau, pour chaque décennie de notre période d'étude. L'intérêt d'une telle démarche est l'observation de ces données dans une double dimension : entre les unités d'observation, ainsi qu'entre les coupes *(Figure 5. 7. b.)*, alors que la méthode des moindres carrés ordinaires n'autorise que la multiplication des régressions pour chaque couple de périodes *(Figure 5. 7. a.)*. Mécaniquement, la taille de l'échantillon augmente considérablement, puisque tous les enregistrements à toutes les périodes sont pris en compte dans l'estimation de la régression. Les données individuelles permettent donc un suivi : par exemple, si deux coupes permettent de mesurer un taux de chômage à 10 %, le panel permet de savoir s'il s'agit des mêmes personnes. A partir d'une équation économétrique simple :

$$y = a + b.x + c.z + \varepsilon \tag{19}$$

où la variable dépendante *y* a pour explicatives les variables *x* et *z*, de coefficients *a*, *b*, *c* et d'aléa *e*, on peut déduire l'équation appliquée à des données de panel, pour une observation *i* à une période *t* :

$$y_{it} = a + b.x_{it} + c.z_{it} + \varepsilon_{it} \tag{20}$$

Si *a priori*, la modélisation ne change que marginalement, l'originalité d'une telle approche réside dans la spécification de l'aléa ε_{it} :

$$\varepsilon_{it} = u_i + v_t + w_{it} \tag{21}$$

où u_i est un terme dépendant de l'individu *i* constant au cours du temps, v_t est un terme ne dépendant que de la période *t* et w_{it} est un terme aléatoire croisé.

La modélisation sur données de panel permet alors de traiter un certain nombre de biais inhérents à l'étude économétrique. Ainsi, par une information riche *« tant en quantité qu'en variabilité »* (Dormont, 1989), la précision des estimations s'en trouve améliorée. Si de simples régressions linéaires peuvent faire apparaître une colinéarité entre les variables explicatives, alors la variabilité entre les individus et entre les périodes vient la diminuer *(Figure 5. 7. b.)*. Elle permet de prendre en compte des différences structurelles entre les observations – la présence d'aménités, les contraintes liées à l'espace géographique par exemple. Elle prend en compte aussi des évolutions

de type macro-échelle sur le temps long – la croissance générale de la population, les évolutions des structures socio-économiques. Enfin, la variabilité inclut les comportements propres à chaque individu, qui est donc à la fois perçue dans les données prises en coupe ou de manière longitudinale.

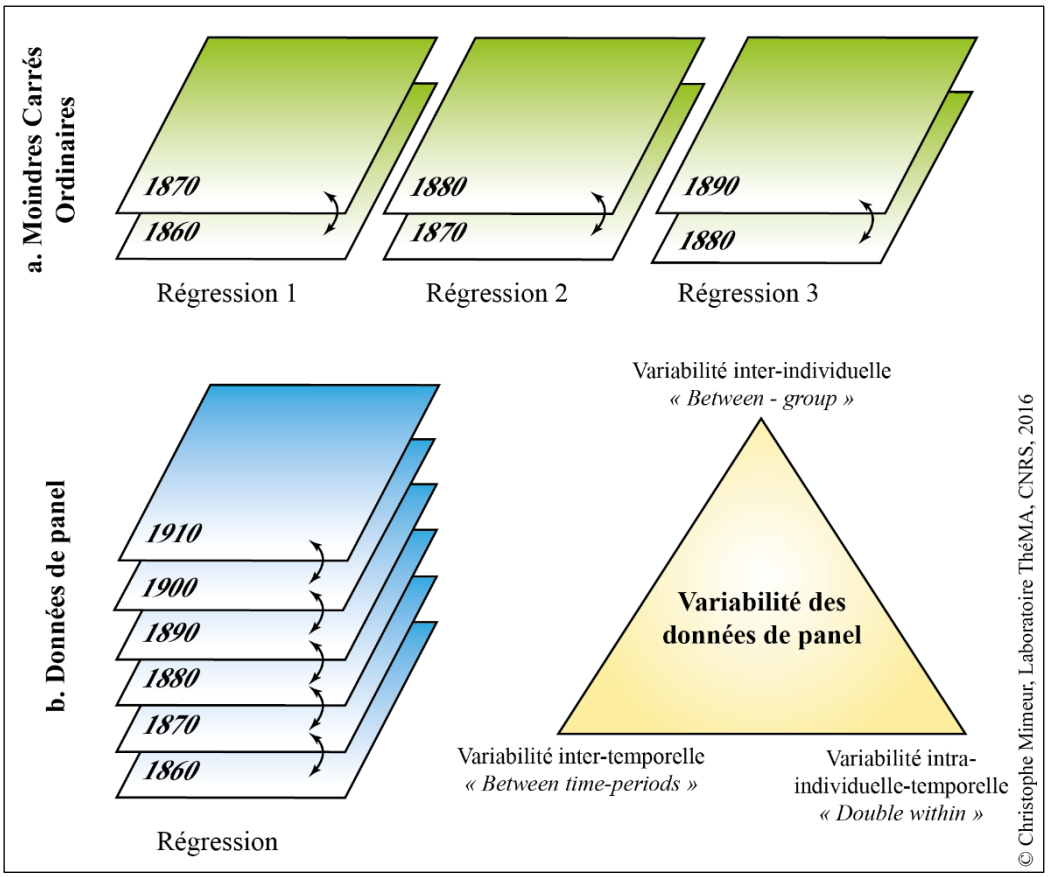


Figure 5. 7. Des moindres carrés ordinaires aux données de panel

En plus de la dimension diachronique, l'une des principales valeurs ajoutées des données de panel est la prise en compte de l'hétérogénéité inobservée. Quand les trois types de variabilité sont accumulés, alors l'erreur dans l'estimation des coefficients dans une simple régression par les moindres carrés ordinaires peut conduire à des résultats contre-intuitifs *(Figure 5. 8)*. Trois méthodes permettent de prendre en compte cette hétérogénéité inobservée, alors même que le modélisateur doit minimiser le nombre de variables omises, quand elles sont effectivement mobilisables. Ces trois méthodes consistent à spécifier le terme d'erreur ε_{it} dans lequel on suspecte le biais :

- La première est qualifiée de « naïve » parce qu'elle met bout-à-bout les données de panel sans se préoccuper de leur nature ni de l'hétérogénité inobservée. Même si les estimations sont souvent fausses, cette méthode permet d'estimer l'apport des autres méthodes, et de parfois les arbitrer ;

- La seconde est le modèle à « effets fixes » : il s'agit d'effets constants, qualifiés d'inobservables mais permanents, qui vont être appliqués en fonction de *i* et de *t* : finalement, les effets de u_i et de v_t se retrouvent sur la constante. Elle prend en compte l'hétérogénéité au niveau de l'ordonnée à l'origine *(Figure 5. 8. a.)*. On peut de même décomposer les effets fixes sur les individus ou sur les périodes, même si la littérature montre que l'effet sur la période est souvent inexistant ou mineur ;

- La troisième est le modèle à « effets aléatoires » : les effets de u_i et v_t sont distribués de manière aléatoire, et s'exercent donc directement sur le terme d'erreur. La méthode de régression vise alors à qualifier ces effets pour affiner les estimations par la suite en saisissant l'hétérogénéité au niveau des pentes *(Figure 5. 8. b.)*. Des tests permettent en outre d'éprouver les deux hypothèses et d'opérer un choix entre effet fixe et effet aléatoire : ces tests seront conduits dans notre troisième partie.

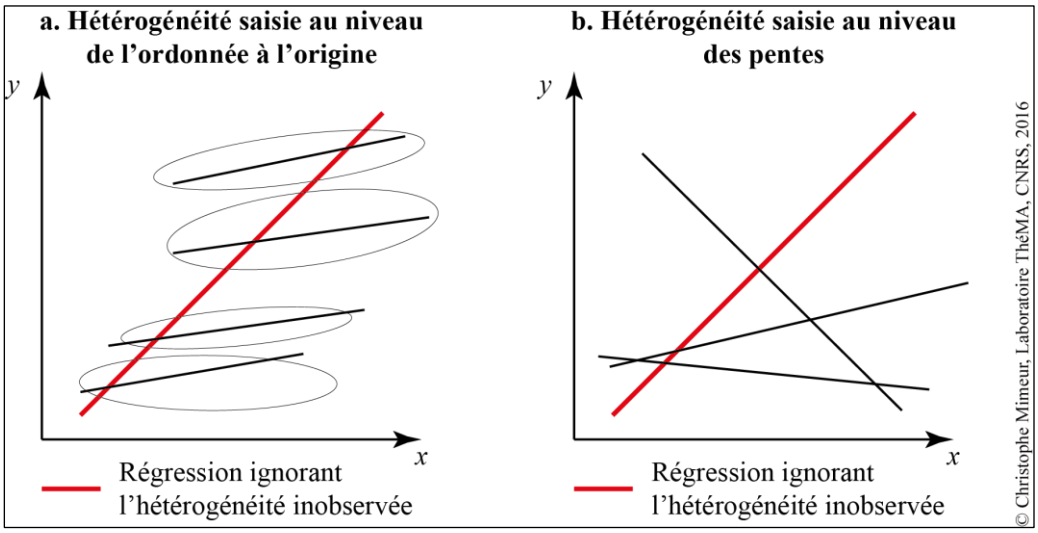


Figure 5. 8. L'hétérogénéité inobservée dans la modélisation en données de panel

Brigitte Dormont propose alors un tableau récapitulatif des avantages et des inconvénients de l'utilisation de données de panel *(Table 5. 2)*. Quand on le confronte à notre question de recherche, on en retient d'abord l'information à une micro-échelle. Bien sûr, si la diachronie est prise en compte par une dimension des panels, alors l'importance de la dimension de coupe garantit une plus grande robustesse des estimations.

Pour autant, les limites de leur utilisation résident surtout sur les estimations intra-individuelles, où le biais de variables omises est fréquent, ce qui incite le modélisateur à recourir à un maximum de variables permettant de rendre compte d'évolutions intra-individuelles, c'est-à-dire variantes dans le temps. Les autres limites évoquées par l'auteur montrent à quel point la réflexion sur l'échantillon à prendre en compte dans les estimations sera déterminante dans leur robustesse. Nous comptons pour cela sur des analyses préalables pour repérer les points aberrants et caractériser

l'échantillon le plus proche de notre questionnement, ce sera l'objet du début de la troisième partie de la thèse.

Panorama de l'utilisation des données de panel

Avantages		Inconvénients	
Caractéristiques des données	**Qualité du modèle**	**Caractéristiques des données**	**Difficultés matérielles**
Données plus nombreuses et plus variables	Estimations plus précises	Moins bonne qualité dans la mesure des variables	Biais de variables omises pour les estimations intra-individuelles
Information au niveau micro-échelle	Données plus adéquates pour l'étude des comportements	Méthodes basées pour la plupart sur l'hypothèse d'un échantillon cyclindré	Difficulté à obtenir le même nombre de périodes pour chaque individu
Importance de la dimension "coupe"	Plus grande robustesse de certaines estimations	Lourdeur des travaux de constitution des échantillons	
Possibilité de mener, à partir du même corps de données, des estimations dans différentes dimensions	Contrôle des spécifications estimées	Certains cas méthodologiques non résolus	

D'après DORMONT, 1989 © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 5. 2. Panorama de l'utilisation des données de panel

A l'issue de cette section, nous avons vu comment la question des effets de long terme des infrastructures de transport a fait l'objet d'études utilisant la modélisation économétrique (Kasraian et al., 2016). Si la cliométrie s'est d'abord développée par la modélisation de la théorie économique et la mesure du passé, alors la croissance des données historiques disponibles a bouleversé les approches, parmi lesquelles les plus empiriques sont devenues nombreuses. Les méthodes sont variées, de la simple régression linéaire par les moindres carrés ordinaires aux tests de causalité. Par ailleurs, nous avons identifié dans la modélisation des données de panel de nombreuses opportunités pour prendre en compte la variabilité entre les unités d'observation ainsi qu'entre les rythmes identifiés entre les périodes, répondant à notre questionnement sur l'effet de l'accélération sur le réseau, de même qu'une capacité à limiter le biais de variables omises. Alors nous pouvons poser l'équation d'un modèle de relation utilisant la dimension du panel entre la croissance de la population POP_{it} et la dotation en réseau RES_{it} :

$$POP_{it} = u_i + v_t + b.RES_{it} + c.X'_{it} + w_{it} \quad (22)$$

Nous avons déjà abordé la définition des variables qui qualifient RES_{it} à partir de l'utilisation des indicateurs qui ressortent des graphes. Toutefois, les raisonnements de cette section nous poussent à définir des variables X'_{it} susceptibles d'avoir un pouvoir explicatif élevé quant à la croissance démographique locale, en faisant appel à la diversification de l'information géohistorique. Il s'agit aussi de limiter le biais d'endogénéité, souvent révélé dans les relations entre réseau et territoire, que nous éprouvons avec une exploration par les variables instrumentales.

5.3. Mobiliser des informations géohistoriques de long terme pour la base de données FRANcE

Les nombreuses études empiriques portant sur la période contemporaine font appel au *« souvenir statistique »* souvent issu des démarches d'observatoires permanents, qui concernent les autoroutes (Bérion, Joignaux et Langumier, 2007) ou les lignes à grande vitesse (Bonnafous, 2014). Il s'agit alors d'enrichir la modélisation des relations entre transport et territoire par des variables participant à l'explication de la variable dépendante : l'objectif est de limiter un biais dans l'estimation du paramètre d'intérêt, ici la dotation en réseau. L'ajout de telles variables permet de limiter le biais de variables omises, en contrôlant l'effet parasite d'un ensemble d'autres facteurs susceptibles d'influencer la variable démographique, même si la spécification en panel y prend aussi sa part.

Sur le temps long, nous sommes ici face aux difficultés inhérentes à la donnée géohistorique, dont les sources sont souvent hétérogènes, et l'information souvent incomplète (Bretagnolle, 2014 ; Delaplace, 2014). Nous nous retrouvons donc face au challenge émergent des SIG-H, où nous misons sur la diversification de l'information, par l'utilisation de données existantes, sans avoir recours à une nouvelle étape de collecte de la donnée. Nous discutons ici des modalités de leur inscription dans la base de données FRANcE, afin d'affiner et préciser la modélisation que nous envisageons dans ce chapitre.

5.3.1. **Intégrer des variables de contexte local**

La modélisation des relations dans la littérature des effets de long terme se heurte souvent à la donnée disponible (Duranton et Turner, 2012 ; Hornung, 2012b), qui restreint la plupart du temps l'angle de vue, tant d'un point de vue spatial que temporel. Le parti-pris de ce travail est bien de travailler à de larges échelles spatiales et temporelles : la recherche d'une donnée historique à de si larges échelles fait face à d'autant plus de difficultés. La première étape consiste à identifier des variables d'environnement, caractéristiques du site d'une unité d'observation d'une part, et de sa situation d'autre part. Dans la modélisation que nous proposons, ces variables tiennent le rôle de

variables de contrôle. L'objectif de l'utilisation de telles variables vise à *« capter les différences dans la croissance future de la population, ainsi que la progressivité »* de l'unité d'observation (Hornung, 2012b). A ce stade, nous définissons donc des variables qui sont invariantes dans le temps, et qui vont donc avoir un effet fixe. Elles permettent de contribuer à explorer la modélisation dans un premier temps, mais ne pourront pas être intégrés dans la modélisation de panel dans la mesure où elles sont invariantes.

Parmi elles, la littérature fait une grande place aux réseaux concurrents : Jeremy Atack pose comme variables de contrôle l'accès à un canal, l'accès à une voie naturelle navigable ainsi que l'accès aux grands lacs du Midwest américain (Atack et al., 2010). De la même manière, Erik Hornung prend en compte l'accès à une voie d'eau et l'accès à une route (Hornung, 2012b). On voit donc ici comme le choix des variables est largement dicté par le contexte géographique et historique d'une part, par la disponibilité de la donnée d'autre part. D'autres études contemporaines portant sur la croissance économique des territoires montrent comment les voies de communication naturelles et les conditions climatiques peuvent aussi avoir un effet notable sur la croissance différenciée, alors que celle des villes fait intervenir une dotation en aménités urbaines (Duranton et Turner, 2012 ; Ottaviano et Pinelli, 2006).

L'histoire des mobilités en France et des pratiques (Flonneau et Guigueno, 2009) nous montre comment la route a constitué pendant longtemps le seul paysage de la vitesse (Guigueno, 2009). En effet, Bernard Le Sueur affirme qu'au début du XIXème siècle, nous avons à faire à *« un système de transport à deux composantes, route et voie d'eau »* (Le Sueur, 1989). A ce titre, nous nous emparons des chemins décrits dans la Carte de Cassini, qu'ils soient parcourus à pied pour les petites ou moyennes distances ou par la traction animale pour de plus longues distances. A ce stade, nous ne prenons pas en compte la croissance d'un réseau postal, que l'on doit considérer comme un système de transport à part entière avec une organisation et une logique, caractéristiques de la notion de réseau (Verdier et Bretagnolle, 2007). Elle le sera dans un second temps de la complexification méthodologique. La carte de Cassini est souvent considérée comme un *« document de référence pour toute étude historique »* (Robert et Costa, 2009). Elle constitue la première carte levée à l'échelle de la France entière, et son édition court de 1749 jusqu'à 1815, période pendant laquelle des corrections ont été effectuées de manière continue. Son intérêt réside *« principalement dans la localisation systématique des principaux éléments d'organisation du paysage »* (Robert et Costa, 2009). Concernant le tracé des voies de communication, *« seul le réseau routier principal et certaines allées sont représentés. Les voies secondaires n'apparaissent pas ou ponctuellement »* (Robert et Costa, 2009). La raison invoquée par les concepteurs de la carte concerne le *« manque de stabilité de cet élément important qui relie les villages entre eux »*[27] : le réseau principal est en revanche dessiné à l'échelle nationale, avec une densité de routes qui demeure non-négligeable.

Pour autant, dans notre challenge émergent, nous nous emparons de cette donnée, non pas pour construire un réseau de transport nécessaire au calcul de temps de parcours, mais comme une donnée de cadrage sur l'armature principale des routes de l'Ancien Régime, parce que nous posons l'hypothèse que ce réseau conserve une forte inertie, à l'heure de la fièvre routière du début du

[27] La Carte de l'Académie, dite de Cassini. http://cassini.ehess.fr

XIXème siècle (Arbellot, 1973, 1990 ; Reverdy, 1995). Il est en effet très difficile de concevoir un modèle comportemental le long des routes de Cassini, fonction des acteurs de l'époque qui sont très différents, et pour lesquels la notion même de vitesse est totalement différente voire inexistante (Guigueno, 2009). En revanche, elle nous renseigne sur la place de chacune des unités d'observation dans l'architecture des routes de l'Ancien Régime. La vectorisation du réseau de routes issues de la carte de Cassini a fait l'objet d'un projet géomatique collaboratif, fruit de la participation de chercheurs de trois institutions principales que sont l'IGN, l'IFSTAAR et l'EHESS (Perret, Gribaudi et Barthelemy, 2015). Le projet s'appuie sur la copie nommée « Marie-Antoinette », commandée très peu de temps avant la chute de l'Ancien Régime, dont les plans ont été mis à jour par la suite jusqu'en 1815. Une lecture de la *Figure 5. 9* montre une répartition des routes inégale à l'échelle de la France : si la moitié Nord est plus dense que la moitié Sud, alors on voit comment les reliefs des Alpes, des Pyrénées et du Massif Central participent à la difficulté de cheminer sur des pentes importantes. La levée de la carte de Cassini montre des tracés très précis et plus nombreux en région parisienne ainsi qu'en Aquitaine. En revanche, le Centre de la France semble être moins bien doté en routes principales jusqu'au début du XIXème siècle. Gardons à l'esprit que la carte de Cassini ne recense que le réseau principal et que la qualité de la levée est fonction de la participation des élites locales, chargées de contrôler le contenu des cartes, proportionnellement au soutien financier accordé par les pays d'états et les généralités (Pelletier, 2013).

Conscients de l'incomplétude de la levée de Cassini du XVIIIème siècle, nous posons toutefois l'hypothèse qu'elle permet de contrôler l'effet parasite d'un certain nombre d'autres facteurs, susceptibles de participer à l'explication de la croissance démographique. Du réseau levé par les chercheurs parisiens (Perret, Gribaudi et Barthelemy, 2015), nous avons fait le choix de retirer l'ensemble des chemins qui ne sont pas raccordés au réseau principal. De cette sorte, nous ne retenons que l'ensemble connexe le plus fort, qui concentre 92 % des tronçons vectorisés. Désormais, une simple variable binaire permet de savoir si une unité d'observation – une commune – est traversée ou non par une voie répertoriée par la carte de Cassini. Ainsi, près de 60 % des communes françaises sont traversées par une voie dessinée sur la carte de Cassini au tout début du XIXème siècle : nous interprétons cette variable comme celle qui montre dans quelle mesure une commune est accessible par le réseau routier principal hérité de l'Ancien Régime et du tout début du XIXème siècle.

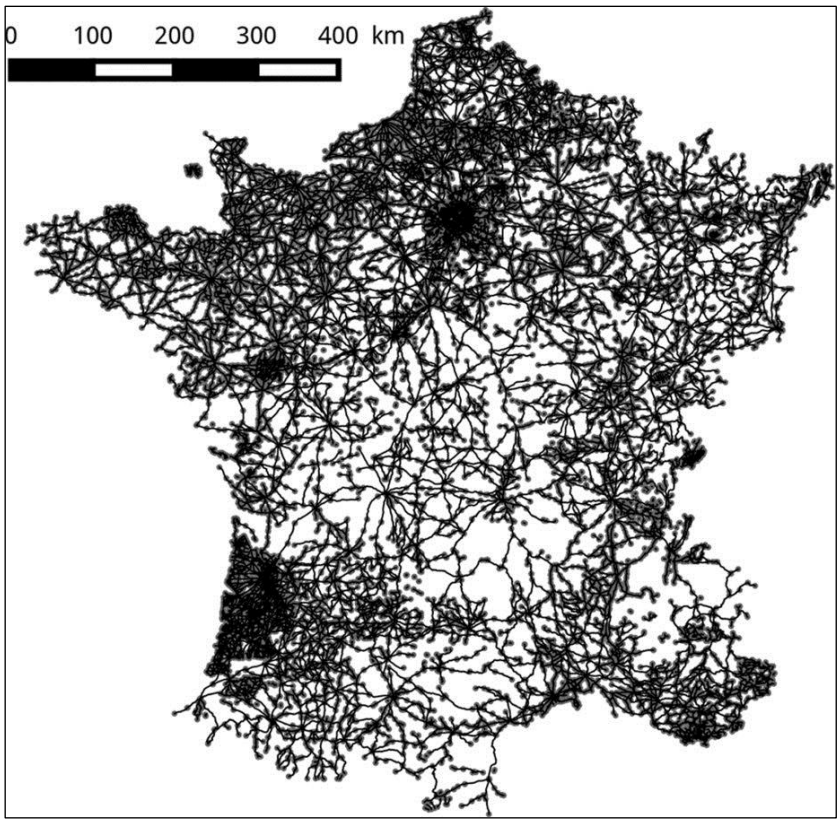


Figure 5. 9. Les voies de communication issues de la Carte de Cassini, d'après PERRET et *al.*, 2015

Par ailleurs, au début du XIXème siècle, *« la navigation joue un rôle économique tout à fait capital »* (Le Sueur, 1989). Elle est notamment centrale à partir de 1820, les bassins fluviaux sont reliés entre eux, par le franchissement d'une ligne de partage des eaux par des canaux : le XIXème siècle est alors celui d'une canalisation généralisée. Deux séries de lois en 1821-1822 et 1836-1837 permettent l'achèvement d'un réseau de plus de 9 000 kilomètres de voies navigables, dont la moitié en canaux : *« certains deviennent des axes importants, comme le canal de Bourgogne, ouvert en 1832, ou celui de la Marne au Rhin, en eau en 1853. D'autres, déjà en dehors des axes nouveaux de trafic, auront une activité plus réduite et souvent plus locale »* (Le Sueur, 1989). Les voies navigables sont souvent considérées comme *« l'épine dorsale du système de transport des économies pré-industrielles »* (Kunz, 1992), alors même que la révolution industrielle amène pour beaucoup *« l'âge d'or de la navigation intérieure »*, même si la concurrence avec les chemins de fer est croissante à mesure qu'avance le siècle. Les impacts d'une telle infrastructure, à l'image du chemin de fer, sont à concevoir à différentes échelles. A cet égard, *« il est difficile de déterminer si la présence ou l'absence du trafic sur les voies navigables a influencé la croissance économique à l'échelle nationale […]. Néanmoins, les canaux pouvaient avoir un impact considérable sur les économies régionales »* (Kunz, 1992). Pour intégrer l'information fluviale dans la base de données FRANcE, nous nous appuyons sur une base de données actuelle, à l'aide de laquelle nous remontons le temps, à l'image de ce qui a été fait pour le réseau ferroviaire français. Il s'agit de la base de données CARTHAGE, qui recense l'ensemble du réseau hydrographique français, enrichie à partir des données IGN, en collaboration avec le ministère chargé de l'environnement ainsi que les agences de l'eau. Après un premier filtre pour ne retenir que les voies navigables, une superposition

avec le réseau hydrographique représenté dans l'ouvrage *« Cinq siècles de transport fluvial en France, du XVIIème au XXIème siècle »* publié en 2003, a permis de remonter le temps, et surtout de dater l'ouverture des canaux entre le XVIIème siècle et aujourd'hui *(Figure 5. 10)*. Dès lors, cette information variante dans le temps avec l'apparition de quelques canaux au cours de notre période d'étude aura sa place dans la spécification en données de panel, au contraire de l'information contenue dans la carte de Cassini.

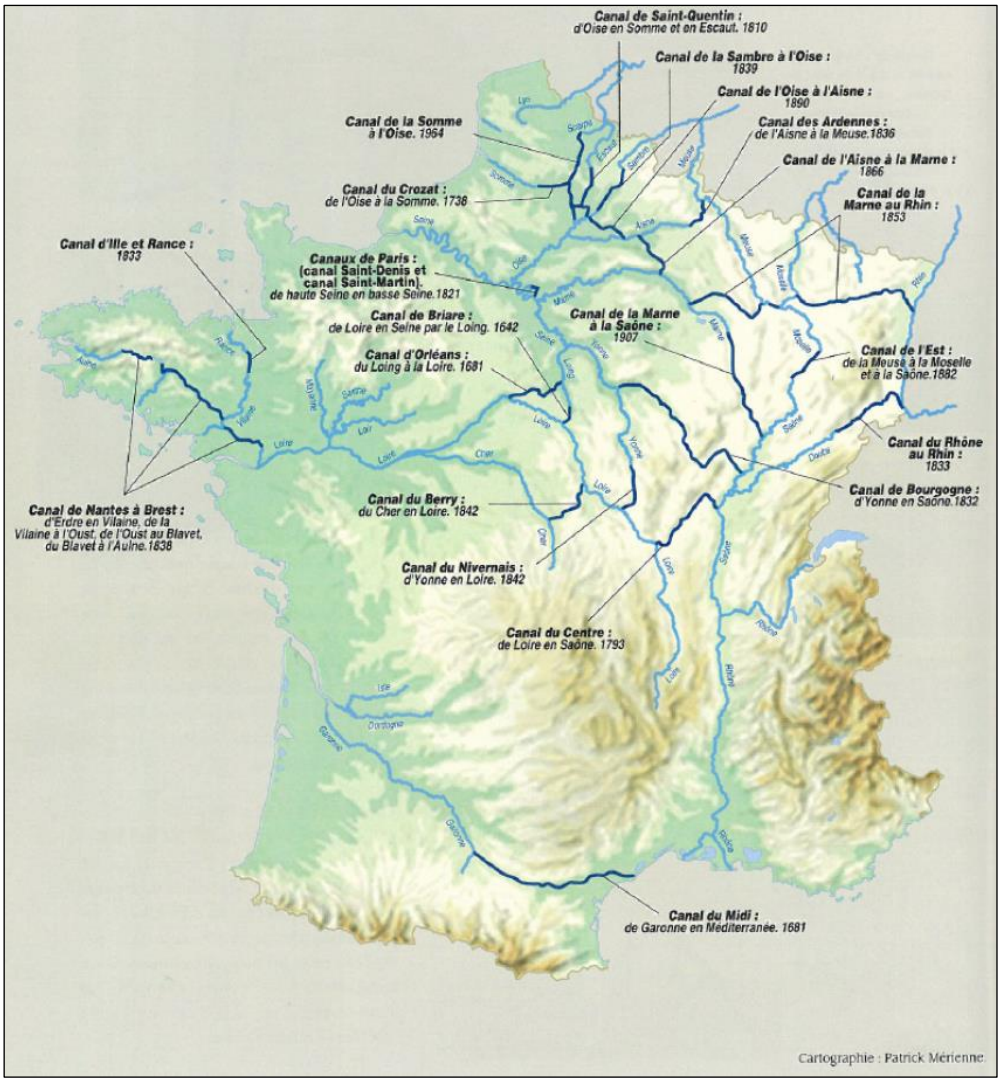


Figure 5. 10. La navigation intérieure depuis le XVIIème siècle, d'après ROBLIN L., 2003. *Cinq siècles de transport fluvial, XVIIème – XXIème siècles*, Editions Ouest-France, 128 p.

Dans la concurrence croissante entre le chemin de fer et la voie d'eau au XIXème siècle, la tarification du transport des marchandises fait l'objet d'une lutte acharnée (Caron, 1997). Alors que les compagnies du Nord et de l'Ouest *« abaissèrent très fortement leurs tarifs de houille et plus généralement de produits pondéreux »* (Caron, 1997), nous posons l'hypothèse que la présence de gisements houillers, qui ont connu un développement important pendant la Révolution Industrielle, peut participer à l'explication de la croissance locale démographique de certaines communes (Rioux, 1989). L'exploitation minière en général, de charbon en particulier, se trouve augmentée de manière exponentielle, parce que les machines en vapeur, surtout les locomotives, sont de grands

consommateurs. Ainsi, cette variable vient atténuer la possible surestimation de la dotation en réseau alors même que la présence d'un gisement peut influer sur la venue d'ouvriers dans la région, de même que sur l'arrivée du chemin de fer pour le transport de la houille. Pour cela, nous nous appuyons sur la base de données SIG Mines France[28], développée par le Bureau de Recherches Géologiques et Minières (BRGM). Ce projet participe à la cohésion de données géoréférencées au sein du BRGM, à destination d'une communauté scientifique large, ainsi qu'aux services en charge de la gestion des ressources minérales et de la prévention des risques. Alors, cette base de données fait aussi bien appel à la base des ressources minérales qu'aux anciennes concessions[29]. Entre 1860 et 1930, la base recense ainsi près de 700 gisements sur le territoire national, à partir desquels on peut distinguer la substance extraite ainsi que les dates d'exploitation. Pour autant, nous devons faire face à certaines données incomplètes qui viennent atténuer l'exhaustivité d'une telle entreprise : nous nous retrouvons dans les défis classiques incombés à la fois aux projets de SIG-H et de la modélisation de panel (Dormont, 1989 ; Knowles, 2005). Mais, la *Figure 5. 11* montre tout de même la répartition des gisements miniers qui intéresse notre période d'étude : cette information invariante dans le temps ne pourra être intégrée qu'aux premières étapes de la spécification du modèle.

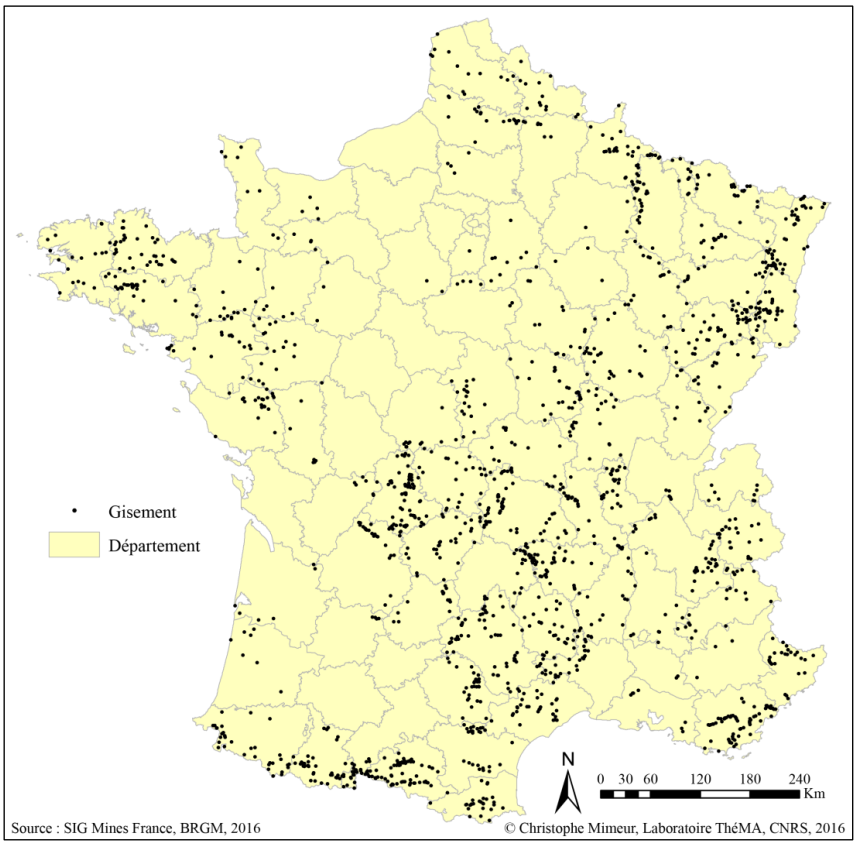


Figure 5. 11. Les gisements houillers sur le territoire français

[28] http://sigminesfrance.brgm.fr

[29] Si les données sont en libre accès sur le site du projet, l'extraction des dates d'exploitation des gisements a fait l'objet d'une collaboration avec les membres du BRGM, afin d'obtenir un fichier unifié. Les données reçues ne permettent pas en revanche d'obtenir une information variante dans le temps. Elles permettront d'explorer le modèle analytique sans pouvoir être intégrées à la spécification en panel.

Cette base permet non seulement de prendre en compte les gisements énergétiques, mais toutes les autres, susceptibles d'influencer la croissance démographique locale. Dans le cadre de construction de connaissances géohistoriques, on retrouve ainsi des zones de concentration de gisements à l'Est de la France, en Lorraine particulièrement. Des gisements importants ressortent également dans le centre de la France, ainsi qu'en Bretagne, les Alpes et les Pyrénées.

Par ailleurs, nous pouvons facilement dériver du MNT *BD_ALTI* de l'IGN des informations quant au site de chaque commune française, pouvant influencer sur sa croissance démographique. Nous retenons dans la base de données FRANcE des informations constantes dans le temps, telles que la surface, l'altitude ou le dénivelé de chaque unité d'observation (Baum-Snow, 2007 ; Baum-Snow et al., 2012). Ainsi, jusqu'à maintenant, les variables retenues pour limiter le biais d'une surestimation de la variable d'intérêt se cantonnent largement à des variables constantes dans le temps. Pourtant, là où on place la dimension temporelle au cœur de la modélisation en panel, nous cherchons désormais à insérer dans le modèle des variables variantes dans le temps et dans l'espace. L'accès aux voies navigables y participe déjà lors de la construction des canaux. Pour autant, nous cherchons à diversifier ces variables dans la démarche constructiviste de notre travail. Alors nous posons les questions de l'échelle des données et de leur insertion dans la base de données FRANcE. A l'instar des travaux déjà cités dans la section précédente (Atack et al., 2010 ; Hornung, 2012a), nous souhaitons prendre en compte des variables décrivant la structure économique, parce qu'elle est aussi un facteur local de la croissance démographique (Chi, Voss et Deller, 2006). Dans la perspective d'une modélisation en panel, qui identifie des processus temporels, nous devons envisager l'insertion de données variantes dans le temps, dans les critères de précision et de pertinence que nous nous imposons.

Si des estimations précises requièrent une variable locale propre à chaque individu, le contexte géohistorique nous suggère une approche adaptée qui combine alors un état des lieux à la fois local et environnant. Ainsi, ces nouvelles variables couvrent les larges échelles spatiales et temporelles. Pour cela, nous nous appuyons sur la richesse des données collectées par Jean-Claude Toutain, qui visent à donner de nouvelles clés pour comprendre les disparités régionales en France (Combes et al., 2011). Les données collectées par Jean-Claude Toutain le sont au niveau départemental[30]. D'après ces auteurs, ne travailler qu'avec trois secteurs d'activités n'est pas restrictif dans la mesure où, à ces larges échelles, les bouleversements socio-économiques de notre période d'étude sont sensibles à des niveaux agrégés, que ce soit la montée du secteur industriel, la dépendance croissante du secteur agricole aux autres secteurs, de l'abandon de la polyculture. Ainsi, les données dont nous disposons concernent trois dates pour notre période d'étude : 1860, 1896 et 1930. Elles décrivent le volume d'emplois pour les trois secteurs, ainsi que la valeur ajoutée. Pour autant, cette dernière variable n'est pas disponible pour l'année 1896, ce qui nous conduit à privilégier l'utilisation de la variable *« emploi »*.

Deux difficultés se posent toutefois pour l'insertion de la donnée dans la base de données FRANcE. La première réside dans l'évolution des découpages départementaux sur le temps long : si le département est un héritage de la Révolution Française, le découpage a évolué, influencé par

[30] Pour le détail des sources utilisées dans la collecte de ces données, nous nous référons à l'annexe disponible dans l'article de Combes et al., 2011. La diversité et l'hétérogénéité des sources placent cette entreprise dans les obstacles inhérents à la cliométrie et à la géohistoire.

les conflits géopolitiques. De cette sorte, nous avons reconstitué les départements de 1860, 1896 et 1930, afin de pouvoir affecter la variable économique à chaque commune. Nous nous appuyons sur la méthode adoptée par Pierre-Philippe Combes et ses coauteurs : les changements les plus notables sont à chercher en Île-de-France, où il n'existait que les départements de la Seine et Seine-et-Oise ; ainsi que les départements à l'est de la France. Le Haut-Rhin et le territoire de Belfort sont unifiés, alors que les départements Meurthe et Moselle le sont aussi pour 1860, Meurthe-et-Moselle et Moselle pour les années ultérieures. La seconde difficulté réside dans l'utilisation de la variable dans la modélisation en panel : nous ne disposons en effet que de trois dates alors que notre approche considère chaque décennie. Dès lors, nous conservons les données de 1860 jusqu'à la décennie 1890, puis les données de 1896 viennent enrichir les prochaines décennies.

Aussi, considérée de manière brute, la variable d'emplois est intrinsèquement liée aux niveaux de population des départements. Pour ne pas biaiser les estimations dans la modélisation de panel, nous misons sur la construction de deux indicateurs, permettant de qualifier la spécialisation ainsi que la diversité sectorielle dans la structure économique départementale (Combes et al., 2011). Pour chaque département *d* et secteur d'activité *s*, on définit donc :

$$SPE_{d,s} = \frac{Emp_{d,s}}{\sum Emp_s} \tag{23}$$

$$DIV_d = \frac{1}{\sum_s (SPE_{d,s})^2} \tag{24}$$

De cette sorte, les variables visant à décrire la structure économique du pays dans notre période d'étude participent à une meilleure estimation de notre modélisation d'une part, en répondant aux critères de pertinence et de précision d'autre part. Les résultats révélés par les auteurs montrent une variation semblable entre la croissance de la population et la croissance des emplois entre 1860 et 1930 (Combes et al., 2011). Par ailleurs, en analysant les inégalités à l'échelle des départements et des régions, ils montrent comment la croissance des inégalités est plus grande entre les régions qu'à l'intérieur d'une même région.

Ainsi, cette sous-section a permis d'introduire de nouvelles variables dans la base de données FRANcE, afin de limiter le biais de variables omises, et donc ne pas surestimer l'impact du réseau sur la croissance de la population. S'il est plus aisé de compléter la base avec des données environnantes et invariantes dans le temps, l'ajout de variables variantes dans le temps répond aux enjeux de la modélisation de panel, en nous renvoyant aux défis de l'information géohistorique. Pour autant, nous suspectons toujours à ce stade un biais d'endogénéité inhérent à la circuité des relations entre évolution territoriale et évolution du réseau. Dans la perspective géohistorique que nous revendiquons dans notre travail, nous misons sur les potentiels enseignements à retirer des réseaux passés pour améliorer les estimations de notre modélisation, à l'heure où la méthodologie statistique semble se sophistiquer par le recours à l'instrumentation (Duranton et Turner, 2012 ; Hornung, 2012b).

5.3.2. **Les « voies » de l'instrumentation : principes et stratégies**

Dans leur revue de littérature, Dena Kasraian et ses co-auteurs identifient non seulement les avancées des SIG comme un potentiel pour étudier les relations de long terme entre infrastructure de transport et leur impact sur le territoire (Kasraian et al., 2016), mais aussi la diversification des méthodes et des approches pour estimer, sinon montrer, la force de ces relations. S'ils n'évoquent pas dans leur liste la modélisation en panel que nous envisageons dans notre travail, ils listent parmi les méthodes utilisées la modélisation par les variables instrumentales, afin de limiter le biais d'endogénéité, inhérent à la modélisation de relations entre phénomènes intimement liés.

Quand nous replaçons cette réflexion dans notre question de recherche, notre revue de la littérature dans la première partie de la thèse nous laisse penser que la croissance démographique locale d'une part, et le développement du réseau ferroviaire d'autre part, participent à des phénomènes de détermination simultanée, que la littérature nomme *« congruence »* (Offner, 1993b) mais aussi *« coévolution »* (Bretagnolle, 2014). C'est dans ce cadre que nous développons une réflexion sur la sophistication de la modélisation par le recours à l'instrumentation : après avoir décrit les principes statistiques de cette méthode, largement empruntée à la démarche économétrique, nous définissons une stratégie instrumentale dans les humanités numériques, qui suppose alors un nouveau recours aux critères de pertinence et de précision dans un raisonnement géohistorique, quand on envisage une expansion du temps long.

Le recours aux variables instrumentales est souvent présenté comme un moyen d'éliminer le biais des estimateurs, celui de l'endogénéité. Cette technique rencontre un vif succès dans l'étude des relations entre transport et territoire, à travers des approches diversifiées (Atack et al., 2010 ; Duranton et Turner, 2012 ; Hornung, 2012b ; Kasraian et al., 2016). Le recours à l'instrumentation est alors commun dans la modélisation économétrique mais relativement sous-exploité dans une perspective historique (Atack et al., 2010).

Rappelons dans un premier temps quelques grands principes qui permettent une meilleure convergence des estimations dans un modèle explicatif. Nous retenons un exemple souvent cité par les économistes pour illustrer l'apport d'une telle méthode (Arcand et Labart, 2009). Dans les études portant sur les relations entre niveau d'éducation et niveau de salaire, la littérature suspecte fréquemment une endogénéité, que l'on pourrait attribuer à l'omission d'une variable mesurant le talent d'un individu. Parce que cette variable est difficilement mesurable, il faut imaginer le recours à des variables qui sont sensiblement corrélées avec la variable explicative, mais suffisamment indépendante de la variable à expliquer : dans cet exemple, on pose alors l'hypothèse que le niveau d'éducation des parents est corrélé avec l'éducation de l'individu, mais indépendant de son salaire. Nous retenons alors du principe de l'instrumentation la substitution de la variable suspectée d'endogénéité par une autre, dont on pose l'hypothèse qu'elle est indépendante avec la variable à expliquer. Autrement dit, dans notre modèle de relation entre population et territoire, il s'agit d'adopter une stratégie permettant d'identifier des variables fortement corrélées avec la dotation en réseau mais suffisamment indépendante avec la croissance démographique locale.

Parallèlement aux exigences méthodologiques et aux spécifications statistiques, notre

réflexion géohistorique nous amène à une double stratégie dans le choix des variables instrumentales. De plus, si l'utilisation de variables instrumentales dans les modèles de régression linéaire ne suppose que les critères précédemment cités, alors la modélisation de panel nécessite d'avoir une réflexivité temporelle, parce que la variable utilisée doit pouvoir traduire les trois types de variabilité des estimations sur des données de panel.

Dans le premier temps de la stratégie, on pose donc l'hypothèse que la *« fourniture de nouvelles infrastructures de transport […] dépend de la croissance de la population actuelle et prévue »* (Duranton et Turner, 2012) : là est identifiée l'endogénéité du modèle. Dans leur approche appliquée à la croissance urbaine aux Etats-Unis face à leur dotation en infrastructures autoroutières, ces chercheurs suggèrent l'utilisation d'un instrument qui reflète le *« niveau des infrastructures de transport à un moment donné, dans la passé »* ainsi que la *« situation géographique favorable à la construction d'infrastructures »* (Duranton et Turner, 2012). Dans leur modèle étudiant la période 1980-2000, ils utilisent ainsi le réseau d'autoroutes américaines de 1947 et le réseau ferroviaire tel qu'il était déployé en 1898. Dans le second cas, d'autres auteurs envisagent la construction de variables instrumentales postulant sur la situation favorable à la construction d'infrastructures, parce qu'elles conditionnent souvent par ailleurs la densité, l'urbanisation, le niveau de population (Atack et al., 2010). Cette remarque nous apparaît d'autant plus consistante que nous avons identifié ces aspects comme des variables de contrôle de la croissance démographique locale. Nous n'envisageons donc pas ici de les considérer comme des éléments de notre stratégie d'instrumentation.

En revanche, dans une perspective historique, le recours aux variables instrumentales se place dans la veine des études ne prenant en compte qu'une variable binaire d'accès au réseau, Erik Hornung et Jeremy Atack construisent une variable instrumentale consistant à traduire la présence d'une commune sur un itinéraire direct entre deux grandes villes, identifiant des grandes lignes (Atack et al., 2010 ; Hornung, 2012b). La construction de telles lignes droites traduit un potentiel de passage d'une voie ferrée, pour instrumenter l'accès au réseau : le postulat des chercheurs est que les projets de construction s'attachent à privilégier la ligne droite, sans quoi la déviation d'une ligne engendre de forts coûts de construction. Dès lors, on postule que les villes connectées en dehors de ces lignes droites le sont pour des raisons endogènes. Si ces études ont une portée sur le long terme, elles n'en restent pas moins micro-régionales.

Par ailleurs, la modélisation des données de panel nous pousse plutôt à privilégier l'approche stratégique de Duranton et Turner à partir des réseaux historiques (Duranton et Turner, 2012), parce que l'instrumentation fait souvent appel à *« l'utilisation de variables retardées »* (Le Gallo, 2002). Si l'une des principales difficultés réside dans la rareté de telles variables, d'autres évoquent leur déficit de qualité, tant elles peuvent parfois avoir des impacts sur de longs délais, et donc ne pas respecter les principes économétriques de l'instrumentation. Pour autant, l'approche géohistorique dans les humanités numériques nous poussent à la création de variables instrumentales s'appuyant sur des réseaux historiques : dans ce cas, il ne s'agit pas de prendre en compte l'état de la dotation en réseau *« l'année précédente, la décennie précédente »*, en postulant que le niveau d'infrastructure d'il y a vingt ans n'a pas d'impact sur la croissance démographique locale que l'on cherche à analyser. L'objectif est de mobiliser la longue durée, ou ce que certains appellent même la *« très longue durée »* (Grataloup, 2015) : on pose alors la question dans la section

suivante du recours à des réseaux historiques ainsi que de leurs conditions d'implémentation dans la base de données FRANcE, pour en faire des instruments de la dotation contemporaine en réseau face à la croissance démographique locale des communes françaises.

5.3.3. **Route et réseaux historiques pour instrumenter FRANcE**

Dans une perspective géohistorique et pour la construction d'instruments de la dotation en réseau ferroviaire depuis le XIXème siècle, nous misons ici sur la création de variables décrivant le *« niveau des infrastructures de transport à un moment donné, dans la passé »* (Duranton et Turner, 2012). Dans une modélisation de données en panel, la prise en compte de la variabilité temporelle nous pousse à avoir recours à au moins deux points dans le temps long. L'objectif est alors triple :

- Avoir recours à la structure de réseaux historiques, tout en questionnant le concept de réseau en fonction des contextes ;
- Questionner les modèles comportementaux pour envisager l'accessibilité par ces réseaux ;
- Evaluer l'adaptabilité de la démarche exposée dans le chapitre précédent pour aborder la question de l'accessibilité sur le temps long par les graphes.

Envisager les réseaux historiques implique un recours à l'histoire des routes de France, de leur définition jusqu'à leur usage sur la longue durée. Nous nous appuyons pour cela sur les écrits du spécialiste français Georges Reverdy, qui a étudié dans un manuel les routes depuis l'Antiquité jusqu'à nos jours (Reverdy, 1995). Alors que la question des routes a souvent été étudiée de manière monographique par l'historiographie française, les essais d'histoire générale à de larges échelles sont plutôt rares (Reverdy, 1995). C'est un géographe, Roger Dion, qui donna une introduction à la géographie de la circulation en France. Si, à l'image de notre postulat pour la construction du multigraphe F, l'individu a toujours cherché à cheminer, nous nous intéressons dans notre stratégie instrumentale aux *« voies construites, plus élaborées que les simples pistes résultant du piétinement des passants »* : la voie romaine est donc *« une route construite du temps des Romains, avec ses ouvrages […], sa plate-forme et sa chaussée […] et son tracé […], résultant de l'intervention des géomètres et des arpenteurs »* (Reverdy, 1995).

Dans les critères de pertinence et de précision d'une part, dans ceux de la construction d'indicateurs d'accessibilité d'autre part, nous voyons dans les voies romaines des infrastructures le long desquelles *« on effectue à pied de forts longs parcours »* (Chevallier, 1972)[31], même si l'auteur mentionne également l'existence de quelques autres véhicules qui servent à transporter les messages. La vitesse des déplacements n'est que rarement évoquée dans les sources anciennes : quelques comparaisons entre les écrits d'Hérodote ou de Strabon font osciller les vitesses entre 3 km/h et 5km/h, avec parfois quelques contradictions dans les écrits de l'époque. Pour autant, le

[31] L'auteur cite ici l'auteur Tacite dans *Annales*, III, 9.

manque d'informations exhaustives à nos larges échelles nous pousse à préférer un modèle comportemental exclusivement piéton. Ainsi, nous restons heurtés aux difficultés d'incomplétude et d'incertitude des données géohistoriques. Dans le cas des voies romaines, ce sont des difficultés tant archéologiques qu'épigraphiques (Chevallier, 1972). Sans entamer un nouveau processus de collecte, nous saisissons en revanche une source fondamentale, et reconnue comme telle en 2000[32], le *Barrington Atlas of the Greek and Roman World*. C'est le résultat d'un projet de 12 ans qui a fédéré plus de 200 chercheurs autour de Richard Talbert, à de larges échelles spatiales – l'ensemble du monde romain – et temporelles – entre 550 av. JC et 640 ap. JC. Malgré d'autres essais à des échelles plus réduites, cet atlas est finalement une première à de telles échelles, alors que pendant longtemps les études antiques ont délaissé l'espace géographique, et surtout sa cartographie (2000). En dépit des prudences historiographiques, nous rappelons ici le caractère exploratoire de l'instrumentation par les réseaux sur la très longue durée, en parallèle de notre positionnement pour la capitalisation de l'information géohistorique. Dans les humanités numériques, nous nous appuyons sur le *Digital Atlas of Roman and Medieval Civilizations*. Sous l'égide de l'historien Michael McCormick et du géographe Giovanni Zambotti, cette équipe de l'Université de Harvard a compilé un certain nombre de sources historiques, dont le *Barrington Atlas*, dans un SIG-H *(Figure 5. 12)* disponible en libre accès[33].

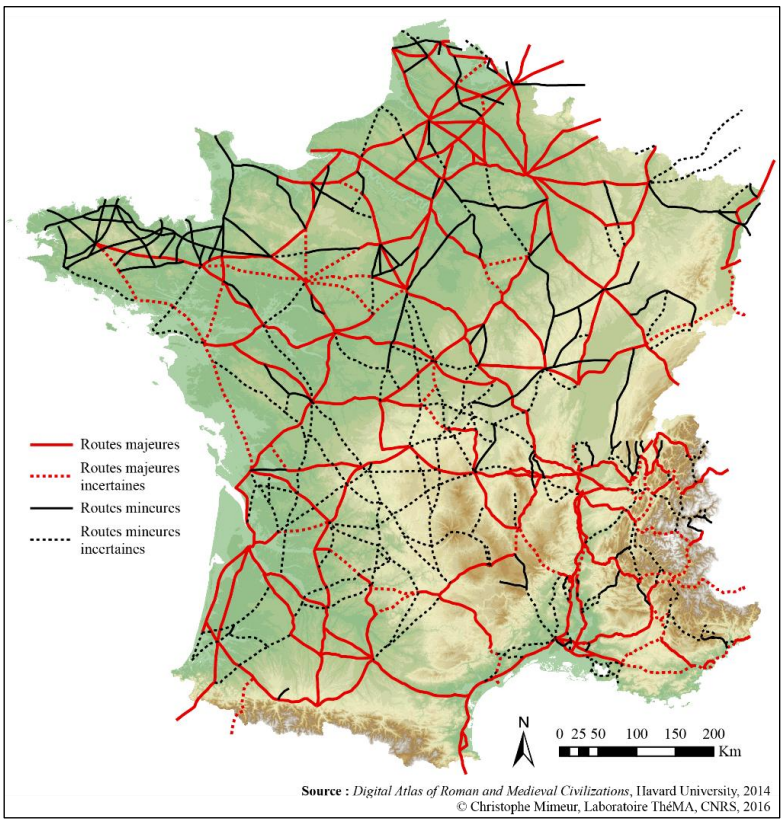


Figure 5. 12. Les principales voies romaines entre 550 av. JC et 640 ap. JC

La structure de voies romaines montre bien des logiques d'axes, centrées sur Lyon, par les

[32] Cet ouvrage a reçu le titre de *Best Professional/Scholarly Multivolume Reference Work in the Humanities*.
[33] http://darmc.harvard.edu

deux rives du Rhône, la Saône, la Moselle, le Rhin. On observe de larges transversales au sud de la France ainsi qu'au Centre, alors que les plus fortes densités de routes majeures, dont le caractère est certain qui plus est, sont dans la partie Nord du pays, même si Paris ne semble pas être un carrefour à la mesure de l'Etoile de Legrand sur le réseau ferroviaire français. Au Nord, les régions de Langres ou Reims sont des carrefours plus importants, là où le Nord-Ouest semble rester à l'écart des routes majeures, mais où les routes mineures viennent densifier le réseau en Bretagne. Beaucoup d'historiens identifient alors une priorité militaire dans l'établissement des voies, car elles *« devaient guider les invasions. Elles sont jalonnées de trésors monétaires. Les campagnes de l'empereur Julien s'expliquent en fonction du réseau routier »* (Chevallier, 1972). Les réseaux les plus incertains sont surtout ceux qui viennent densifier le réseau au centre de la France, mais qui ne remettent pas en cause la logique d'axes que nous venons de décrire.

L'utilisation de cette donnée pour l'évaluation de temps de parcours dans notre stratégie instrumentale suppose un certain nombre d'adaptations pour reproduire la méthodologie développée dans le chapitre précédent. Sans réitérer la succession des traitements géomatiques, nous revenons sur les principales étapes de la modélisation. A l'inverse d'un réseau ferroviaire discontinu, nous postulons que l'entrée sur le réseau de voies romaines est continue : en langage de graphes, nous devons donc théoriquement multiplier les nœuds du graphe à l'infini. Pour autant, si chaque commune est espacée de 2,3 kilomètres en moyenne d'une autre sur le territoire français, nous choisissons de segmenter chaque tronçon de voies romaines tous les 2 kilomètres. Par la suite, nous calculons l'impédance, la durée nécessaire pour traverser le tronçon, à partir de la typologie de routes disponible dans le *Barrington Atlas* et de la grille que nous avons adaptée, que nous estimons être capable de prendre en compte des différentiels entre les classes et les cheminements en-dehors des voies, davantage que des temps de parcours réels *(Table 5. 3)*. Ainsi, sur les quelques 24 000 kilomètres de voies romaines répertoriées par les chercheurs, près de 60 % sont des voies qualifiées de certaines. Les autres routes sont incertaines : nous faisons ainsi l'hypothèse que leur praticabilité était sans doute légèrement inférieure à celles dont les traces sont les plus évidentes. Pour autant, dans le *Barrington Atlas*, les routes les plus incertaines ne sont renseignées que par des directions, et ne sont pas repris dans le SIG-H conçu par l'équipe d'Harvard, ce qui limite considérablement le risque d'erreur historique. Le croisement systématique au MNT suit le même processus que pour le graphe triangulé des communes. Désormais, on a donc un graphe R, décrivant les cheminements sur les voies romaines, qui peut être juxtaposé à la triangulation par le graphe P traduisant les cheminements hors des voies *(Figure 5. 13)*, pour obtenir le multigraphe ROM.

		Kilomètres		Pondération de vitesse
Routes majeures	Certaines	11382	*41,80%*	1
	Incertaines	2936	*10,78%*	1,5
Routes mineures	Certaines	5291	*19,43%*	2,5
	Incertaines	7621	*27,99%*	3
Triangulation				5

Source : *Digital Atlas of Roman and Medieval Civilizations*, Havard University, 2014
© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 5. 3. Répartition des voies romaines dans la France actuelle

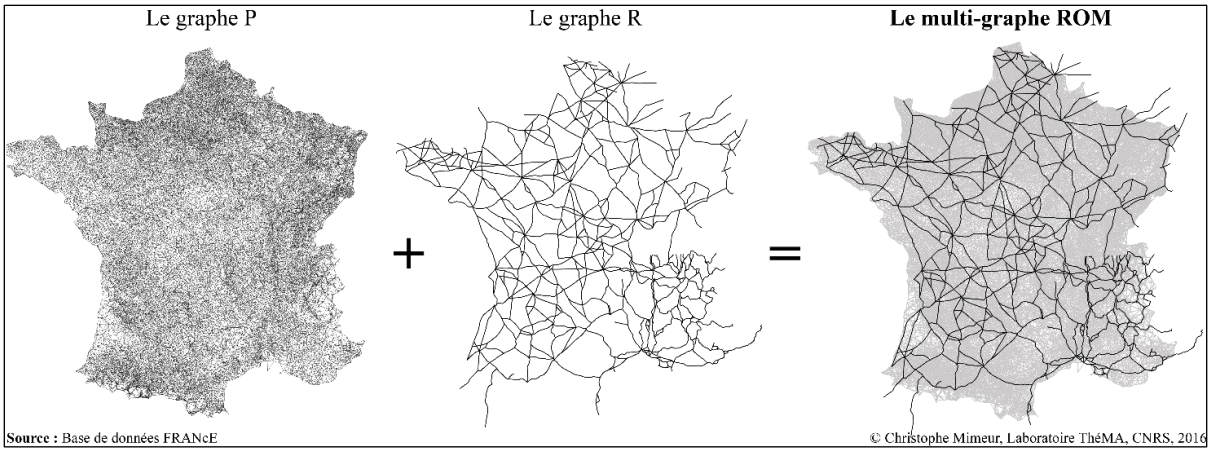


Figure 5. 13. Le multigraphe ROM

A partir du multigraphe ROM, nous sommes désormais en capacité d'analyser dans la prochaine partie de la thèse comment la prise en compte des voies romaines participe à l'instrumentation de la dotation en réseau ferroviaire au XIXème siècle. Toutefois, dans une modélisation en panel, nous avons besoin d'un second point, à minima, pour tirer profit des variabilités qu'il est alors possible de modéliser. La question du devenir des voies romaines a souvent été posée : en dépit d'une forte inertie dans la structure des réseaux, beaucoup d'historiens s'accordent à dire que *« les anciennes voies sont devenues rapidement impraticables »* (Reverdy, 1995). Ainsi, pendant quelques siècles, *« les pélerinages, et en premier lieu celui de Saint-Jacques-de-Compostelle, ont constitué […] un réseau routier spécifique »* (Reverdy, 1995), alors que ces chemins ont disparu à la fin du XVIIème siècle. L'histoire médiévale semble alors laisser une large place aux réseaux urbains et aux pavages des rues, sans que des études générales évoquent les grands itinéraires entre les foyers de commerce et de foire dans la France médiévale, en partie en raison d'une relative absence d'un pouvoir central, susceptible de recenser de telles initiatives. Il n'en reste d'ailleurs souvent que des ponts de pierre, alors que les multiples chemins de l'époque ont souvent été *« complètement transformés au cours des phases ultérieures de l'extension urbaine »* (Reverdy, 1995). Notre stratégie nous pousse alors à chercher un réseau plus loin dans le temps.

Les historiens tendent à dater la création d'un réseau de poste *« qui allaient constituer l'armature principale du réseau jusqu'au milieu du XIXème siècle »* (Reverdy, 1995) en 1480. Si l'obtention de tels réseaux nécessite le recours à de nouvelles sources[34] pour établir un réseau au XVIème siècle, nous nous rapprochons des travaux initiés par Anne Bretagnolle et Nicolas Verdier quant à la constitution d'un corpus des routes de postes entre le XVIIème et le XVIIIème siècle (Bretagnolle, 2009 ; Verdier et Bretagnolle, 2007). Le réseau de la poste aux chevaux semble alors jouer un rôle fondamental dans les modes de déplacement de l'époque (Bretagnolle, Giraud et Verdier, 2010), parce qu'il constitue le premier véritable réseau de transport public à l'échelle du territoire national (Arbellot, 1973). Ce réseau s'est particulièrement densifié entre 1708 et 1833, passant de 10 420 à 27 822 kilomètres (Bretagnolle, Giraud et Verdier, 2010). Alors que l'on parle

[34] On ne citera ici que *La guide des chemins de France*, publié à Paris chez Charles Estienne en 1552.

plus volontiers d'un maillage de voies romaines, on peut véritablement parler d'un réseau de poste à cheval, même s'il n'est pas pensé comme cela au départ (Lepetit, 1986), dans la mesure où *« il superpose deux réalités coexistantes : d'une part, un système de relais, tenus par des maîtres de poste possédant un certain nombre de chevaux mis à disposition de l'institution postale ; d'autre part, une série de tronçons routiers, qui dépendent rapidement des Ponts et Chaussées »* (Bretagnolle et Verdier, 2005). Anne Bretagnolle et Nicolas Verdier ont mis à notre disposition le réseau postal de 1833 *(Figure 5. 14)*, date à laquelle il est en réalité le plus dense depuis le XVIIIème siècle : le long de ce réseau, la vitesse moyenne des déplacements est estimée à 15 kilomètres par heure, en dépit de sources généralisées et harmonisées (Bretagnolle, 2003), dans de solides voitures à quatre roues, souvent bâchées de cuir, *« avec trois places dans le coupé plus un siège dans le cabriolet à côté du cocher, emmené par quatre ou cinq chevaux de trait léger »* (Studeny, 2009).

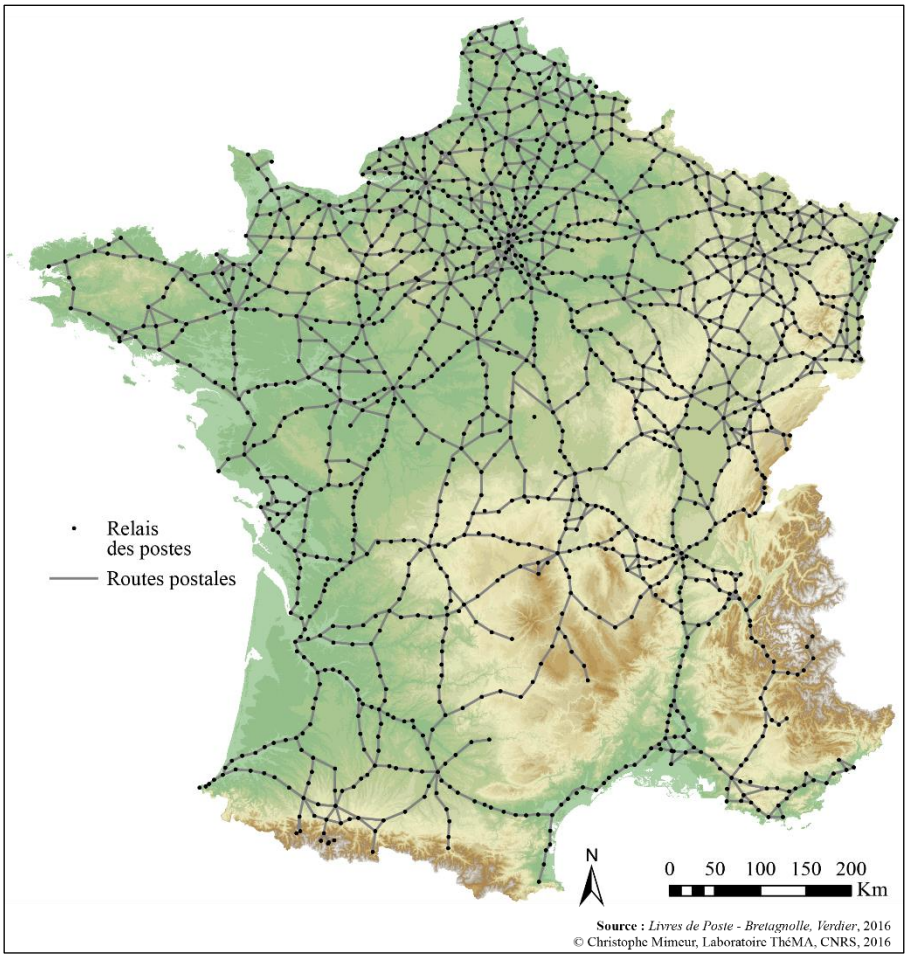


Figure 5. 14. Le réseau postal en 1833

La partie septentrionale de la France bénéficie d'une plus grande densité de routes que la partie méridionale. A l'instar des voies romaines, nous retrouvons ici une logique d'axes, parmi lesquels la vallée du Rhône et de la Saône, ainsi que la transversale entre l'Aquitaine et le Languedoc. En revanche, la centralité du réseau postal semble bien s'identifier à Paris : au fur et à mesure du déploiement de ce réseau, le centre de gravité du réseau s'est largement rapproché de Paris, même si les dernières extensions du réseau ont visé à atténuer le déséquilibrage entre le Nord et le Sud de la France (Bretagnolle et Verdier, 2005). Dans notre stratégie instrumentale

exploratoire, il faut toutefois noter ici que ce réseau ne fournit que des vitesses homogènes le long des routes de postes. Pour autant, nous misons sur ce réseau pour apprécier les différentiels d'accessibilité.

Alors que les chercheurs, pour modéliser l'accessibilité, se sont basés sur des rapports du simple ou double pour distinguer les routes de poste des autres cheminements, nous nous contentons d'adapter notre méthode développée dans le chapitre précédent, qui permet de reprendre leur hypothèse selon laquelle *« les segments de rabattement […] imaginés correspondent en quelque sorte au lacis dense des simples chemins de terre »* (Bretagnolle, Giraud et Verdier, 2010) appréhendé par la triangulation dans notre travail. De la même manière que pour le réseau ferroviaire, le réseau de poste à cheval est discontinu dans la mesure où on ne peut se déplacer qu'à ces vitesses supérieures seulement à partir du moment où l'on se rend dans un relais pour emprunter un véhicule adéquat. Pour sophistiquer l'approche des auteurs, nous choisissons de croiser les tronçons des routes de poste avec le MNT, pour tenir compte des dénivellations, dont les chevaux sont tout aussi sensibles que la force humaine. Les traitements géomatiques sont dans ce cas tout à fait analogues à ceux effectués pour la construction du multigraphe F : quand on juxtapose le graphe P issu de la triangulation entre les communes françaises, au graphe C des relais de postes à cheval de 1833, on obtient le multigraphe POS *(Figure 5. 15)*.

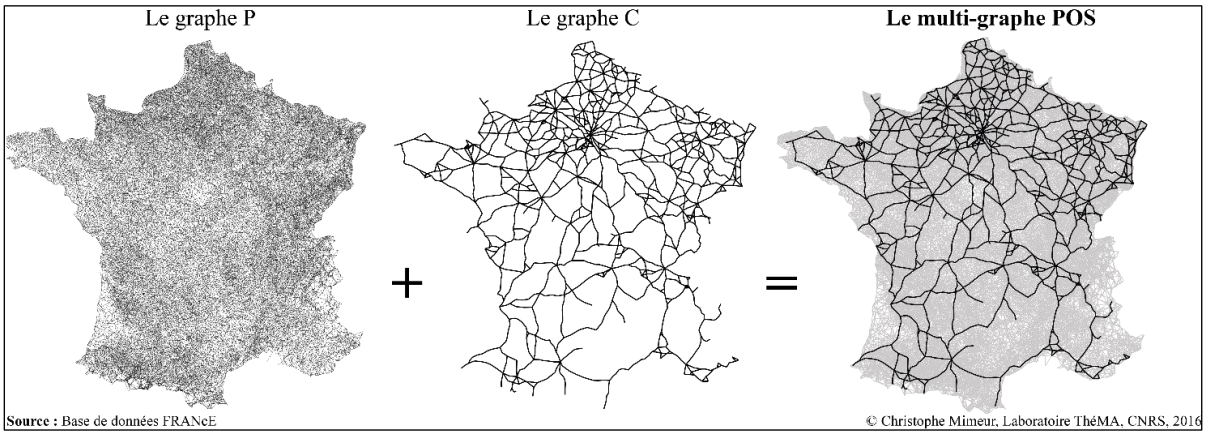


Figure 5. 15. Le multigraphe POS

Ainsi, à partir du maillage des voies romaines d'une part, du réseau postal de 1833 d'autre part, nous avons désormais deux points dans le temps pour instrumenter la dotation en réseau ferroviaire. De la même manière que pour le multigraphe F, on peut adapter les indicateurs d'accessibilité sans difficulté à partir du formalisme des graphes *(cf. 5.1)*. Ces deux points permettent de prendre en compte la variabilité dévolue aux données de panel. Ils participent à la prise en compte d'un biais d'endogénéité lié à la concomitance de deux phénomènes intimement liés.

Conclusion

Ce cinquième chapitre visait à construire un modèle analytique pour l'étude des relations entre croissance démographique locale et dotation en réseau. Sur le long terme, de nombreux travaux empiriques se cantonnaient aux dépenses d'investissement public à un niveau macro-économique, à une description simple de la dotation en réseau à un niveau micro-économique, par la vérification de grandes théories économiques (Kasraian et al., 2016). Notre modèle analytique poursuit trois autres objectifs : d'abord complexifier l'approche du réseau par l'accessibilité géohistorique, puis valoriser et capitaliser l'information géohistorique, en questionnant l'espace et les échelles d'une part, en saisissant la dimension temporelle de ces informations pour affiner les relations dans le modèle d'autre part. Ainsi, des méthodes statistiques les plus classiques jusqu'à des modèles spatiaux raffinés, nous misons dans le cadre de cette thèse sur la prise en compte de la variabilité temporelle dans la formulation du modèle par une structuration en données de panel, parce que nous donnons parallèlement une forte empreinte spatiale aux variables à introduire dans le modèle analytique.

Cette conclusion permet de définir les différents degrés de mobilisation de la base de données qui seront entrepris dans la troisième partie de la thèse pour étudier l'effet du réseau sur les dynamiques démographiques. La *Figure 5. 16* permet ainsi de préciser le statut que nous donnons à chacune des familles de variables que nous avons construites et collectées tout au long de ce chapitre. Il s'agit d'abord d'étudier le potentiel descriptif et explicatif de la complexification de l'accessibilité géohistorique dans la phase de croissance du réseau ferroviaire français *(Figure 5. 16. b.)*. Par la suite, il s'agit d'envisager les relations qu'entretiennent cette dotation en réseau et le système de peuplement *(Figure 5. 16. a. et b.)*. Enfin, il s'agit bien de complexifier l'approche par une modélisation économétrique des effets du réseau, en misant sur la diversification de l'information géohistorique *(Figure 5. 16. c.)*. L'utilisation de variables de contrôle participe alors à la prise en compte de situations locales pouvant contribuer à l'explication d'une situation démographique locale, et vient donc minimiser le biais de variables omises, déjà atténué par la dimension en panel. Par ailleurs, une seconde stratégie consiste à introduire des variables instrumentales pour limiter le biais d'endogénéité inhérent à l'approche que nous privilégions.

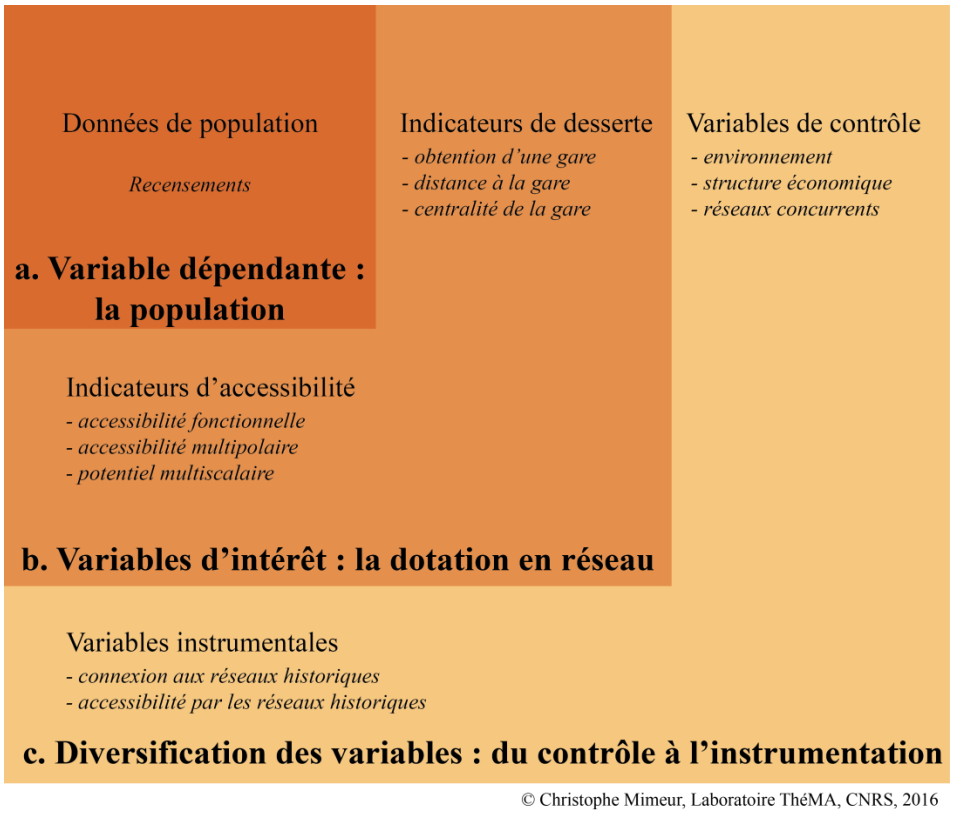


Figure 5. 16. Les degrés de mobilisation de la base de données pour étudier l'effet du réseau

La formulation que nous avons donnée à notre modèle analytique nous conduira de plus à préciser l'échantillon sur lequel portera l'étude de l'effet du réseau de manière à limiter le biais de sélection. Dans ce cadre, la mobilisation des deux premiers degrés semble opportune pour le définir au début de notre troisième partie. Par la suite, la spécification du modèle prend en compte l'ensemble des degrés définis dans la *Figure 5. 16.* A ce stade de nos investigations, nous misons sur la complexification progressive du second bloc pour affiner les estimations, ainsi que sur la mobilisation des variables instrumentales pour limiter l'endogénéité.

Par ailleurs, la large utilisation du multigraphe dans ce chapitre, adaptable à d'autres réseaux, nous conduit à explorer une seconde valeur ajoutée d'un tel formalisme, où le langage mathématique peut nous permettre d'interroger et de tester des interactions entre niveau de population et création du réseau, à partir de règles simples.

Chapitre 6. MODELISER LES INTERACTIONS DANS UN CONTEXTE GEOHISTORIQUE

Introduction

Dans les défis des projets géohistoriques d'une part, et dans la place centrale que l'on donne à l'accessibilité dans cette partie d'autre part, notre réflexion nous pousse à diversifier les approches méthodologiques dans la géographie théorique et quantitative, et de les adapter aux objectifs de notre travail de recherche. Si le précédent chapitre s'attache à capitaliser l'information géohistorique dans une perspective analytique de l'effet structurant de l'infrastructure de transport sur la croissance démographique locale, alors nous interrogeons ici l'inverse de la relation, tel qu'il est évoqué dans la controverse scientifique soulevée dans le premier chapitre (Offner, 1993b ; Pumain, 1982). Pour autant, des auteurs soulignent sa relative rareté dans des perspectives de long terme (Kasraian et al., 2016). La démarche de ce chapitre s'inscrit toujours dans le challenge émergent des SIG-H, en exploitant pleinement les capacités de l'objet *« graphe »* construit dans le *Chapitre 4*. Elle trouve sa place dans notre positionnement géohistorique et constructiviste, visant à qualifier spatialement les observations empiriques évoquées dans le *Chapitre 3*, en interrogeant les capacités de notre instrument de recherche à mieux qualifier la phase de croissance du réseau ferroviaire français (Garrison et Levinson, 2014).

A ce stade, il s'agit de complexifier l'approche par une modélisation des interactions qui permet de croiser l'information démographique et le développement réticulaire, en postulant que le chemin de fer se dessinerait *« en fonction d'une hiérarchie de tailles et de dynamismes urbains qui préexistait à son installation »* (Pumain, 1982). On se place de ce point de vue dans une démarche empirique à l'aide de la base de données FRANcE : ce n'est plus ici un modèle analytique mais un modèle d'évolution basé sur un formalisme mathématique et algorithmique. Ce chapitre vise alors à construire un modèle d'évolution d'un réseau de transport, à partir de règles simplifiées, mais qui s'inspirent de mécanismes fondamentaux décrits tout au long de la thèse et dans les théories d'arrière-plan. Appliquées à un modèle théorique, ces règles ont vocation à être appliquées dans le cadre de la base de données FRANcE de manière à confronter les résultats issus de la démarche de modélisation avec la réalité de la construction du réseau ferroviaire français, et ainsi de qualifier les convergences et divergences dans la construction de nouvelles connaissances géohistoriques.

Dans ce cadre, nous reviendrons d'abord sur la démarche de modélisation en géographie et plus spécifiquement celle de la croissance des réseaux de transport, dans la géographie des transports, mais pas seulement (6.1.). Alors nous proposerons l'inscription de notre modèle dans la littérature existante, en l'adaptant à nos questions géohistoriques, ce qui nous permettra de décrire et qualifier les éléments du modèle, ainsi que ses règles et ses paramètres (6.2.). Enfin, nous proposerons une exploration théorique du modèle, de manière à qualifier son fonctionnement (6.3.), ce qui nous permettra d'envisager par la suite son application empirique dans le cadre de la croissance du réseau ferroviaire français dans la base de données FRANcE.

6.1. La modélisation des interactions entre transport et territoire

Jusqu'ici, le terme de modélisation a revêtu plusieurs acceptions, englobées par la définition proposée par François Moriconi-Ebrard : *« énoncé théorique fondé sur la simulation schématique et formalisée, dans un langage donné (verbal, mathématique, graphique, ...), d'un objet de connaissance et, au-delà, d'une réalité »* (Lévy et Lussault, 2003). Il semble vain de définir ce terme de manière précise et exhaustive. Nous nous contenterons de revenir ici sur les acceptions utilisées jusqu'à maintenant, pour envisager un degré supérieur de la modélisation dans ce chapitre. Les premières modélisations proposées dans le cadre de cette thèse ont d'abord été des modélisations graphiques, parce que l'une de leurs principales propriétés est didactique (Durand-Dastès, 2001). La seconde acception de la modélisation adoptée jusqu'à maintenant réside dans la construction et la formalisation d'une information géohistorique (*Chapitres 4 et 5*), *« influencée par la nature de la réalité qu'elle sert à étudier et par les traditions et les fondements épistémologiques des disciplines qui la pratiquent »* (Durand-Dastès, 2001). La troisième est à chercher dans la formalisation d'un modèle analytique (*Chapitre 5*) répondant à des hypothèses, reposant sur l'effet d'une nouvelle infrastructure sur la croissance démographique locale. Celle envisagée ici est encore d'un autre ordre.

6.1.1. **Un nouveau degré dans la modélisation des interactions géohistoriques**

L'objectif de ce chapitre est bien de caractériser la phase de croissance du réseau, temporellement et spatialement : la démarche de modélisation dynamique s'inscrit donc dans le cadre épistémologique que nous avons défini pour cette thèse. Alors que le chapitre précédent a proposé de formaliser le processus géographique par un processus explicatif, il s'agit désormais *« d'imaginer que l'explication des évolutions observées relève pour une part de dynamiques évolutives plus générales, communes à un certain nombre de systèmes, et que l'on peut abstraire dans des modèles, mathématiques ou informatiques »* (Pumain, 2003). Ce n'est plus l'effet de l'infrastructure que l'on cherche à expliquer ici mais le processus d'adaptation et de causalité inverse : la phase d'expansion du réseau répond-elle à une entreprise d'échelle globale, d'échelle locale (Dupuy, 1993 ; Pumain, 1982), suivant quelles temporalités ? A partir de quelles règles simples de construction peut-on envisager la formation et la morphogénèse du réseau ferroviaire français ? Dans quelles mesures ces règles permettent-elles de reproduire la forme, la structure et la hiérarchie du réseau effectivement construit ? Ainsi, passer de l'étude de l'effet à l'étude des processus nous fait passer d'une modélisation statique à une modélisation dynamique. Dans le même temps, d'un point de vue méthodologique, se posent d'autres questions : *« jusqu'à quel point modéliser le « monde réel » permet-il de le comprendre ? Pour nous géographes, que cachent nos modélisations de la complexité spatiale lorsqu'elles tentent de la rendre intelligible ? »* (Saint-Gerand, 2005). C'est la confrontation de ces questions thématiques et méthodologiques, dans le challenge émergent des SIG-H pour la construction de connaissances géohistoriques, qui occupe les prochaines lignes, en confrontant les aspects théoriques, méthodologiques et appliqués à notre question.

Dans les cadres logiques de la modélisation définie par François Durand-Dastès, le recours à la modélisation dynamique en géographie réside dans la mise en œuvre de *« simulations qui constituent un deuxième temps de la déduction. [Elles] […] peuvent incorporer des données plus ou moins largement issues d'observations ; elles aboutissent à des résultats concrets, qui expriment alors ce à quoi pourrait ressembler la réalité si elle fonctionnait conformément aux hypothèses de base du modèle »* (Durand-Dastès, 2001). Cette approche méthodologique est toutefois à exposer de manière d'autant plus didactique que ce *« type de raisonnement […] a pu soulever des réticences de la part des praticiens de certaines sciences sociales »* (Durand-Dastès, 2001), dans notre perspective géohistorique en particulier. L'intérêt d'une telle approche n'est pas à chercher dans les techniques divinatoires, mais dans des projections statistiques et territoriales, à même de pouvoir être comparées à des situations par ailleurs connues. Cette méthode se rapproche des nombreuses réflexions quant aux fondements d'une théorie de la géographie *« permettant de comprendre l'évolution des systèmes territoriaux […] qui conduisent à l'organisation et à la différenciation des entités localisables »* (Pumain, 2003), notamment par la confrontation systématique des résultats obtenus par la simulation avec des observations empiriques afin de tester la validité des hypothèses formulées (Durand-Dastès, 2001). Cette démarche trouve toute sa place dans notre posture méthodologique où les effets de connaissances sont à contrôler de manière empirique.

Nous prenons ainsi appui à la fois sur le contexte géohistorique, sur des théories d'arrière-plan – telles que la théorie des réseaux (Curien, 2000) – et sur les observations empiriques à partir de la base de données FRANcE. La question posée dans ce chapitre est celle de l'émergence d'un phénomène, celui de la diffusion de l'infrastructure ferroviaire, dans l'espace et dans le temps, malgré que le *« statut épistémologique de l'émergence dans les modèles […] demeure […] ambigu »* (Pumain, 2003) : l'ambiguïté est à chercher selon nous dans le compromis entre simplification et réduction. Pour rendre intelligible un modèle, aussi bien que reproductible, la formalisation des règles doit être simplifiée pour vérifier des propriétés fondamentales des relations entretenues entre deux objets (Haggett, Frey et Cliff, 1977). L'utilisation de tels modèles nous rapproche alors des théories de la complexité *« la complexité sociale peut-être rendue plus intelligible si l'on accepte d'énoncer des propositions, lois ou modèles, regroupant ces processus et leurs résultats les plus fréquents, selon une opération dite de simplexité »* (Pumain, 2014).

Quand on l'applique à notre question, nous identifions dans cette démarche la mise en exergue de propriétés qui participent à l'explicitation de changements, de transformations, d'auto-organisation, d'émergence, d'imprévisibilité (Pumain, 1998). Denise Pumain a montré les capacités heuristiques et didactiques de ce type de modélisation (Desmarais et Ritchot, 1998). Elle répond de plus à nos réflexions sur les échelles d'analyses, parce que les théories de l'auto-organisation que nous venons de citer concernent des systèmes qui sont observables à plusieurs niveaux (Pumain, 2003) :

- *« Le niveau microscopique constitué d'un très grand nombre d'éléments en interaction » ;*
- *« Le niveau macroscopique où apparaissent des phénomènes collectifs non-directement prévisibles à partir de la connaissance des états des éléments microscopiques ».*

C'est bien ce que nous interrogeons dans notre question et notre démarche : dans les larges

échelles spatiales et temporelles, cette modélisation permet de suggérer des mécanismes interactifs où l'individu est la structure élémentaire de notre raisonnement, la commune française. Dans cette théorie, l'interdépendance entre chacune des 36 000 communes suggère par ailleurs la mise en valeur de configurations macroscopiques. L'application de la théorie de la complexité d'une part, de la modélisation dynamique d'autre part, participe ainsi à bien *« identifier des attracteurs dans une dynamique […] à un échelon macrogéographique, dans des systèmes complexes où les causalités circulaires aux niveaux plus microgéographiques sont légion »* (Pumain, 2014).

Nous retenons dans cette démarche trois intérêts (Pumain, 2003) pour caractériser une dynamique de croissance d'un réseau, d'un point de vue spatial et temporel :

- La modélisation dynamique a pour vocation de dépasser l'opposition entre idiographie et nomothétie, trouvant ainsi un écho particulier dans notre processus d'acquisitions de nouvelles connaissances géohistoriques : pour cela, elle s'appuie sur la description de propriétés générales, à appliquer à un espace défini, de manière théorique ou empirique ;
- La modélisation dynamique vise à *« repenser la diversité géographique »* (Pumain, 2003), dans laquelle le processus permet d'apprécier la hiérarchie dans les interactions entre les objets, à partir de leurs propriétés initiales, faisant d'eux des émetteurs et des attracteurs qui évoluent au cours du processus ;
- La modélisation dynamique permet de mettre en relation des configurations géographiques dont les impacts sont transcalaires.

Ces trois intérêts sont à lire de manière successive, à la manière dont la démarche du modélisateur s'opère. Ainsi, on peut replacer ce processus dans les ambitions de notre instrument de recherche, d'après les contours que nous en avons dessinés au début de cette partie *(cf. 4.1, p. 141)*. La première phase est celle de l'initialisation du modèle, dont l'un des objectifs est la conceptualisation des données, le *Chapitre 4* y est consacré. La seconde est celle de l'exécution du modèle : ce chapitre a vocation à exposer la question, les déterminants et les paramètres de la démarche de modélisation. La troisième est celle de l'évaluation du modèle (Durand-Dastès, 2001 ; Sanders, 2001) : elle est d'abord théorique dans ce chapitre, elle sera appliquée à la base de données FRANcE dans la troisième partie de la thèse. C'est bien la structure du réseau généré, comparée à la structure effectivement construite qui sera au cœur de nos attentions, de même que la hiérarchie produite dans les deux cas.

Si la question de la modélisation dynamique en géographie a particulièrement intéressé l'évolution du système de villes sur le temps long (Pumain, 1997 ; Sanders, 2001), la question des interactions entre transport et territoire a très tôt fait l'objet de la considération des parties prenantes de la modélisation dynamique, en géographie, mais également du côté des sciences expérimentales. La prochaine section, à travers un état de l'art des modèles d'interactions entre transport et territoire, permet de mieux identifier les propriétés que nous attribuons à notre modèle, et la formalisation des règles que nous souhaitons y appliquer.

6.1.2. **Les familles de modélisation de la croissance des réseaux**

La section précédente a montré comment les questions de la causalité inverse et du rôle d'une structure préexistante peuvent être investiguées par la modélisation dynamique décrivant la croissance des réseaux. Nous proposons ici de présenter quelles approches sont régulièrement entreprises dans la littérature, pour mieux identifier l'identité de notre propre modèle d'évolution du réseau de transport ferroviaire. La lecture de l'état de l'art sur la modélisation dynamique des interactions entre transport et territoire s'appuie d'une part sur la grille de lecture proposée par Hadrien Commenges mais aussi par les lectures de David Levinson, complétées par nos propres investigations qui viennent enrichir les typologies déjà construites par les auteurs que nous mobilisons (Commenges, 2013 ; Xie et Levinson, 2009). Dans tous les cas, les auteurs font bien une distinction entre l'étude causale des infrastructures, faisant appel à l'empirisme et aux modèles statistiques (Banister et Berechman, 2004), et de l'étude interactive de l'évolution des réseaux, qui nous occupe à présent, rapprochant la théorie de l'empirisme.

C'est bien dans le second cas que l'on se place ici : si l'étendue des travaux reste importante, la grille proposée montre la diversité des approches, en y ajoutant une composante d'identification, permettant de voir quelles sont les principales tendances parmi ces familles. La première grille *(Figure 6. 1. a.)* reprend les items de description des approches de modélisation définies par Hadrien Commenges. Cela nous permet de distinguer les travaux selon l'objet : la demande en transport, classiquement abordée par la modélisation en 4 étapes, l'interaction entre transport et occupation et l'évolution intrinsèque des réseaux de transport. La formalisation permet de distinguer les différentes méthodes utilisées pour modéliser les relations entre les objets à partir de modèles plus ou moins mathématisés tels que le modèle gravitaire pour envisager les interactions spatiales. La temporalité s'étale entre le court terme – où les changements sont ponctuels – et le long terme – où les changements sont perceptibles jusqu'à l'échelle du siècle. La spatialité désigne l'échelle à laquelle le problème est posé – de l'intérieur d'une ville à une relation entre un système de villes. Enfin l'arbitrage entre théorie et empirisme participe d'un mouvement entre formalisation théorique d'un modèle et vérification empirique des hypothèses.

La seconde grille confronte les cinq tendances remarquées dans la modélisation des réseaux *(Figure 6. 1. b.)*, envisagées de manière chronologique dans la littérature (Xie et Levinson, 2009) mais dont les frontières temporelles restent floues. A chaque fois, nous nous efforçons de résumer les principales approches en appliquant la carte d'identité définie dans la première grille. Cela nous amène à détecter les évolutions en termes d'objet, de formalisation des relations, mais aussi l'évolution des préoccupations, qui se traduisent en termes d'échelles spatiales et temporelles. On voit ainsi comment le long terme a été à l'origine des processus de modélisation de l'évolution des réseaux, mais a cependant été longtemps délaissé au profit d'analyses de court terme et de moyen terme pour des modèles de prévision. Ainsi, c'est depuis les années 2000 que la science des réseaux a réinvesti les dynamiques de long terme, dans des processus d'optimisation.

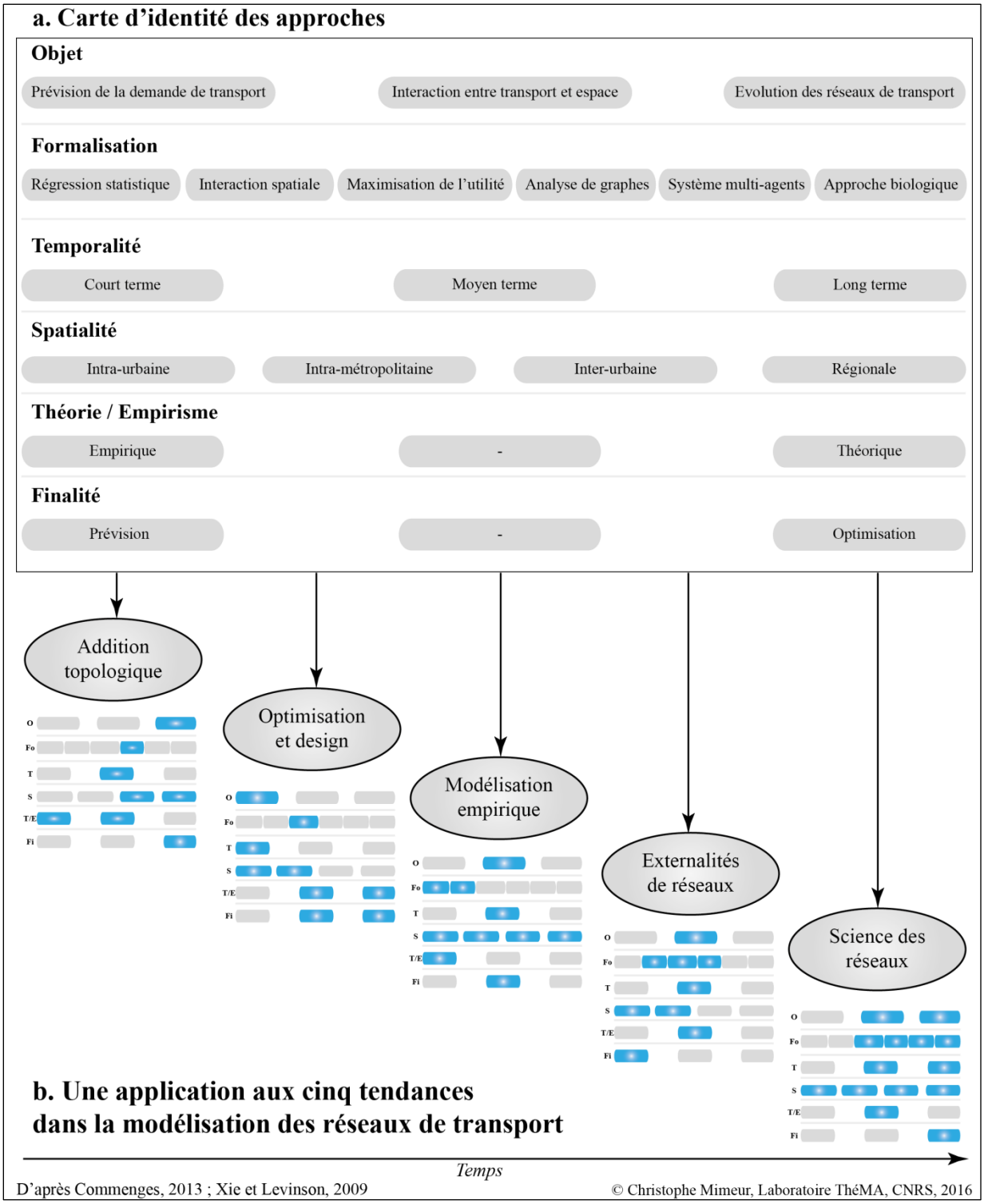


Figure 6. 1. La modélisation dynamique des interactions entre réseau de transport et territoire – Carte d'identité des approches et application aux cinq tendances dans la modélisation des réseaux de transport

a. *Les approches traditionnelles des interactions entre transport et espace*

Dans les années 1960, c'est bien dans la géographie des transports que sont entreprises les premières études traitant de la phase de croissance des réseaux (Xie et Levinson, 2009). Ils trouvent une résonance particulière ici parce que ce sont les premiers à utiliser le formalisme des graphes (Garrison, 1960), en questionnant les propriétés topologiques des réseaux de transport, et plus particulièrement le phénomène d'addition. Le processus formalisé par Garrison est alors très simple : à chaque étape, on connecte la ville la plus importante qui n'est pas encore connectée. La plupart de ces travaux cherche à montrer les propriétés de connectivité et de connexité des réseaux constitués (Kansky, 1963), par l'opération de règles simples, heuristiques et intuitives, qui ne sont pas systématiquement posées de manière mathématique. Dès 1971, Black propose de modéliser le réseau ferroviaire dans le Maine, aux Etats-Unis à partir d'un potentiel d'accessibilité fondé sur les populations du milieu du XXème siècle : il montre la logique centre/périphérie inhérente à la construction du réseau (Black, 1971) à partir de la construction d'un réseau en arbre depuis Portland. Le processus d'extension des réseaux répond à des logiques d'investissement. On construit à chaque étape un lien fort, jusqu'à ce qu'un seuil soit atteint. A l'époque, les capacités de calcul ont largement contraint les investigations, même si les travaux traitent davantage de réseaux à l'échelle régionale voire nationale, sur des temporalités moyennes et longues. D'après notre lecture de la littérature, ce type de modélisation est resté *« dormant »* (Xie et Levinson, 2009) entre ces travaux du début des années 1970 et le début des années 2000 (Commenges, 2013).

Par la suite, la modélisation des interactions s'est davantage tournée vers l'angle opérationnel, à partir des années 1970. La socio-économie des transports, l'économie et l'aménagement sont les disciplines privilégiées, avec une réduction croissante des échelles tant spatiales que temporelles. L'intérêt se concentre de plus en plus aux échelons intra-urbain ou intra-métropolitain. Ils représentent la plupart des démarches qui s'attachent à des processus d'optimisation et de design des réseaux (Xie et Levinson, 2009), dans une temporalité qui est souvent courte. Dégager des structures semblent une entreprise vaine à de telles échelles, tant une décision prise à l'échelle locale peut être très efficiente à cette même échelle mais peut apparaître aléatoire à des échelles plus larges. Les points de vue locaux cherchent alors à améliorer le flux dans le réseau, en traitant des phénomènes de congestion à travers l'appréhension de comportements individuels : la confrontation de scénarios et la prédiction participent à la prise de décision, dans la gestion des carrefours et de flux de circulation, en fonction de la capacité et du débit des différents éléments du réseau, fortement hiérarchisés.

La modélisation de la croissance des réseaux a reçu un regain d'intérêt dans les années 1980 avec la complexification des processus économiques liés aux interactions des acteurs. Dans ce cadre, la plupart des démarches modélisatrices s'appuie sur un réseau théorique, souvent de taille réduite, parce que les approches sont souvent désagrégées et requièrent une puissance de calcul importante. Les économistes questionnent alors les régimes de propriétés et leur influence dans les investissements pour la croissance des réseaux, à partir de mosaïques de règles juridictionnelles (Xie et Levinson, 2009), et d'interactions entre les externalités procurées par le réseau. S'il est donc difficile de définir des échelles d'analyses, le cadre intra-métropolitain ou intra-urbain reste souvent privilégié. A l'échelle régionale, ce sont des processus d'inertie et de *« path dependence »* qu'il s'agit de mettre en exergue, même s'il reste difficile de juger du moment à partir duquel le processus

de décision peut être qualifié d'exogène.

A ce propos, la modélisation classique des interactions entre transport et territoire s'attache finalement à une description exogène du réseau de transport. A partir de scénarios d'aménagement et de système de transport prédéfini, la modélisation classique consiste alors à estimer les flux en fonction de l'occupation du sol, en suivant des règles simples parmi lesquelles la friction de la distance énoncée par Tobler est la plus commune. Ce type de démarche s'inscrit alors dans la prévision de la demande de transport, où l'hypothèse d'évolution d'une infrastructure est étudiée par ses répercussions sur les flux qu'elle est susceptible de générer. Ainsi, l'évolution du réseau de transport est envisagée de manière exogène dans le réseau : il est défini *a priori* et ne répond pas forcément à des objectifs d'optimisation (Commenges, 2013).

La place que nous souhaitons donner à notre modèle d'évolution d'un réseau de transport est finalement différente de celle des deux dernières familles exposées. Ce qui distingue notre approche des autres est la place que l'on assigne au réseau de transport dans la démarche de modélisation. Alors que les démarches les plus communes considèrent l'évolution du réseau comme le paramètre qu'il s'agit de tester, notre approche privilégie une évolution endogène du réseau de transport : l'infrastructure est le sujet même de la modélisation. Dans la question que nous posons dans ce chapitre, la modification des flux d'une part et la structure même du réseau d'autre part sont donc à envisager de manière endogène. Ces deux éléments, répartitions des flux et structures, sont eux-mêmes influencés par des paramètres d'entrée, les niveaux de population dans notre cas : alors nous ne plaçons pas ce modèle dans des perspectives de prévision ou de prospection mais bien dans des perspectives d'optimisation.

b. *La modélisation de l'optimisation dans la science des réseaux*

Notre propos a déjà montré comment le traitement de l'objet réseau est pluridisciplinaire. Si aujourd'hui, une attention particulière est portée par les géographes mais aussi les sociologues à propos des réseaux sociaux, l'application de la théorie des graphes, comme formalisme mathématique, a élargi le cercle des intérêts portant sur l'évolution des réseaux. Si le graphe n'est qu'un prétexte graphique pour certains, cet objet revêt des propriétés géométriques et mathématiques qui sont aussi explorées par les sciences expérimentales.

Il n'est donc pas si étonnant de trouver une littérature portant sur l'évolution du réseau du côté des sciences expérimentales. Il est en revanche beaucoup plus surprenant de trouver des références aux réseaux de transport chez les physiciens, de lire des conceptions d'agencement et d'optimisation des réseaux de transport dans la revue *Nature*. *« Traditionnellement, les physiciens modélisent les dynamiques d'un réseau de transport comme un processus d'optimisation, avec un intérêt particulier dans l'extraction de caractéristiques agrégées du réseau à partir de fonctions simples de coûts »* (Xie et Levinson, 2009) : on reconnaît volontiers dans cette acception les enjeux que nous avons posés dans la modélisation dynamique des réseaux dans notre perspective

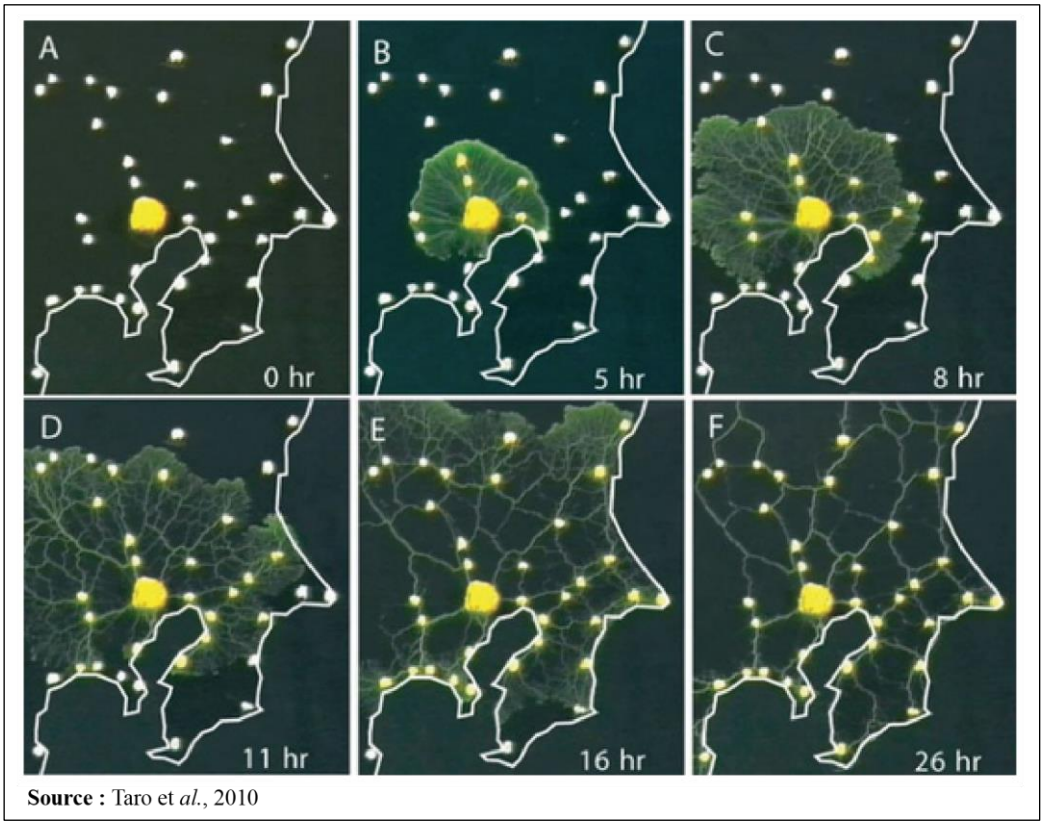
géohistorique. Là encore toutefois, la plupart de ces formalisations, basées sur l'analyse de graphes, s'appuie souvent sur des réseaux théoriques, qualifiés parfois d'idéaux, *« capables de reproduire des propriétés topologiques de réseaux existants, en considérant le réseau comme la réalisation d'un optimum pour une quantité donnée »* (Louf, Jensen et Barthelemy, 2013). Finalement, le langage informatique et mathématique semble souvent l'emporter sur des considérations thématiques et contextualisées.

De premiers travaux se sont attachés à caractériser la connexité dans un réseau, ce sont ceux de la première tendance que nous avons identifiée *(Figure 6. 1. b.)*, parmi lesquels on voit comment le plus proche voisin est connecté dans l'optimisation du réseau d'une part, et comment la distribution de la hiérarchie des nœuds obéit à une loi puissance (Xie et Levinson, 2009). La croissance des études sur des réseaux complexes montre en effet comment les sciences expérimentales voient dans les comportements des réseaux des processus invariants d'échelle (Albert et Barabási, 2000 ; Barabási et Frangos, 2014), selon lesquels l'apparition de nouveaux nœuds dans le graphe se traduit par leur rapprochement vers les nœuds les mieux connectés, au détriment des autres. Ce modèle d'attachement préférentiel montre une évolution de la structure des réseaux de trois manières différentes : la croissance du réseau par l'apparition de nouveaux nœuds, leur attachement par l'apparition de nouveaux liens et l'évolution des propriétés du réseau.

Ainsi, plus spécifiquement pour étudier l'évolution de la structure du réseau de transport, la complexification de l'approche requiert la modification des propriétés tant locales que globales des réseaux. Dans un processus d'optimisation d'un modèle de trafic, des physiciens proposent une structure générique en arbre, à partir de laquelle les changements de paramètres font apparaître différents types de réseaux en arbre (Barthélemy et Flammini, 2006). De ces réseaux, ils déduisent la formation d'une hiérarchie faisant apparaître des structures en hub, séparées dans l'espace (Louf, Jensen et Barthelemy, 2013). Pour autant, la plupart de ces conclusions s'appuie sur des répartitions théoriques des villes, répondant à une distribution de la population issue d'une loi puissance, et les conclusions tirées sont des conséquences de larges échelles à partir d'indicateurs globaux. L'approche que nous envisageons est en revanche empirique, s'appuyant sur la réalité géographique et sur la répartition historique des populations sur le territoire français, dans une perspective de comparaison entre un réseau guidé par les règles que nous lui attribuons et un réseau effectivement construit.

De manière périphérique mais néanmoins croissante et finalement encore plus originale, des approches biologiques ont cherché à résoudre le problème de l'optimisation des réseaux de transport, au croisement de la théorie et de l'empirisme. En effet, le *Physarum Plasmodium*, un organisme à la frontière du végétal et de l'animal, a la particularité de se répandre en réseau quand il cherche à s'alimenter (Tero, Kobayashi et Nakagaki, 2006). A la recherche d'un minimum d'efforts pour un maximum de rentabilité, ce micro-organisme fait l'objet de nombreuses expériences *in-situ*, dont la plus célèbre est la reproduction *(Figure 6. 2)* du réseau métropolitain de Tokyo (Tero et al., 2010). Des expériences en laboratoire ont aussi permis de tester la reproductibilité du réseau d'autoroutes au Mexique (Adamatzky et al., 2011). A l'intérieur de boîtes de Petri, les chercheurs disposent des flocons d'avoine, respectant l'hétérogénéité spatiale d'un

territoire, et proportionnellement à la population des points. Après le choix d'un point de départ sur lequel est déposé le micro-organisme, l'expérience montre comment il se propage au début du processus dans de nombreuses directions, de manière aléatoire et exhaustive. Après avoir identifié l'ensemble des points qu'il doit rejoindre, le micro-organisme ne conserve que les lignes qu'il estime utile à son approvisionnement, à travers l'établissement d'un réseau minimum, en éliminant des chemins qu'il juge redondant. En revanche, si les chercheurs identifient les propriétés qui le poussent à réduire les chemins qu'il envisage au départ, ils ont davantage de difficultés à présager des règles qui le conduisent à former un réseau très dense au début du processus (Niizato, Shirakawa et Gunji, 2010), à travers un enchevêtrement où il semblerait vain de chercher des structures. On voit ce phénomène dans les premières huit heures du processus *(Figure 6. 2)*, où le micro-organisme n'a pas encore terminé d'explorer tous les points qu'il commence déjà son processus de sélection des itinéraires. Par ailleurs, le processus montre comment il est dépendant du point de départ de l'expansion du *Physarum Plasmodium*. Finalement, au bout de 26 heures, on arrive à reproduire le réseau métropolitain de Tokyo, sans toutefois que l'on puisse y déceler une hiérarchie dans les chemins dessinés. Si le recours à ce type de modélisation semble pertinent dans les processus de sélection d'itinéraires une fois qu'un réseau a atteint son expansion maximum, il semble moins intéressant pour traduire la phase de croissance du réseau, puisque les processus qu'il met en jeu ne permettent pas de caractériser par les mots les phases de naissance, croissance et maturité d'un réseau de transport (Niizato, Shirakawa et Gunji, 2010) et dans sa hiérarchisation progressive.



Source : Taro et *al.*, 2010

Figure 6. 2. La reproduction *in-situ* du réseau métropolitain de Tokyo à l'aide du *Physarum Plasmodium*

Ainsi, si les approches physiques d'une part, biologiques d'autre part, semblent reproduire correctement des processus d'optimisation basés sur des structures de type *hub-and-spoke*, alors notre approche postule que la *« grande majorité des réseaux spatiaux existants ne semblent pas le résultat d'une optimisation globale, mais plutôt de l'addition progressive des nœuds et des segments résultant d'une optimisation locale »* (Louf, Jensen et Barthelemy, 2013). Le recours à la modélisation dans la science des réseaux, à partir du formalisme de graphes, semble donc répondre aux intérêts de la théorie de l'auto-organisation que nous avons dégagée dans la sous-section précédente. Elle se trouve ici enrichie des nombreuses approches développées depuis les années 1960, parmi lesquelles l'attachement préférentiel permet des changements de propriétés aux échelles locales et globales. Les modèles générant des réseaux de transport, obéissant à cette théorie et à ces approches, n'ont pas connu un réel succès dans les deux dernières décennies du XXème siècle (Xie et Levinson, 2009) alors que les premiers essais sont l'initiatives de géographes. D'autres voies s'étaient ouvertes vers l'assignation et le management du trafic, routier en particulier, quand l'automobile a requis presque toute la concentration des modélisateurs (Dupuy, 1999).

En remettant l'évolution même d'un réseau de transport au centre du débat, nous reprenons le postulat faisant l'hypothèse que *« la hiérarchie dans un réseau de transport est une propriété émergente, contrairement à l'idée reçue qui voudrait que la hiérarchie soit le résultat d'une planification »* (Commenges, 2013 ; Yerra et Levinson, 2005). La prochaine section s'attache ainsi à caractériser le modèle de croissance du réseau ferroviaire élaboré avec la base de données FRANcE, à partir de l'identification d'objectifs, d'hypothèses, de règles que nous envisageons pour comparer un réseau simulé du réseau historique.

6.2. La construction d'un modèle de croissance d'un réseau de transport pour FRANcE

La modélisation de l'évolution des réseaux de transport a connu un regain d'intérêt au début des années 2000. Elle se structure surtout autour d'une même équipe de recherche, dirigée par David Levinson. Nous relevons dans la structure de l'ouvrage qu'il a co-écrit avec Feng Xie, *Evolving transportation networks*, les principaux enjeux dans la modélisation de la croissance des réseaux (Xie et Levinson, 2011), en croisant les approches théoriques nécessaires à la formalisation du modèle d'une part, jusqu'à l'exécution même qui aboutit à une évaluation empirique des règles choisies dans le modèle *(Figure 6. 3)* d'autre part, en appui des interprétations de la modélisation dynamique qui ont été faites depuis le début du chapitre.

Cette section est à lire comme la construction de l'identité du modèle, pour identifier sa place dans la littérature existante et la place du modélisateur dans le processus que nous souhaitons voir émerger. Ensuite, il s'agit de formaliser les formes d'initialisation du modèle, à partir desquelles l'exécution du modèle consiste en une articulation entre un module statique et un module dynamique.

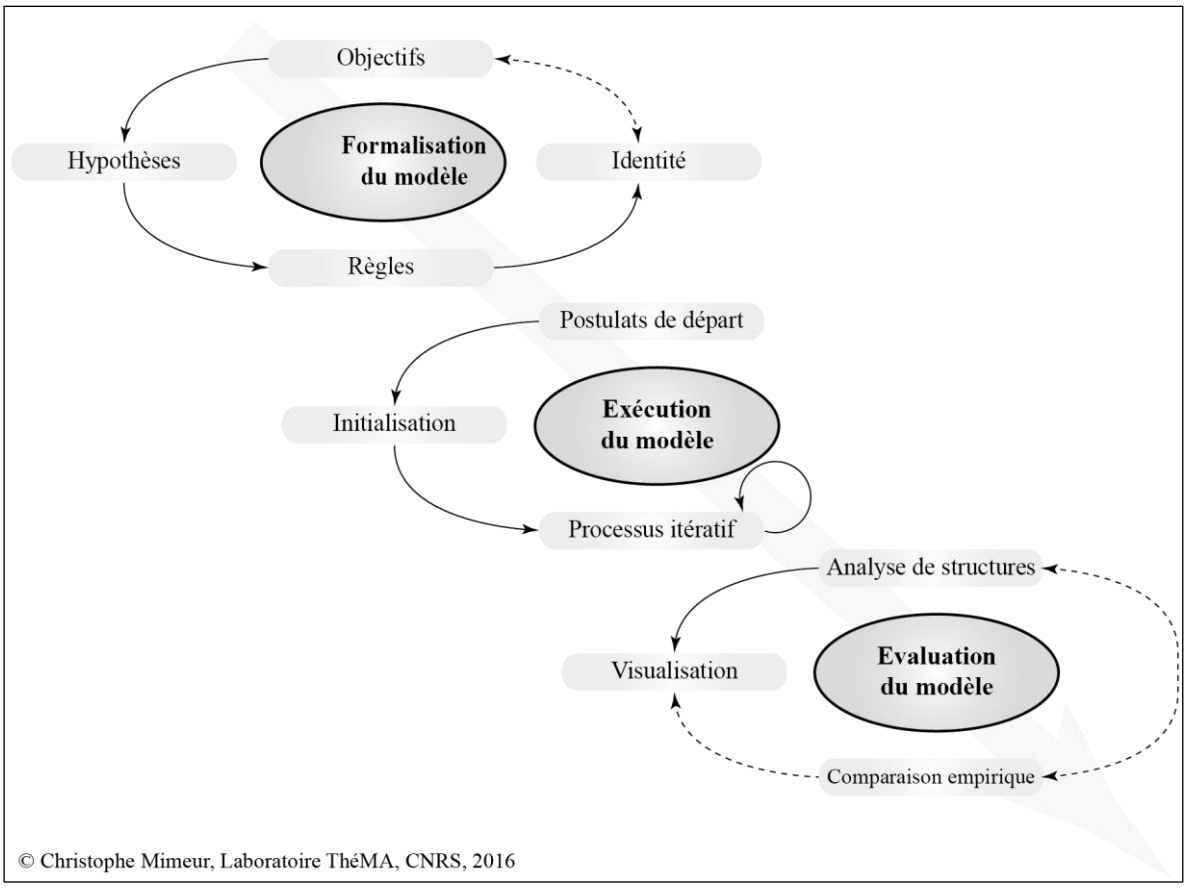


Figure 6. 3. Le processus de modélisation dynamique du réseau de transport ferroviaire

6.2.1. **La définition et le positionnement du modèle**

L'objectif de ce chapitre est bien d'étudier les phénomènes d'émergence et de hiérarchie du réseau de transport. Cet objectif s'inscrit dans le cadre géohistorique qui a vu la création d'un réseau ferroviaire à partir d'une innovation dans la première moitié du XIXème siècle. Ce modèle s'inscrit dans le débat sur les effets structurants des infrastructures de transport. Nous reprenons donc ici des éléments fondamentaux de la théorie des réseaux, abordée dans le premier chapitre. Le modèle part de l'hypothèse simple de la diversité et de l'hétérogénéité spatio-temporelle, dans lesquelles nous donnons un double rôle au modélisateur : il est tout aussi bien l'incarnation d'une *« main invisible »*, d'un laissez-faire dans l'évolution du réseau en réponse à une *« structure urbaine préexistente »* (Pumain, 1982), à travers la multiplication des compagnies privées jusqu'à la seconde moitié du XIXème siècle. La seconde incarnation possible du modélisateur est celui d'un ingénieur des Ponts et Chaussées, aux pouvoirs centralisés et démesurés, où l'objectif de sa politique est de répondre à une efficacité économique la plus rationnelle possible s'il est fortement impliqué, où il mène une *« politique suiveuse »* s'il se contente de suivre la hiérarchie urbaine établie, répondant à un besoin existant (Baptiste, 2003). De cette sorte, nous participons au débat initié par Offner sur les hésitations méthodologiques, en ôtant ici toute considération politique et tout interventionnisme. Le modélisateur prend donc connaissance d'une situation – c'est l'objet de cette sous-section –, il entreprend ensuite un processus de construction – par lui-même ou planifié, de manière normative,

à partir d'un ensemble de règles très simples, à vocation toute aussi démonstrative que didactique. Dès lors, le modélisateur a pour mission dans l'ultime étape de valider des affirmations, telles que celles rencontrées dans littérature, stipulant par exemple qu'*« une rapide analyse du réseau urbain français aujourd'hui montre une adéquation assez bonne entre la taille des villes et la qualité des infrastructures de transport qui les desservent »* (Baptiste, 2003).

Quand on évoque le principe de hiérarchie, il semble incontournable de se tourner vers la théorie des places centrales développée par Christaller (Christaller, 1933), parce qu'elle mobilise les principes d'agglomération, d'accessibilité et d'interaction spatiale, au cœur de notre raisonnement, dans la distribution des populations et des activités. Cependant, dans cette formalisation, la hiérarchie n'est pas appréhendée comme un phénomène d'émergence (Yerra et Levinson, 2005), mais comme le résultat d'une distribution inégale des populations et des activités. Dans des perspectives de long terme, cette formalisation propose surtout des approches statiques comparées (Camagni, 1996) et ne laisse finalement que peu de place à la hiérarchie de réseau. En revanche, ce sont ces trois mêmes principes évoqués plus haut qui sont à l'origine de notre réflexion sur les mécanismes de croissance d'un réseau de transport dans ce chapitre. Ainsi, nous formalisons ici un modèle qui permet d'analyser l'émergence d'une hiérarchie dans le réseau de transport, à partir d'un processus décentralisé et transcalaire, s'inscrivant dans notre positionnement géohistorique à partir de la structure élémentaire de la commune, répondant aux principes d'accessibilité, d'agglomération et d'interaction. De la même manière que la famille de modèles proposée par l'équipe de l'université de Minneapolis, nous cherchons à nous intéresser en sortie à la structure du réseau produit par les mécanismes du modèle. En revanche, s'ils s'appuient sur des espaces théoriques, nous misons sur l'intérêt empirique de l'approche.

Dans tous les cas, nous faisons le postulat ici d'une stabilité de la hiérarchie urbaine sur le temps long (Guérin-Pace, 1993), en faisant le choix d'un niveau initial de la population pour modéliser l'émergence du réseau de transport. En effet, *« malgré les modifications continuelles auxquelles elle est soumise, [la trame urbaine] possède un caractère persistant ou permanent »* (Guérin-Pace, 1993). En plus de la répartition de la population qui est définie *a priori*, notre modélisation, à visée empirique, a pour ambition de dépasser le classique espace homogène et isotrope, toutefois souvent utilisé dans les théories économiques traditionnelles. La méthodologie développée dans le *Chapitre 4* participe à la prise en compte de la rugosité de l'espace, telle que le relief. Par ailleurs, la structure adaptative de notre modèle de données nous autorise à imaginer d'autres critères pour déterminer d'autres formes de friction dans l'espace. Parmi eux, on peut imaginer raisonnablement que les contraintes topographiques ont un rôle non-négligeable dans les tracés des voies ferroviaires[35]. A partir de là, il s'agit de styliser, c'est-à-dire définir une stratégie de comportement, capable de créer un réseau de transport.

Notre approche privilégie l'optimisation d'un réseau de transport à partir d'un processus d'efficience économique. Alors, la littérature privilégie des modèles économiques simples, parmi lesquels l'arbitrage entre coûts et bénéfices est le critère de construction du réseau (Louf, Jensen et

[35] A ce propos, la technologie ferroviaire a vu une transition entre la grande vitesse ferroviaire depuis le XIXème siècle et la très grande vitesse ferroviaire du XXème siècle. La première privilégie les courbes pour éviter les fortes pentes alors que les tracés de LGV tendent à privilégier la ligne droite, quitte à emprunter des trajectoires à fort pourcentage.

Barthelemy, 2013). Ce type d'approche permet l'évaluation systématique de la place de chaque entité dans un ensemble plus global, et est de plus désormais utilisé dans la plupart des processus d'évaluation des investissements en infrastructures depuis les années 1960 (Delaplace, 2014). Ainsi, dans la diversité des modèles de l'équipe américaine, ces derniers ont en commun de prendre en compte le flux qui transite sur chaque lien du réseau, de le traduire en termes de potentiels revenus. C'est à ce stade qu'est formalisée l'interaction spatiale entre les structures élémentaires du modèle. C'est aussi à partir de ces mêmes flux que l'on peut estimer les investissements (Black, 1971 ; Levinson et Yerra, 2006) à faire sur le réseau. Alors naturellement, dans notre cas, la construction de l'infrastructure ferroviaire se reporte sur les liens les plus rentables, capables de dégager finalement davantage de revenus. Dans un processus d'optimisation d'une part, dans notre approche misant sur l'impact de la différenciation des vitesses d'autre part, l'investissement se traduit par l'amélioration de la vitesse le long de chaque lien candidat à la construction d'une infrastructure ferroviaire, répondant finalement au principe de la maximisation de l'utilité par un processus de sélection. Le formalisme des graphes, mis au centre de notre instrument de recherche depuis le début de cette seconde partie, semble le meilleur candidat pour modéliser ces mécanismes dans le processus : les indicateurs d'intermédiarité participent à la définition du flux le long d'un réseau. Dans le langage mathématique et informatique, cet arbitrage entre gains et coûts de cette optimisation s'inscrit dans les problèmes NP-Difficile, dans la mesure où il cherche la minimisation de la moyenne des distances et que la résolution ne peut être qualifiée de très efficace, venant justifier là encore le rapprochement entre géographie et sciences informatiques dans notre questionnement.

Après avoir fixé les objectifs et les règles comportementales que l'on souhaite attribuer à notre modèle, il s'agit de cadrer son application dans l'espace et dans le temps. La temporalité de ce modèle s'inscrit naturellement dans des perspectives de long terme. Là encore, la théorie des réseaux nous montre comment le processus de construction du réseau ferroviaire est caractérisé par des temporalités lentes et cumulatives (Garrison et Levinson, 2014), venant justifier le recours au modèle itératif, par ailleurs très souvent mobilisé dans la littérature (Gastner et Newman, 2006 ; Louf, Jensen et Barthelemy, 2013). Alors que les modèles de génération s'appuient sur des itérations dites *« gourmandes »* (Louf, Jensen et Barthelemy, 2013) selon des critères géométriques. Notre approche diffère ici parce que nous postulons que, sur la longue durée, et compte tenu du contexte géohistorique, la situation qui préexiste n'est pas vierge et des relations entre les structures élémentaires préexistent. Ce ne sont alors pas des propriétés géométriques qui peuvent prévaloir, dans notre cas, c'est plutôt l'interaction spatiale dans cette situation qui n'est pas vierge qu'il convient de concrétiser.

La spatialité de ce modèle répond aux cadres fixés par notre question de recherche. La visée est donc transcalaire dans la mesure où ces règles appliquées à l'échelle locale participent à l'émergence de structures à l'échelle globale. A la différence des nombreux modèles intra-urbains, le nôtre s'inscrit à l'échelle inter-urbaine, pour l'ensemble du territoire français. Toutefois, des focus sur certaines zones paraissent naturellement possibles : l'une des questions qui se posent est alors le rôle de la situation globale dans la construction du réseau observée à l'échelle régionale.

Enfin, à la différence des nombreux modèles évoqués jusqu'à présent dans la littérature, la richesse de la base de données FRANcE nous autorise à évaluer le mécanisme d'investissement à

l'origine de ce modèle dans le domaine empirique, par la comparaison de la situation réelle avec la situation simulée, alors que de nombreux modèles s'appuient sur des structures initiales théoriques et restent évasifs quant à la nature du réseau modélisé (Commenges, 2013). Cette comparaison peut s'appuyer d'abord sur les indicateurs globaux de structures du réseau. Comment le modèle se comporte-t-il, est-il capable de reproduire globalement un réseau de transport répondant à des critères globaux de structures, et notamment de distribution des flux ? Toutefois, le questionnement nous pousse à observer la morphologie, le tracé ainsi que la mise en place d'une hiérarchie, perçue comme une *« structure émergente »* : dans quelle mesure le modèle participe à la reconstruction historique du réseau ferroviaire français, dans sa phase de croissance jusque dans le premier quart du XXème siècle.

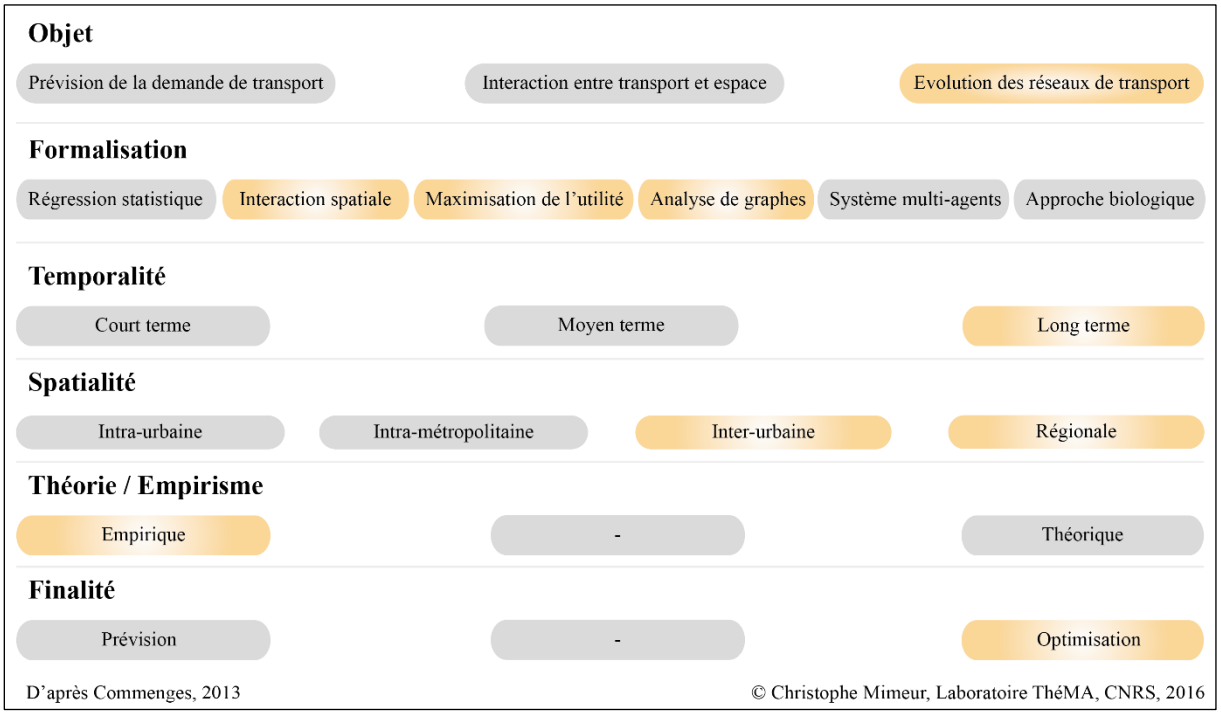


Figure 6. 4. La carte d'identité du modèle d'évolution du réseau ferroviaire

6.2.2. L'initialisation du modèle : d'une situation préexistante à l'interaction spatiale

Dans la démarche globale de la modélisation de l'évolution du réseau de transport, la seconde étape consiste à l'exécution du modèle. A partir du fonctionnement général, il s'agit de définir pratiquement les étapes nécessaires à la simulation d'un réseau de transport : de la prise en compte d'une situation initiale à la fin du processus itératif *(Figure 6. 3)*. Cela requiert la description précise des mécanismes, qui passe nécessairement par un formalisme mathématique et informatique. Elle participe toutefois à une meilleure compréhension du fonctionnement global, et donne de premières pistes à l'étape ultime de la démarche, qui consiste en l'interprétation du modèle

simulé (Durand-Dastès, 2001).

La situation initiale du modèle correspond à un réseau de projets transactionnels, tel que nous l'avons défini dans la théorie des réseaux *(Figure 6. 5. a.)*, et qui maximise l'étendue des relations possibles. La littérature sur la croissance des réseaux s'appuie le plus souvent sur des modèles théoriques, depuis de simples grilles régulières jusqu'à des représentations cylindriques ou encore plus complexes (Baptiste, 2003 ; Levinson et Yerra, 2006 ; Xie et Levinson, 2011). C'est dès cette étape que le formalisme du graphe participe à la modélisation d'une situation existante : dans notre cas, une situation avant les premiers rails et les premières locomotives. Le graphe permet donc de matérialiser la diversité spatio-temporelle des villes à partir des nœuds : le nœud est renseigné par la population initiale *(Figure 6. 5. b. et Figure 6. 6. a.)*. Dans la lignée des hypothèses posées dans le *Chapitre 4*, les liens du graphe P représentent les segments théoriques, les liens, entre chaque nœud, caractérisés par un attribut qui traduit le coût nécessaire à sa traversée – son impédance. Dans la situation initiale, cette première phase se caractérise ainsi par la triangulation de Delaunay entre les 36 000 communes, qui, dans le domaine empirique, permet de traduire l'hétérogénéité et l'anisotropie de l'espace, parce que le coût de déplacement piéton traduit un temps de parcours croisé avec la longueur et le dénivelé de chaque lien *(cf. Chapitre 4)*.

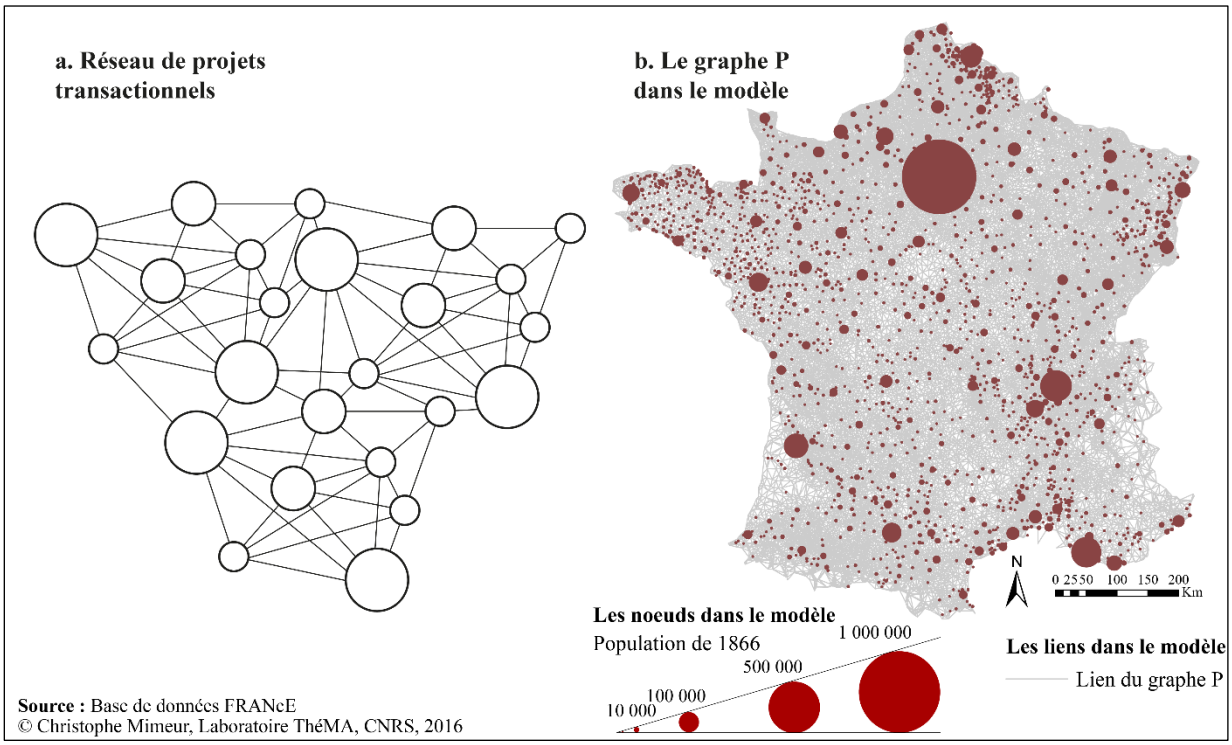


Figure 6. 5. La phase d'initialisation du modèle : entre graphe théorique et empirique

La suite de l'initialisation du modèle consiste en un module d'interaction spatiale : il s'agit du module statique dans le modèle d'évolution du réseau de transport. Cette étape vise alors à vérifier l'hypothèse selon laquelle la croissance du réseau ferroviaire répond à un besoin existant, traduit théoriquement et empiriquement par le réseau de projets transactionnels. Il s'agit de la première étape qui permettra ensuite le passage de l'*« infrastructure »* au *« réseau »* (Dupuy, 1999), par la mise en œuvre d'un modèle d'interaction spatiale. Pratiquement, elle permet d'évaluer les flux potentiels entre chaque paire de nœud, en s'appuyant dans notre cas sur les masses contenues dans chaque nœud du graphe *(Figure 6. 6. b.)*.

Le modèle gravitaire est la traduction physique analogique du postulat posé dans la démarche de la modélisation de l'évolution de réseau de transport : *« deux corps s'attirent en raison directe de leur masse et en raison inverse de la distance qui les sépare »* selon la loi de gravitation universelle de Newton *(Figure 6. 6. b.)*. La masse des communes est alors considérée à la fois comme un potentiel d'émission et comme un potentiel d'attraction *(Figure 6. 6. a.)*. Dans ce modèle, on utilise donc une formalisation très simple du modèle gravitaire, où le potentiel d'échange entre deux communes est donné par :

$$FLUX_{ij} = \frac{P_i . P_j}{d_{ij}^2} \tag{25}$$

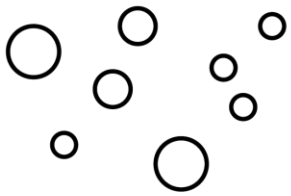
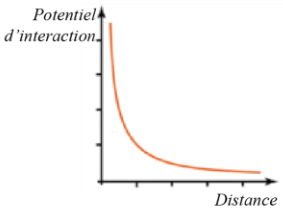
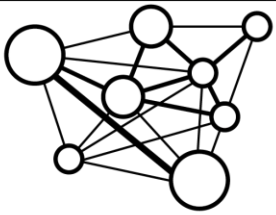
où le flux entre une commune *i* et une commune *j* est donné par le produit de leur population P_i et P_j, rapporté au carré de la distance entre les deux communes. L'exposant est a priori égal à 2, comme il est posé dans le modèle de Newton, et représente l'intensité de la friction occasionnée par la distance *(Figure 6. 6. b.)*.

L'étape ultime de la phase initiale du modèle consiste à répartir les flux estimés par le modèle d'interaction spatiale dans le graphe P (*Figure 6. 6. c.)*. Le formalisme des graphes permet de définir un chemin pour chaque origine et destination, à partir du graphe théorique et triangulé. Dans la mesure où le domaine empirique rend compte d'une structure spatiale hétérogène et anisotrope, la méthode choisie pour affecter les flux est celle du plus court chemin, à partir des attributs des liens – l'impédance. L'algorithme de Dijsktra est le procédé utilisé, comme dans la construction des indicateurs d'accessibilité (Dijkstra, 1959). Ces chemins-là peuvent dès lors être qualifiés de chemins optimaux dont le coût généralisé, comprenant les contraintes physiques et la distance, est le plus efficient. Pour mesurer la fréquence de passage des chemins optimaux dans chaque lien du graphe, la méthode se rapproche du calcul de la centralité d'intermédiarité des liens. Il traduit dans quelle mesure un lien est le passage privilégié de nombreux chemins optimaux. Plus particulièrement, elle sert à évaluer quels liens sont les plus empruntés, quantitativement et qualitativement, en fonction de la génération des origines et des destinations, et des flux estimés à partir du modèle gravitaire *(Figure 6. 6. b.)*. La formule est donnée par :

$$\delta(l) = \sum_{i,j} \sum_{p \in \pi P(i,j)} \left(\frac{1_{l \in p}}{|\pi P(i,j)|} \right) . C(l) \, . FLUX_{ij} \tag{26}$$

où δ est l'intensité du flux qui traverse un lien *l*, qui somme, pour chaque couple d'origine et de destination, $|\pi P(i,j)|$ comme l'ensemble des plus courts chemins entre *i* et *j*, avec *C(l)* l'impédance de ce même lien et $FLUX_{ij}$ le flux estimé par le modèle gravitaire.

Désormais, l'ingénieur économiste peut adopter à partir de cette situation un point de vue cinétique, dont l'objectif est de déterminer des liens sur lesquels vont se concentrer les flux. C'est à cette issue que les revenus potentiels sont évalués, et que le module dynamique du modèle d'évolution du réseau de transport entre dans l'exécution. En effet, si le modèle gravitaire constitue dans notre démarche empirique la meilleure solution pour estimer les flux en fonction des propriétés des nœuds et des liens du graphe, *« c'est un modèle statique, qui ne prend pas en compte l'évolution de la configuration, en particulier celle engendrée par les flux »*[36]. Ainsi, pour faire évoluer la configuration du réseau de transport en fonction de la structuration des flux engendrés, notre modèle dispose d'un module dynamique, qui sera à même de réalimenter le module statique, prenant en compte l'évolution des configurations du réseau.

Phases de l'initialisation	Représentation théorique	Application au modèle d'évolution d'un réseau de transport
a. Hétérogénéité spatio-temporelle		Les communes i sont les noeuds du graphe et sont caractérisées par des masses.
b. Interaction spatiale		***Evaluation des potentiels d'échange*** Le potentiel d'interaction entre deux communes est proportionnel à leurs masses et à l'inverse de leur distance.
c. Une conception maximaliste		***Repartition spatiale des potentiels*** Les chemins optimaux permettent de mesurer la fréquence de passage des flux potentiels.

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Figure 6. 6. Modéliser le besoin existant dans le modèle d'évolution du réseau de transport

[36] PUMAIN Denise, « Modèle gravitaire », *Hypergéo*, Encyclopédie en ligne, www.hypergeo.eu

6.2.3. **Le module dynamique pour générer le réseau de transport**

La modélisation dynamique du réseau de transport consiste en l'amélioration des performances des liens du graphe intégrés dans le modèle. Elle traduit non seulement l'implantation d'une infrastructure mais le renforcement progressif de la vitesse, créant une hiérarchie du réseau ainsi construit. La modélisation classique prend en compte plusieurs configurations de réseau construites de manière exogène (Commenges, 2013 ; Louf, Jensen et Barthelemy, 2013) afin d'en comparer l'efficacité. La démarche est différente ici, dans la mesure où le module consiste à étudier l'évolution des configurations de manière endogène, à partir de la dernière étape décrite dans le module statique. Elle permet d'appréhender sur quels liens le potentiel revenu est important, mais sur lesquels on postule toutefois que les performances ne sont pas suffisantes. C'est le résultat du calcul de $\delta(l)$. L'étude de la part de chaque lien à partir de cet indicateur permet d'identifier les liens sur lesquels il faut investir, c'est-à-dire augmenter la vitesse. Alors que la modélisation classique s'attache à étudier la modification des comportements des individus de manière désagrégée – *agent based* –, notre modèle vise donc à modifier le comportement-même des liens du graphe, on parle alors de modélisation *link based* (Levinson et Yerra, 2006). Le modèle modifie ainsi les attributs du graphe. Pour autant, la dynamique n'affecte que la répartition des flux et non les masses introduites à l'initialisation du modèle, qui sont donc mobilisées à plusieurs reprises.

Une première approximation tend à se rapprocher des comportements classiques rencontrés dans l'économie des transports, à l'échelle de l'agglomération (Commenges, 2013). Elle consiste à systématiquement sélectionner le lien le plus rentable à partir de l'évaluation économique. La procédure effectuée consiste à améliorer l'efficacité de ce lien en réduisant son coût, qui traduit la friction de l'espace. Dans ce cas, la modification est unique, systématique et non cumulative, et constitue une itération dans le modèle (Commenges, 2013). C'est la méthodologie utilisée dans le modèle SERT qu'a développé Hadrien Commenges dans sa thèse, qu'il qualifie comme une boucle de renforcement systématique ou boucle inductive, à l'échelle d'une agglomération.

Si l'approche est ici la même, la procédure de modification du réseau suit une autre approximation. Elle s'inspire davantage des modèles développés par l'équipe américaine (Levinson et Yerra, 2006), dans laquelle nous renforçons le processus endogène. A une échelle plus large, nous posons l'hypothèse que le processus de modification peut affecter plusieurs liens à la fois. Cette hypothèse trouve son sens dans le contexte géohistorique, où la multiplication des initiatives privées d'une part, des grandes compagnies d'autre part, et plus ou moins guidées par l'Etat, sont à l'origine de la forte croissance du réseau sur l'ensemble du territoire dans la seconde moitié du XIXème siècle (Caron, 1997). La modification de l'attribut des liens est donc systématique et la modification peut être cumulative : le renforcement d'un lien peut avoir lieu dans plusieurs itérations du même processus.

L'effet de la modification consiste à modifier l'attribut du lien, c'est-à-dire de diminuer la vitesse de friction. Il s'agit d'un paramètre défini avant le processus de modélisation. Ainsi, l'objectif endogène est de concentrer les flux, dans la dimension cinétique décrite dans la théorie des réseaux. Ce processus paraît de plus adapté à nos échelles d'études dans la mesure où nous ne recherchons pas le maillage le plus complet, mais le plus efficient, dans lequel nous misons sur une inertie temporelle. Cette inertie est directement traduite dans le modèle par la possibilité cumulative de renforcement. La modification de l'attribut de coût d'un lien est donnée par :

$$C_l^{t+1} = \frac{C_l^t(1-\alpha(\delta_l))}{\sum C_L^t(1-\alpha(\delta_L))} \quad (27)$$

où *C* est la nouvelle impédance du lien *l* au temps *t+1*, diminuée du paramètre α défini par le modélisateur, en fonction du coût de chaque lien *l* au temps *t*. Cette valeur est rapportée à la somme des nouveaux coûts de tous les liens *L* au temps *t*.

Si la fin de chaque itération modifie la vitesse sur les liens, alors elle recommence par un détour vers le module statique : les nouvelles vitesses sont intégrées dans une nouvelle répartition des flux. Une nouvelle évaluation a donc lieu à chaque itération qui consiste à répartir les mêmes flux issus de l'estimation par le modèle gravitaire mais cette-fois ci dans une situation où le réseau de transport a été amélioré par l'itération précédente. Ainsi, alors que dans le modèle gravitaire simple, les distances en vol d'oiseau ne permettent pas de différencier les relations entre deux villes qui n'ont pas la même infrastructure de transport, l'approche par les temps de parcours nous permettra de différencier une liaison qui accueille une infrastructure lourde et rapide d'une liaison sans infrastructure, et ce au cours-même du processus de modélisation. Le processus est terminé à l'issue d'un certain nombre d'itérations, qui constitue le second paramètre du modèle. En revanche, aucune contrainte de continuité dans le réseau n'est introduite : pratiquement, il s'agira d'évaluer comment le modèle est capable de créer un réseau connexe ou un réseau dont les tronçons sont disséminés.

L'architecture globale du modèle reprend les trois schémas classiques de la modélisation dynamique *(Figure 6. 3)* : une phase d'initialisation et une phase d'exécution, qui seront suivies d'une phase d'évaluation, à la fois théorique puis empirique. Le modèle lui-même participe à la succession de trois modules *(Figure 6. 7)* :

- Appliquée à la question posée dans ce chapitre, la phase initiale du modèle répond à une structure préexistante dans le domaine empirique ;
- Le module statique participe à l'initialisation du modèle en formalisant les règles de l'efficience économique qui constitue le postulat général du chapitre, dont l'objectif est une analyse coût-bénéfice de tous les liens entre les communes françaises ;
- Le module dynamique formalise le mécanisme de construction de la nouvelle infrastructure, à travers la diminution de sa force de friction, c'est-à-dire la diminution de son coût généralisé par la diminution des vitesses ;

- L'articulation de ces deux modules permet l'évolution endogène de la configuration réseau et la distribution des flux à l'échelle globale, rendant cumulatif le mécanisme d'efficience économique.

Le modèle que nous proposons ici est donc déterministe. La prochaine section vise désormais à étudier le comportement général du modèle pour asseoir par la suite nos interprétations empiriques dans la troisième partie de la thèse.

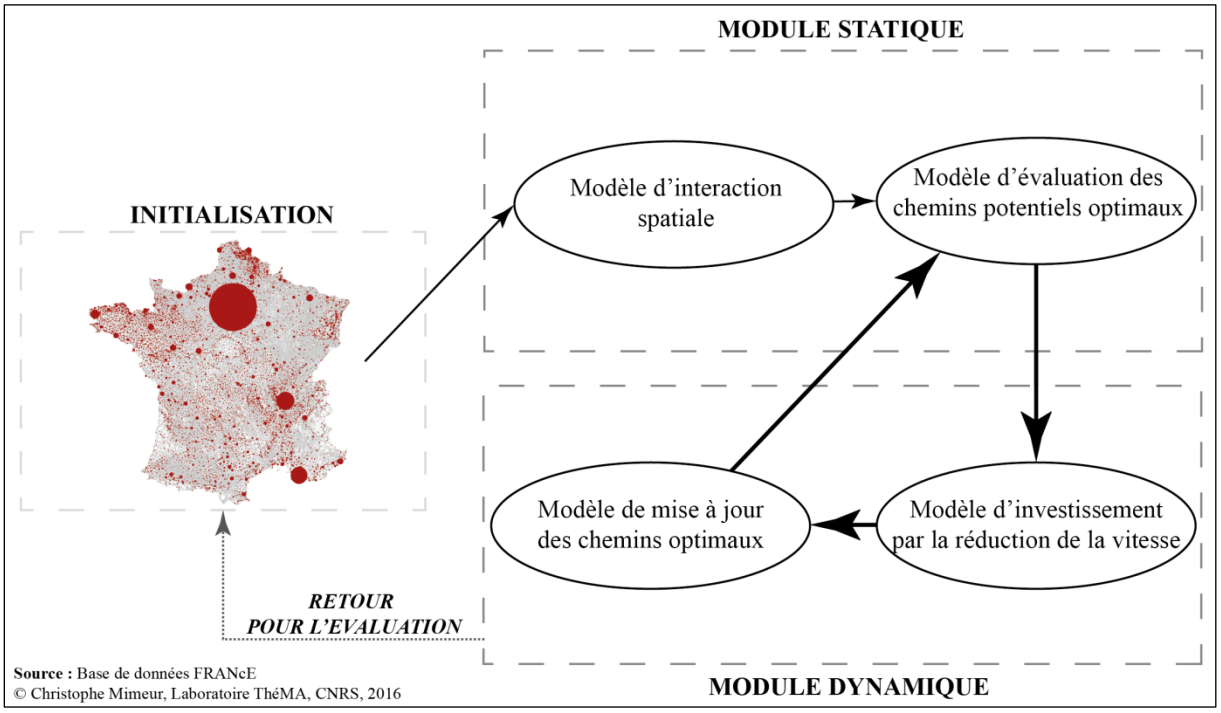


Figure 6. 7. Le modèle d'évolution du réseau de transport : du modèle statique au modèle dynamique

6.3. Explorer le modèle d'évolution de réseau de transport

Cette dernière section a pour objectif d'explorer le modèle d'évolution de réseau de transport à partir de situations stylisées. Un espace stylisé se définit comme un espace théorique, qui ne repose pas sur la représentation d'un espace effectivement rencontré dans la réalité, mais qui se réfère à une forme de représentation spatiale bien connu des géographes. Il s'agit d'étudier le comportement général du modèle, pour qualifier ses mécanismes d'un point de vue algorithmique et statistique bien sûr, mais surtout les phénomènes spatiaux à l'œuvre à partir des espaces stylisés et connus, et d'évaluer dans quelles mesures le modèle permet de reproduire une structure émergente de réseau, qui s'appuie sur une forme et une hiérarchisation (Levinson et Yerra, 2006) interprétables dans le contexte géohistorique de la thèse.

Dans la droite lignée du *Chapitre 4*, l'exécution du modèle participe à la diversification et aux passerelles entre les outils géomatiques *(Figure 4. 2)*. Le modèle est ainsi implémenté grâce au package *igraph* du logiciel *R*. L'interfaçage avec les librairies spatiales permet d'étudier les sorties des modèles théoriques directement à partir du logiciel, alors que l'ensemble des résultats peut tout autant être transféré vers un outil SIG. Le modèle d'évolution du réseau de transport consiste pratiquement à l'exécution d'un script sous le logiciel *R* présenté en *Annexe 2*. La première étape consiste en l'initialisation du modèle. A partir d'informations spatiales issues d'un logiciel SIG ou de l'architecture d'un SGBD, il s'agit de créer un objet graphe, reconnu par *igraph*. Cette transformation permet d'alléger l'information qui ne se résume plus qu'à des liens et des nœuds. La géométrie des liens n'est alors plus requise. L'initialisation consiste ensuite à préciser quels champs permettent de traduire la position des nœuds, leur masse, ainsi que les attributs des liens – l'impédance dans notre cas. A partir de là, une représentation graphique ou cartographique du graphe théorique initiale est possible.

Par la suite, le script consiste en la définition des fonctions permettant d'exécuter le modèle d'interaction spatiale, le modèle d'évaluation des flux qui en résultent dans le graphe P, le modèle d'investissement qui consiste à renforcer les liens du graphe. Une fonction finale permet d'exécuter l'ensemble de ces modules à l'intérieur d'une boucle qui permet l'exécution successive de chacun d'entre eux à chaque itération, permettant de cette manière la mise à jour de l'évaluation des flux dans la nouvelle configuration. La durée d'exécution du modèle est essentiellement fonction de la taille du graphe à l'initialisation et du nombre d'itérations choisi par le modélisateur[37].

Ce script est hybride : il nous permet dans cette section de modéliser des espaces stylisés mais il est tout aussi adapté à l'exécution du modèle à partir de graphes issus de la base de données FRANcE. Dans la chaîne méthodologique de la modélisation dynamique, la phase d'évaluation est l'ultime étape qui contribue à la construction de connaissances *(Figure 6. 3)*.

6.3.1. **Une exploration théorique du modèle par des espaces stylisés**

La plupart des approches ont d'abord en commun d'utiliser l'interprétation visuelle pour évaluer la structure du réseau généré par un modèle. Le premier objectif est donc d'observer la forme générale du réseau, sa capacité à reproduire des situations viables dans un contexte réel. Cela passe souvent par la comparaison avec des réseaux modélisés graphiquement, tels les réseaux maillés, en arbre ou en étoile (Bavoux et al., 2005), en gardant à l'esprit la démarche instrumentale du modèle. Ce procédé répond à la démarche constructiviste et abductive du travail, où notre connaissance du contexte géohistorique et des théories d'arrière-plan participe autant à l'évaluation du modèle que les indicateurs statistiques. Ainsi, la nature du modèle ne requiert pas une confrontation systématique d'individu à individu (Durand-Dastès, 2001).

Le second objectif à poursuivre est l'étude des capacités circulatoires du réseau de transport généré. Les propriétés issues du modèle gravitaire questionnent ces aspects, au moins autant que la

[37] Les capacités du logiciel R ne permettent pas pour l'heure la prise en compte des 36 000 communes dans le processus. En revanche, elles parviennent à prendre en compte l'ensemble des liens du graphe triangulé.

forme générale du réseau. On interroge de cette sorte la capacité du modèle à produire une hiérarchisation du réseau simulé. Cet objectif se trouve justifié par le fait qu'*a priori,* tous les liens sont potentiellement affectés par le processus de génération de manière cumulative. C'est donc bien ce processus de hiérarchisation qui permet de traduire celui de la sélection de liens au détriment d'autres pour générer le réseau de transport. Dans ce cas, les indicateurs de hiérarchie, qui traduisent la distribution des flux et des vitesses (Levinson et Yerra, 2006), sont à mobiliser dans l'expérimentation et l'évaluation du modèle. La démarche consiste à confronter des structures, et surtout leur distribution (Durand-Dastès, 2001), dont l'interprétation visuelle sera un complément.

L'exploration théorique consiste à étudier le comportement général du modèle et la sensibilité des deux paramètres du modèle. L'objectif n'est pas de définir des valeurs arrêtées des paramètres mais d'appréhender leur impact, ce qui amènera dans une étape ultérieure à la sélection de paramètres en fonction de ce que le modélisateur souhaite montrer. Ils sont :

- Le **paramètre α** défini dans le modèle d'investissement permet de réduire le coût nécessaire pour franchir un lien par la réduction de sa vitesse. La normalisation systématique des coûts permet au modélisateur de faire varier α entre 0 et 1. L'objectif est de voir comment le modèle réagit selon que le différentiel de vitesse appliqué est faible ou fort par rapport au coût initial ;
- Le second paramètre est le **nombre d'itérations** à effectuer au sein du même processus. L'objectif est d'étudier le comportement du modèle dans le processus de sélection d'axes sur lesquels l'investissement doit se concentrer et comment l'évolution du réseau de transport peut atteindre une situation stabilisée. Il s'agira alors de qualifier quelle est la distribution des flux et des vitesses de cette forme stabilisée et quelle interprétation géohistorique il peut en être fait.

Pour cela, nous nous appuyons sur la mise en place d'une grille à partir de situations stylisées. Les trois modèles utilisés reprennent les trois objets de l'initialisation : des nœuds qui matérialisent des communes, des liens qui représentent les liaisons piétonnes existantes avant l'apparition d'un nouveau réseau. Nous appliquons une grille carrée qui n'est pas circulaire, au sein de laquelle les relations sont matérialisées par un voisinage de Von Neumann[38]. Trois types d'espaces sont modélisés dans le plan d'expérimentation :

- Un espace qualifié d'homogène, où la distribution des communes est uniforme à l'intersection de chaque lien *(Figure 6. 8. a.)* ;
- Un espace monocentrique, où la distribution des populations suit un gradient centre/périphérie *(Figure 6. 8. b.)* ;
- Un espace polycentrique, où la distribution des populations suit un tirage aléatoire dans une distribution logarithmique *(Figure 6. 8. c.)*.

[38] Une telle formalisation signifie que l'espace est finie et est fixé par les limites de la grille.

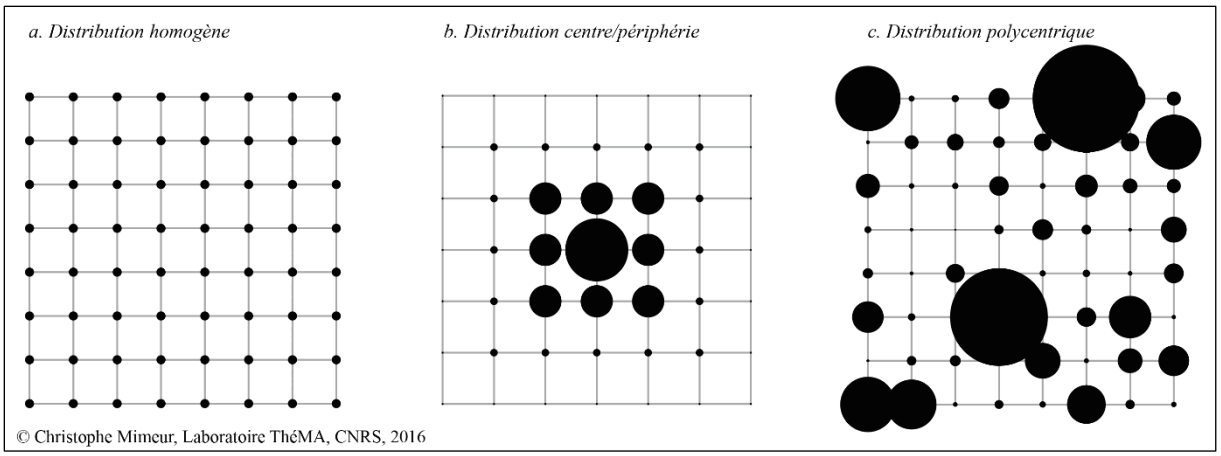


Figure 6. 8. Trois espaces stylisés pour l'expérimentation du modèle

Les liens ont dans tous les cas le même coût au moment de l'initialisation, de manière à pouvoir comparer les résultats à la fin de chaque processus. L'étude de situations où l'espace est anisotrope est laissée au domaine empirique, dans la troisième partie de la thèse. Dans l'expérimentation, la distribution théorique des plus courts chemins suit donc un gradient centre/périphérie, et est calculée à l'aide de la centralité d'intermédiarité. L'objectif est notamment d'étudier comment les configurations spatiales des trois approches stylisées viennent modifier ce gradient dans l'évolution endogène du réseau de transport.

6.3.2. De la structure à la hiérarchie de l'évolution du réseau de transport dans les espaces stylisés

En confrontant les structures théoriques au modèle d'évolution du réseau de transport, nous suivons le même protocole d'analyse dans les trois cas. La première étape consiste à comparer la distribution des flux telle que le réseau initial le permet – qui ne tient donc compte que des liens – , ainsi que la distribution des flux issue de l'application du modèle gravitaire – qui tient compte à la fois des liens et des nœuds du graphe. La confrontation de ces deux structures permet l'exécution du modèle en modifiant les liens du réseau initial. Dans un premier temps, le paramètre α est fixé à 0,5 : la vitesse est divisée par deux, et c'est une valeur souvent utilisée dans la littérature (Commenges, 2013 ; Levinson et Yerra, 2006). Par la suite, nous utiliserons des valeurs extrêmes $\alpha = 0,1$ et $\alpha = 0,9$ pour tester le comportement du modèle quand il provoque des différentiels très forts ou très faibles à chaque itération. Enfin, la démarche sera exploratoire pour examiner d'autres valeurs intermédiaires. Pour chaque valeur de α, nous exécutons le modèle avec 2, 4, 6, 8, 10, 20, 30, 40, 50 et 100 itérations. Nous faisons le choix de ne pas utiliser des indicateurs de structures de graphes, dans la mesure où il a été démontré que pour deux réseaux sensiblement différents, surtout par leur hiérarchie, ces indicateurs peuvent présenter des indicateurs semblables (Dupuy et Stransky, 1996 ; Pumain et Saint-Julien, 2001).

a. *Le modèle face à une distribution homogène*

Dans le cas d'une distribution homogène des communes, la distribution des flux à partir des liens d'une part, à partir des nœuds par le modèle gravitaire d'autre part, est semblable dans la mesure où les deux objets du graphe sont homogènes initialement *(Figure 6. 9)*. Lors de l'exécution du modèle, si deux chemins ont une valeur identique, alors il choisit de manière aléatoire le chemin sur lequel l'évaluation économique aboutit à une modification des vitesses.

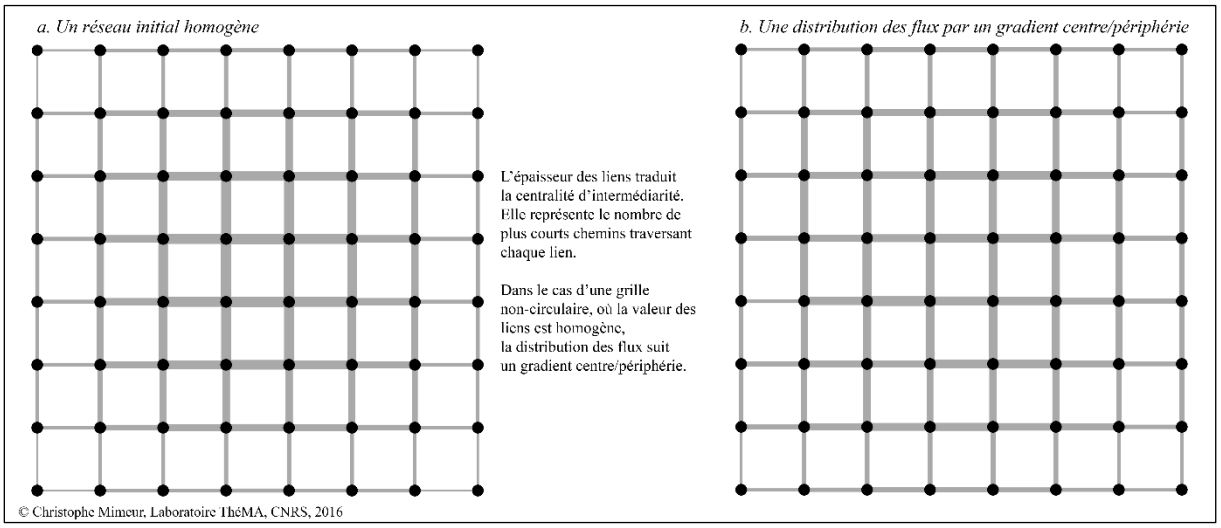


Figure 6. 9. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution homogène

La *Figure 6. 10* montre l'évolution du réseau de transport entre 2 et 20 itérations. Après 2 itérations, c'est à partir du centre que les liens sont prioritairement améliorés. Pour autant, tous les liens sont modifiés. La distribution montre bien ces deux liens choisis et un grand nombre d'autres modifiés dans une moindre mesure. Après 4 itérations, des premières sélections apparaissent : le flux se concentre sur certains liens, faisant diminuer l'importance d'autres, surtout en périphérie. Après 20 itérations, la situation est stabilisée. Nous avons sélectionné ici les 25 liens qui concentrent le plus de flux, afin d'en étudier la structure. Ils dessinent un réseau qui semble partir d'un point central, d'où partent 4 branches. Les nombreux chemins *ex-aequo* influent sur la distribution aléatoire des axes les plus structurants. Cependant la structure du réseau est connexe et une première interprétation visuelle ne montre pas de rupture incohérente dans la hiérarchie mise en place dans le réseau généré.

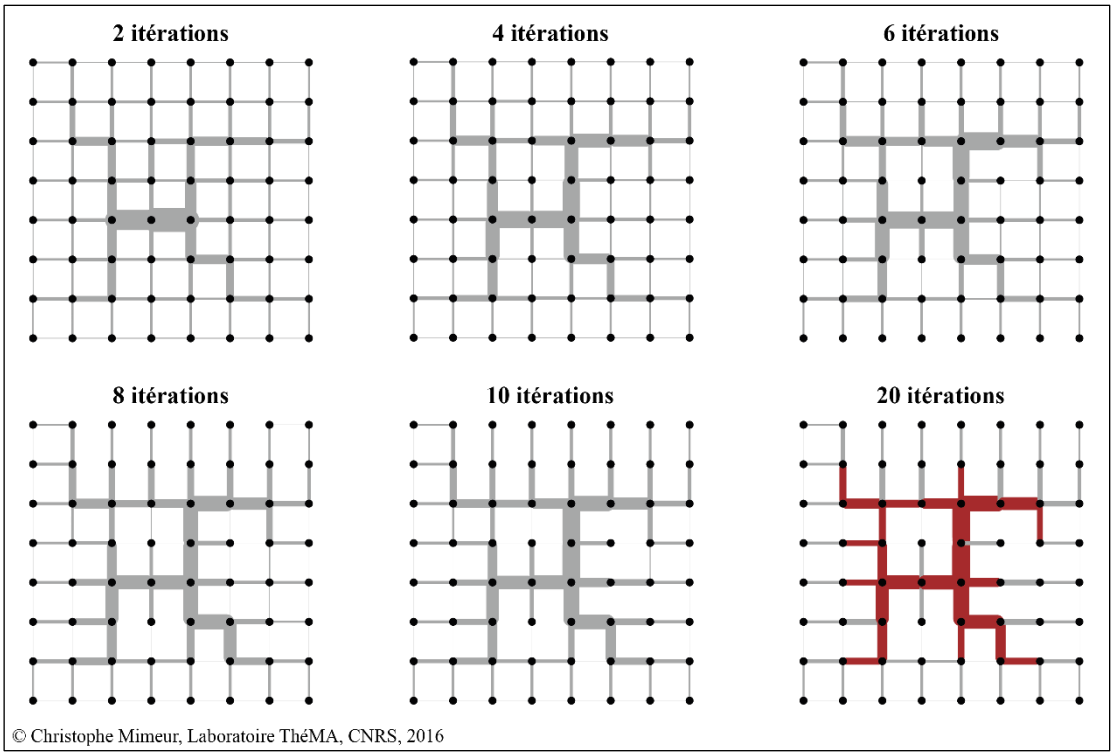


Figure 6. 10. Génération du réseau de transport avec α = 0,5

Il s'agit aussi d'analyser la forme de la distribution des flux à mesure que le réseau est généré *(Figure 6. 11)*. L'axe des abscisses mentionne le nombre de plus courts chemins tandis que l'axe des ordonnées traduit leur fréquence. La distribution des flux semble correspondre à une structure hiérarchique : forcément, nous optons pour une minimalisation des tronçons à construire pour le réseau ferroviaire. La situation est stabilisée après 20 itérations : c'est cette distribution qui nous intéresse plus particulièrement. Auparavant, on voit que le processus de sélection n'opère pas encore : les valeurs élevées sont encore homogènes et concernent de nombreux liens. Après 8 itérations, une rupture s'opère et c'est à ce moment-là que la forme du réseau, interprétée de manière visuelle, semble se stabiliser. Enfin, entre 10 et 20 itérations, une nouvelle sélection s'opère avec la chute des valeurs intermédiaires. Dans la distribution qui nous intéresse, après 20 itérations, une rupture se dessine entre des liens améliorés en faible proportion et sont nombreux, d'autres sont améliorés de manière conséquente et sont plus rares. On place la limite à 30, à l'endroit où la dernière sélection est la plus probante :

- Un premier ensemble de liens, qui sont nombreux et proposent des améliorations modérées : dans notre contexte géohistorique, nous pouvons les rapprocher des voies dont l'objectif est la desserte locale ;
- D'autres, presque aussi nombreux, proposent des améliorations plus importantes : ce sont des voies intermédiaires ;
- Certains proposent des améliorations conséquentes mais sont en nombre réduit : ce sont des voies rapides, dont le coût d'investissement est élevé.

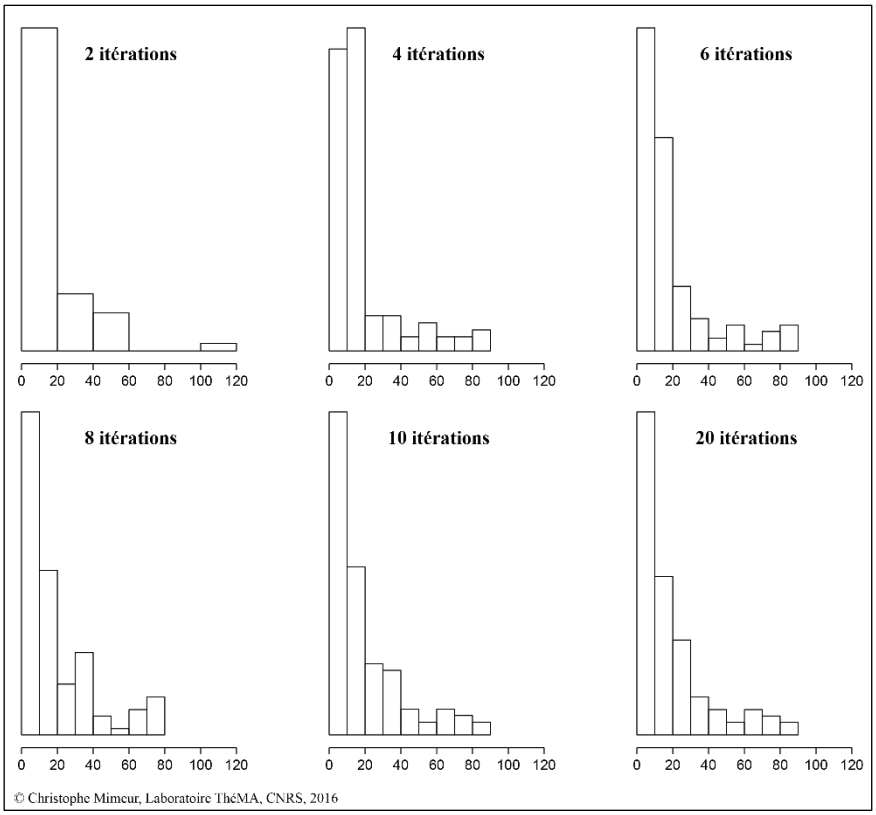


Figure 6. 11. Distribution des flux selon le nombre d'itérations avec α = 0,5

D'autres valeurs de α permettent d'étudier le comportement du modèle dans les processus de sélection et de hiérarchisation. Avec le paramètre α = 0,1, on remarque que tous les tronçons sont concernés par l'amélioration du réseau après 8 itérations *(Figure 6. 11. a.)*. Les différentiels de vitesse ne sont pas assez marqués pour qu'une sélection de liens se fasse. L'inversion de la distribution vers les axes les plus forts n'apparaît qu'au bout de 20 itérations, même si tous les tronçons restent encore impactés jusqu'à 30 itérations. La sélection n'est pas synonyme de hiérarchisation, car les différentiels ne sont pas encore assez forts. Le processus de hiérarchisation ne se renforce qu'après 40 itérations, même si la rupture entre les axes modifiés et ceux qui ne le sont pas n'est pas encore marquée. Il faut attendre 50 itérations pour qu'un véritable processus de sélection doublé d'une hiérarchisation opère et aboutisse à une situation équilibrée similaire à celle obtenue avec α = 0,5.

Avec le paramètre α = 0,9, le processus observé est quasiment inversé. Les différentiels de vitesse imposés sont grands. Deux groupes se distinguent : les tronçons très peu affectés, et les tronçons très affectés. Dès 4 itérations, le processus laisse de côté certains tronçons : la sélection opère, mais pas encore la hiérarchisation *(Figure 6. 11. b.)*. Ici, encore davantage que pour des valeurs inférieures, le caractère *ex-aequo* pousse à la sélection aléatoire de chemins au cours du même processus. On retient ici que la sélection des liens à améliorer est plus forte, même si la forme du réseau obtenu laisse apparaître des discontinuités dans la distribution hiérarchique. Pour information, des valeurs intermédiaires ont été testées : entre 0,5 et 0,8, le réseau se stabilise de la même manière que pour α = 0,5, en marquant la hiérarchie de manière plus forte.

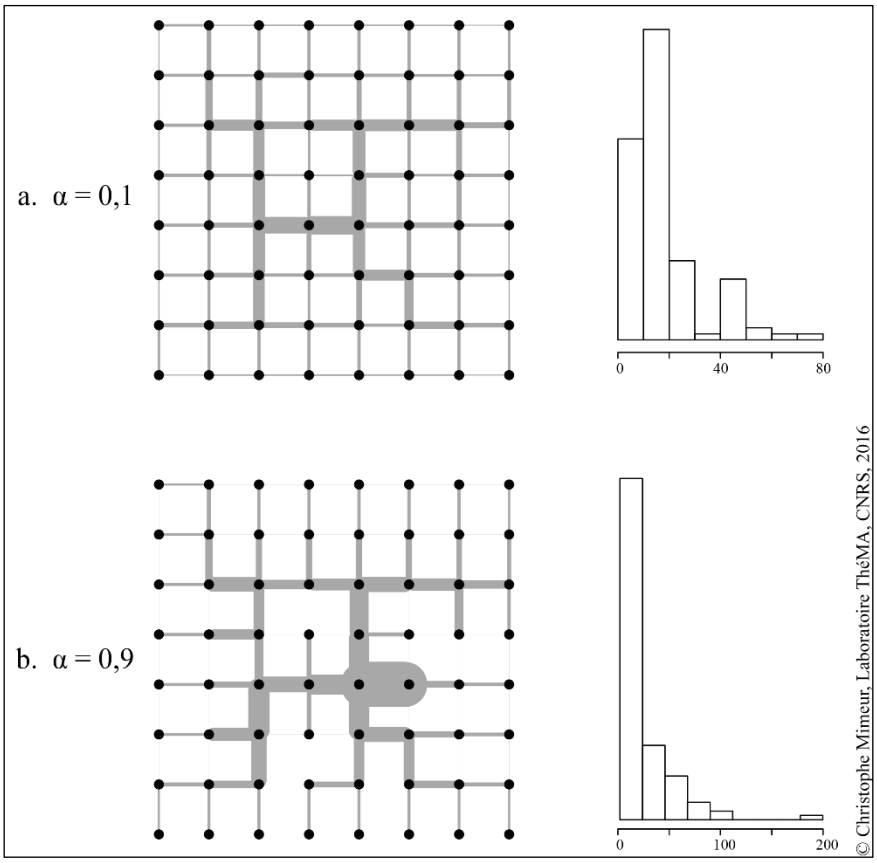


Figure 6. 12. Génération du réseau de transport après 8 itérations

De manière globale, cette première expérimentation montre que le modèle d'investissement participe à la connexité du réseau construit, ce qui constitue un point positif pour la poursuite de nos investigations. A mesure que l'on augmente le paramètre α, la sélection est opérée de plus en plus tôt, l'augmentation du paramètre d'itérations participe quant à elle à la hiérarchisation de la distribution. Si les différentiels sont petits, la sélection est tardive et la hiérarchisation n'est pas prononcée. S'ils sont grands, la sélection est rapide mais la hiérarchisation montre une large part d'aléatoire. Il s'agit désormais d'étudier le comportement du modèle dans d'autres espaces stylisés, encore mieux connu du géographe.

b. *Le modèle face à une distribution centre/périphérie*

A la différence de l'exemple précédent, la structure des communes, qui suit un gradient centre/périphérie, vient renforcer ce même gradient dans la répartition des flux issue du modèle gravitaire. La forme entre la distribution des flux selon les liens *(Figure 6. 13. a.)* et selon les nœuds est la même, mais accentuée dans le second cas *(Figure 6. 13. b.)*. Quel impact cette confrontation a-t-elle dans l'exécution du modèle de génération du transport ?

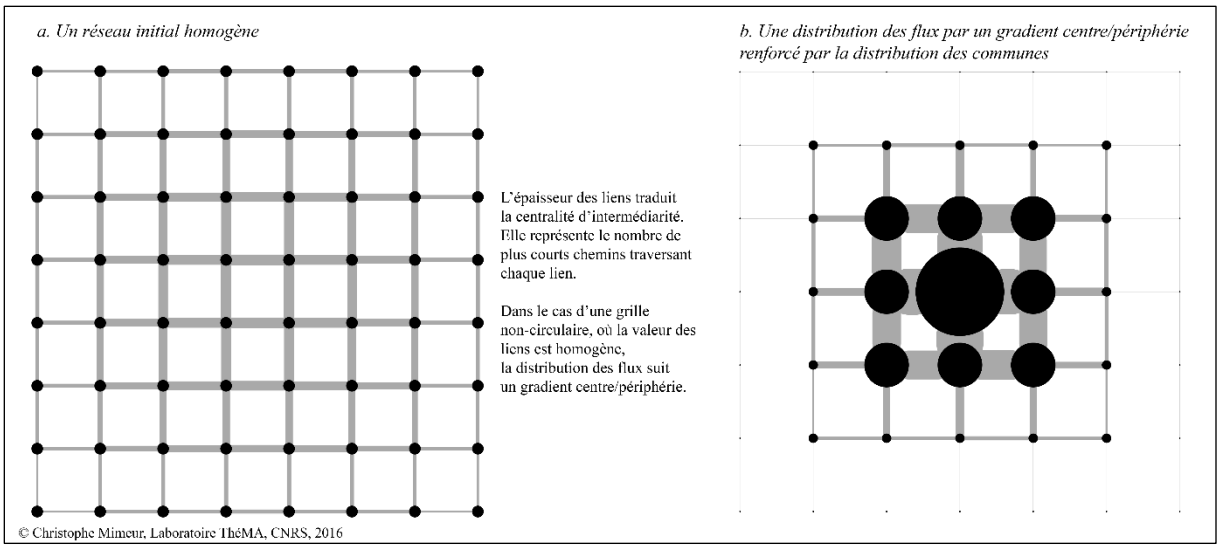


Figure 6. 13. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution centre/périphérie

Nous ne revenons pas ici en détail sur chacune des configurations avec la valeur centrale du paramètre α. On se contente de confronter la forme obtenue après un équilibre obtenu à la suite de 8 itérations *(Figure 6. 14)* pour α = 0,5. Avant cet équilibre, seuls les liens entre le nœud central et la première couronne ainsi qu'entre ces nœuds périphériques sont améliorés : c'est ce que le modèle gravitaire suggère dans une certaine mesure. La génération du modèle de transport opère une sélection après 8 itérations, où seulement 50 % des liens concernés sont améliorés dans le processus. Le réseau crée une nouvelle hiérarchie dans les centres secondaires, où certains deviennent centraux pour en rejoindre d'autres, comme c'est le cas du centre voisin du nœud principal au sud de ce dernier. Alors cette sélection permet l'amélioration de liens vers des centres périphériques, qui étaient pourtant mis de côté au départ. La distribution correspondant à cette situation stabilisée nous éclaire davantage sur le mécanisme du modèle *(Figure 6. 14)* :

- Seule une minorité de tronçons est affectée par le modèle : la distribution reprend l'élément de séparation spatiale, principe de l'accessibilité géohistorique, qui fait une claire rupture entre un tronçon doté d'une infrastructure et celui qui ne l'est pas ;
- La hiérarchisation dans les tronçons améliorés semble reproduire les contraintes d'investissement dans la construction d'une infrastructure : les liens les plus améliorés, les plus rapides, sont alors les moins nombreux, parce que les plus chers.

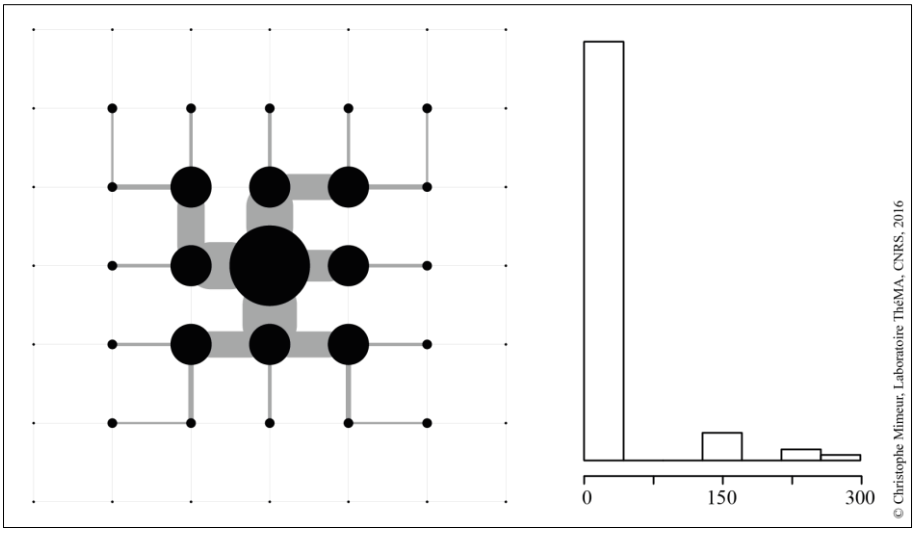


Figure 6. 14. Génération du réseau de transport dans un espace monocentrique après 8 itérations avec α = 0,5

Il s'agit désormais de confronter cet espace stylisé aux valeurs extrêmes du paramètre α. Ainsi pour α = 0,1, le modèle n'opère pas une sélection, parce qu'encore une fois, le différentiel de vitesse n'est pas assez fort : il se comporte en suiveur des flux imposés par le modèle gravitaire. Ce processus s'entretient jusqu'à 20 itérations : s'il y a quelques liens fortement améliorés, la distribution montre que les autres gains ne sont que minimes et ne marque pas de rupture avec les liens qui ne sont pas du tout impactés. C'est seulement à partir de là que l'accumulation des améliorations sur le réseau entraîne un phénomène de sélection *(Figure 6. 15. a.)*. Cette sélection s'opère par une suppression progressive d'arcs intermédiaires, dont le résultat aboutit à une hiérarchisation des axes les plus forts dans un dernier temps, après 50 itérations.

Avec α = 0,9, le processus de sélection est immédiat et concerne les tronçons adjacents au nœud central. Les itérations suivantes cherchent ensuite à choisir des liens de manière à rejoindre les centres périphériques, sans à aucun moment que la dernière couronne ne bénéficie d'une infrastructure. Le processus vise ensuite au renforcement de quelques axes, afin de pouvoir reporter des investissements pour rejoindre les centres les plus périphériques. Toutefois après 20 itérations, la structure visuelle du réseau généré montre, certes une connexité, mais un déficit de continuité dans la succession hiérarchique des liens *(Figure 6. 15. b.)*. Ainsi, alors que deux axes forts se dessinent entre les centres secondaires, les performances chutent d'un tronçon à l'autre, pour rejoindre le nœud central d'une part, pour rejoindre les nœuds les plus périphériques d'autre part. On ne constate donc pas de réelle rupture entre les liens améliorés et ceux qui ne le sont pas.

De ce modèle centre/périphérie, le comportement du modèle en fonction de α conforte les conclusions apportées dans le modèle homogène, renforcé par une structure préexistante. Le modèle semble ainsi d'abord opérer un processus de sélection à partir duquel il fait émerger une hiérarchie dans les liens les plus rentables à l'investissement d'une infrastructure.

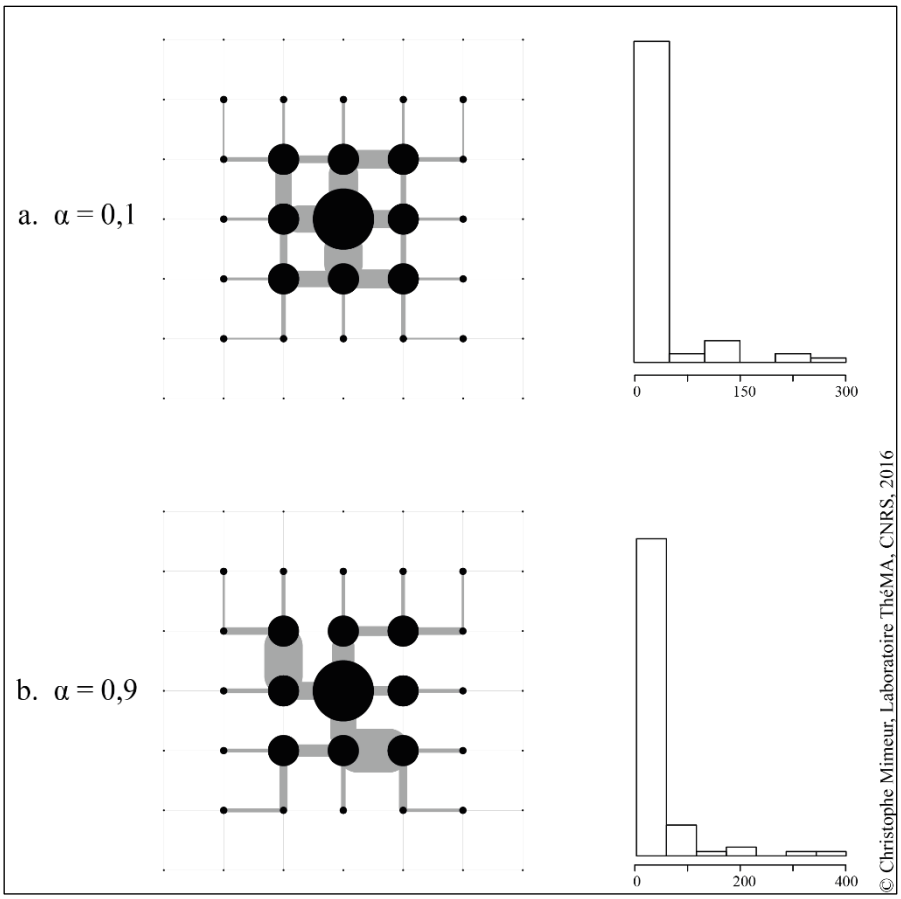


Figure 6. 15. Génération du réseau de transport dans un espace monocentrique après 20 itérations

c. *Le modèle face à une distribution polycentrique*

La dernière proposition de configuration théorique est un espace polycentrique. Dans la mesure où la distribution des communes suit une loi logarithmique, la répartition voit 2 centres principaux, dont l'un est au cœur de l'espace étudié, et l'autre sur la bordure nord. Environ 5 centres intermédiaires sont répartis sur les bords de l'espace et légèrement au centre. De petites villes jusqu'à de petits bourgs sont répartis sur le reste de l'espace. La répartition des flux selon le modèle gravitaire montre une relation à privilégier entre les deux centres *(Figure 6. 16. b.)*. Pour autant, des axes vers les centres secondaires se dessinent également, et les chemins les plus courts passent parfois par de petits villages. Les règles d'efficience économique conservent-elles cette forme de réseau, loin du gradient centre/périphérie imposé par la forme initiale *(Figure 6. 16. a.)* ?

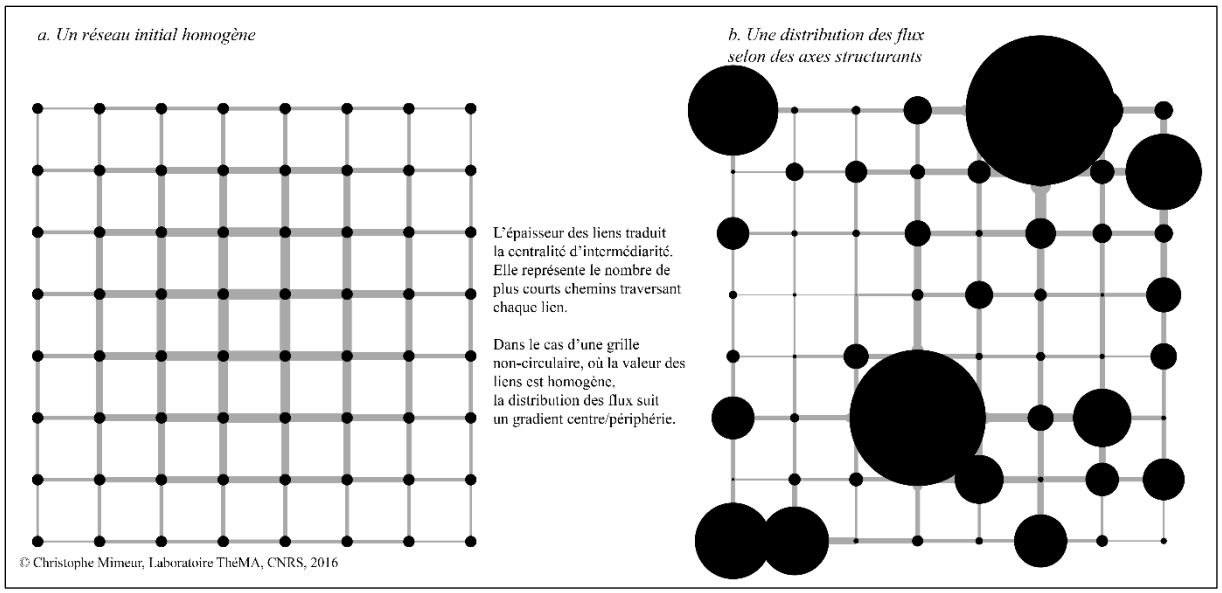


Figure 6. 16. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution polycentrique

Comme dans les deux cas précédents, on s'intéresse d'abord à la forme du réseau créé avec $\alpha = 0,5$. Le début du processus tend à générer un réseau proche de la distribution issue du modèle gravitaire. Le mécanisme consiste alors à rejoindre tous les points les plus forts. Si l'axe rejoignant les deux principaux centres paraît le plus prioritaire, alors le modèle tend à le privilégier et à le renforcer au fur et à mesure. Ainsi, après 10 itérations, cet axe est le plus fort *(Figure 6. 17)* et concentre le plus de flux. La sélection concerne les axes ayant une valeur supérieure à 20. A partir de ce lien principal, on observe la création de ramifications vers des centres secondaires : elle participe à l'établissement de carrefours, parfois avec 4 branches autour d'un même centre pourtant assez petit. Dans ce cas, on mise sur un effet structurant de l'infrastructure, parce que la taille n'induit pas a priori de faire de ce centre un carrefour important. La structure du réseau généré semble répondre à un réseau en arbre, en opérant un processus de sélection par rapport à la distribution issue du modèle gravitaire. La sélection privilégie les centres, quitte à en exclure certains, comme c'est le cas d'un centre périphérique au nord-ouest de l'espace. Après 10 itérations, on observe donc un réseau dont le nombre d'axes est minimisé.

Après 10 itérations, la distribution semble répondre aux critères déjà évoqués dans les autres cas : on observe une nette rupture entre les liens qui ne bénéficient d'aucune amélioration et les autres *(Figure 6. 17)*. Le processus de hiérarchisation est en revanche ici plus difficile à saisir. Certes, l'axe principal entre les deux principaux centres semble se détacher des autres. Pour autant, il est moins évident de dégager d'autres éléments hiérarchiques : des liens intermédiaires rejoignent des centres secondaires qui restent à proximité de l'axe central. En dehors, ce sont des liens dont l'amélioration est faible qui sont les plus nombreux, mais dont l'objectif est de rejoindre les centres les plus éloignés. Cet état de fait entre ainsi en cohérence avec le modèle comportemental observé dans les chapitres précédents.

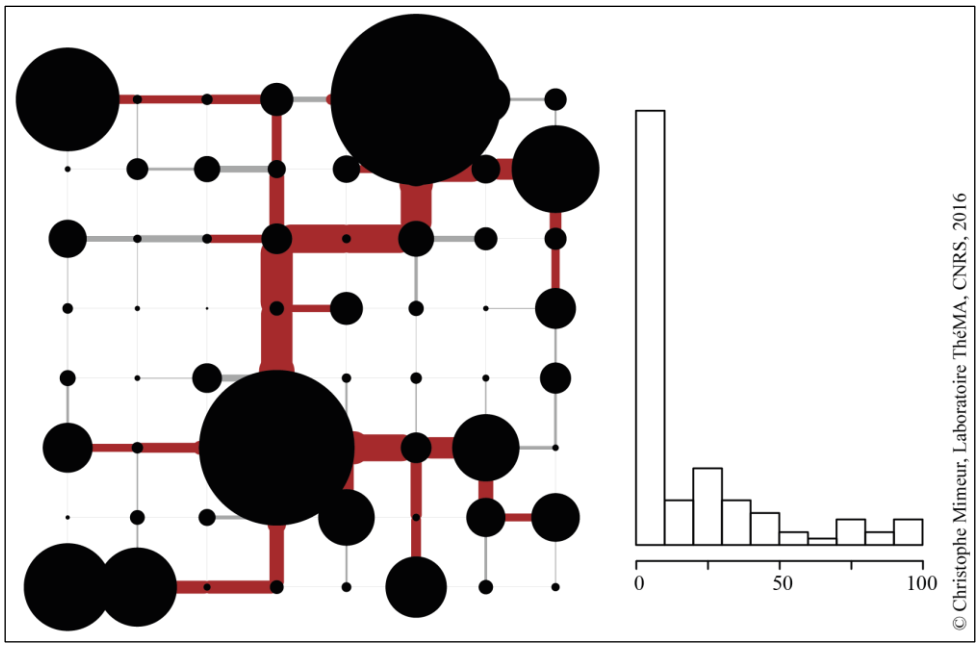


Figure 6. 17. Génération du réseau de transport dans un espace polycentrique après 10 itérations avec α = 0,5

Pour mieux comprendre les processus conjoints de sélection et de hiérarchisation, on utilise dans un premier temps le paramètre α = 0,1 *(Figure 6. 18. a.)*.La structure du réseau généré est cohérente. L'axe entre les deux centres les plus forts se dessine, même si le tronçon le plus amélioré est unique et est adjacent au centre le plus important au nord de l'espace. Pendant longtemps au cours du processus, deux axes continuent de se faire concurrence : d'orientation Nord-Sud, ils cherchent tous les deux à rejoindre les deux points les plus forts, tout en assurant la desserte de centres intermédiaires. Le faible différentiel de vitesse imposé dans le modèle d'investissement ne permet pas au modèle d'opérer un choix[39]. Assez naturellement, la distribution issue n'est pas satisfaisante. On n'observe aucune rupture entre les tronçons qui ne sont pas touchés par l'amélioration et les autres. De plus, le processus de hiérarchisation n'opère que pour quelques liens avec un impact fort.

Pour α = 0,9, le processus de sélection est encore une fois immédiat *(Figure 6. 18. b.)*. L'axe central ressort forcément, et de manière forte, délaissant toute tentative de concurrence observée pour α = 0,1. Si on observe une rupture nette entre les liens sur lesquels portent l'investissement et ceux qui n'en profitent pas, le processus de hiérarchisation à l'œuvre est plus difficile à saisir. Après 10 itérations, la structure du réseau généré paraît cohérente, en dépit de quelques discontinuités dans le réseau. Par exemple, un tronçon existe à l'ouest du nœud le plus fort au nord de l'espace, mais n'est pas suivi des faits par le dessin d'un axe vers l'ouest. La suite du processus, après davantage d'itérations, est le fruit d'une sélection tout aussi intense, mais la distribution des liens améliorés est bouleversée, participant à la création de situations incohérentes, en créant des ruptures de connexité.

[39] Une série de tests a permis de voir que la sélection ne s'opère qu'après 250 itérations, quand l'accumulation des petites améliorations a été suffisante pour faire un choix.

Le modèle a été mis à l'épreuve d'une structure polycentrique, nous permettant d'étudier les comportements spatiaux du modèle. Les conclusions viennent confirmer les mécanismes observés plus haut : le paramètre α participe au phénomène de sélection et de discrimination tandis que le processus itératif participe davantage à l'émergence d'une hiérarchie.

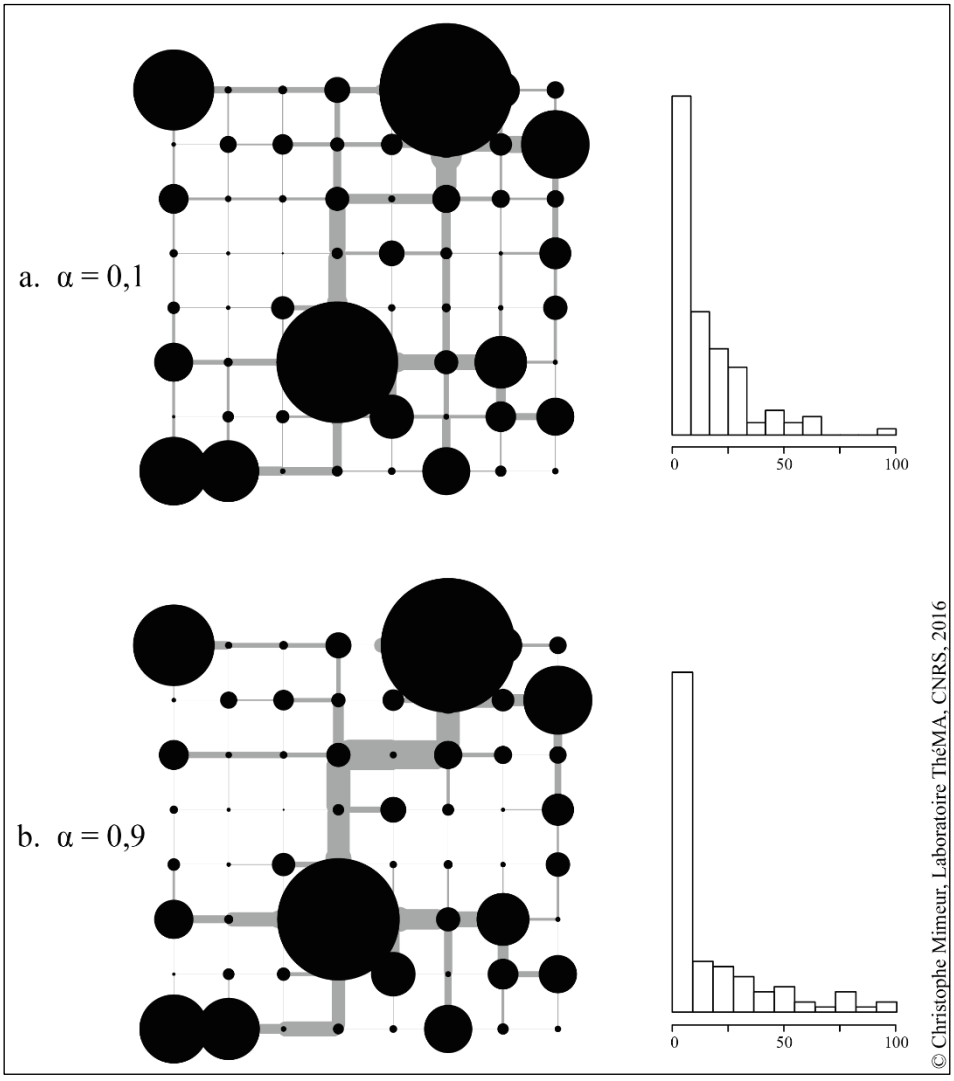


Figure 6. 18. Génération du réseau de transport dans un espace polycentrique après 10 itérations

6.3.3. **Vers une définition des paramètres pour une exécution empirique du modèle**

L'objectif est désormais de tester le modèle sur une base empirique. Les règles formalisées dans les modules statiques et dynamiques sont-elles en mesure de reproduire le processus de croissance du réseau ferroviaire français ? Dans ce cadre, nous avons vu comment l'intérêt ne réside pas forcément dans la fixation de valeurs des paramètres *a priori*, mais est plutôt de voir dans quelle mesure ces paramètres modifient le réseau généré.

Le modèle d'évolution du réseau de transport vise à donc à reproduire une certaine famille de règles élémentaires que l'on suspecte d'avoir un rôle important dans la phase de croissance du réseau : en quoi la prise en compte d'une structure démographique préexistante participe à la structure et à la hiérarchie du réseau en privilégiant les nœuds les plus forts et en concentrant les flux potentiels selon la méthode de la branche commune. La *Figure 6. 19* permet ainsi d'exposer les trois aspects à partir desquels vont se construire la réflexion dans le domaine empirique. Les écarts entre un réseau simulé et un réseau réel sont à lire comme porteurs d'autres facteurs que ceux pris en compte dans le processus de modélisation, comme l'influence du contexte local ou la force d'une décision politique.

Dès lors, l'initialisation du modèle paraît fondamentale : à quel degré cette structure importe-t-elle le plus dans l'architecture globale du réseau de transport ? S'appuie-t-elle sur l'ensemble de la hiérarchie des communes françaises ou se contente-t-elle de ne s'appuyer que les premiers degrés de cette hiérarchie ? Ainsi, une confrontation de plusieurs situations initiales, plusieurs réseaux de projets transactionnels, permet de répondre à cette question *(Figure 6. 19)*.

Si les règles et hypothèses laissent à penser à un laissez-faire, la main du modélisateur est à envisager comme celle d'un planificateur qui répond à un besoin existant selon les règles qu'il s'impose lui-même. Le paramètre α permet donc d'explorer dans quelle mesure la génération du réseau ferroviaire participe à un processus de sélection d'axes structurants le long desquels les investissements doivent se concentrer : de forts différentiels de vitesse laissent penser à une volonté de concentrer les flux, tandis que de faibles valeurs de α tendent à préférer l'expansion du réseau au détriment de la performance circulatoire *(Figure 6. 19)*. Dans l'exécution empirique du modèle, nous aurons ainsi à définir l'intensité du paramètre : faible à 0,1, modéré entre 0,3 et 0,7 et fort à 0,9.

Faire varier le nombre d'itérations dans le modèle questionne le processus de hiérarchisation à l'œuvre sur le temps long, avec l'inertie des réseaux que cela occasionne *(Figure 6. 19)*. En quoi la durée du processus influence-t-elle la mise en place d'une hiérarchie à l'intérieur de la même infrastructure ? Comment ce processus hiérarchique se conjugue avec les choix d'investissement ? Le choix étant déterminant dans la durée d'exécution du modèle, un choix entre 2 et 20 itérations a été fait pour chacune des valeurs du paramètre d'investissement. Par ailleurs, d'autres explorations ont montré une stabilisation du modèle entre 15 et 20 itérations.

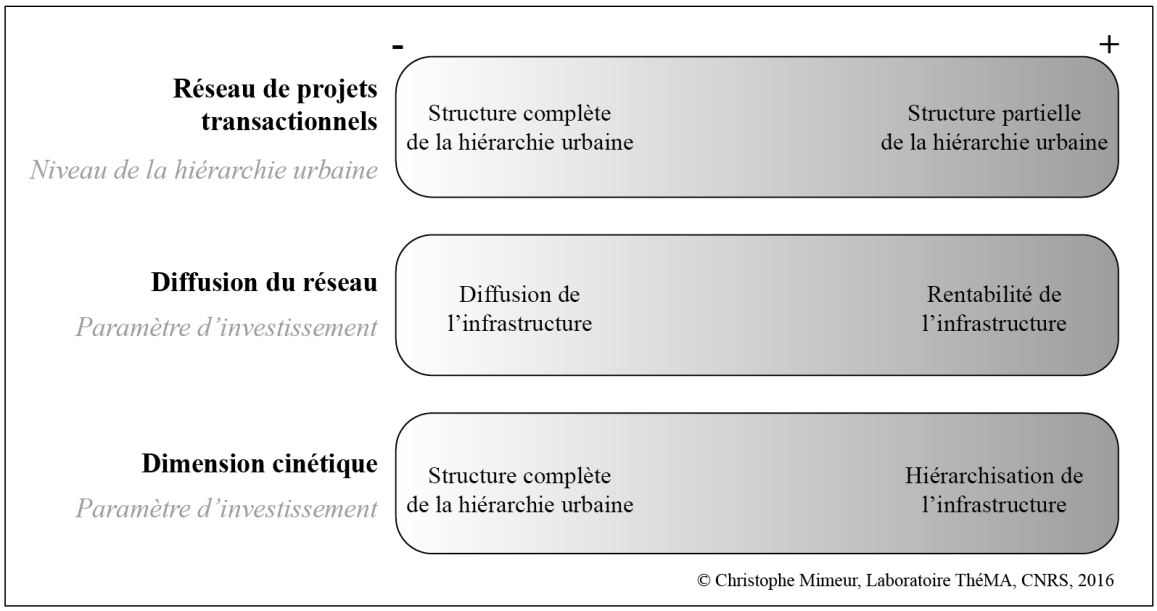


Figure 6. 19. Une signification géohistorique des paramètres du modèle d'évolution du réseau de transport

Ici, nous misons sur la modélisation *link-based*, afin de ne pas influencer sur les points de départ initiaux du réseau. On sait comment les premières lignes sont le fruit d'initiatives privées et comment le processus de mise en connexité a été long. En laissant le modèle décidé d'un investissement à l'échelle du lien – du tronçon – le modélisateur n'influence pas une structure prédéterminée et centralisée, où la construction du réseau répondrait davantage à une succession qu'une accumulation.

Dans ce cadre, le passage d'espaces stylisés au territoire français prend son sens dans la description de ces liens et des nœuds qui les font entrer en communication. L'hétérogénéité de l'espace est bien traduite par le semis de nœuds, dont l'attribut rend compte de la taille démographique. L'anisotropie de l'espace est laissée aux liens : les coûts homogènes des espaces stylisés laissent place à des coûts largement hétérogènes. On pose l'hypothèse que deux facteurs spatiaux participent au coût de chaque lien : la distance et la situation topographique.

Les jalons sont donc posés pour une exécution dont l'objectif reste semblable au développement suivi dans la section. Il s'agit de confronter la forme du réseau généré ainsi que la distribution des flux sur ce nouveau réseau, de sorte à le comparer au réseau réel. Cela pose également les premières interrogations quant à la confrontation d'un réseau réel et d'un réseau simulé, et du recours à la source géohistorique dans une démarche exploratoire dont fait l'objet la troisième partie de la thèse.

Conclusion

Ce sixième chapitre avait pour objectif de poser les fondations d'un modèle dynamique de l'évolution d'un réseau de transport, dans une perspective de longue durée et à une large échelle spatiale, afin de comparer un réseau construit par l'intelligence artificielle et un réseau effectivement construit dans l'histoire du transport français. Si ce type d'approche a marqué les débuts de la modélisation dynamique en géographie, la recherche opérationnelle s'est vite tournée vers des échelles largement plus réduites, tant d'un point de vue spatial que temporel. Elle vise à répondre à des problématiques de management des infrastructures et aux interactions entre système de transport et utilisation du sol, à l'heure où l'évaluation systématique a imposé des études avant et après (Delaplace, 2014). C'est dans le domaine de la science des réseaux (Dupuy et Offner, 2005) et parfois plus étonnamment du côté des sciences expérimentales, qu'un regain d'intérêt s'est manifesté au début des années 2000 (Xie et Levinson, 2009).

Les approches – originales – physiques, voire biologiques, nous ont permis d'aborder la croissance du réseau de transport par un processus d'auto-organisation, à partir de règles simples, qui aboutit à un processus de hiérarchisation du réseau généré (Yerra et Levinson, 2005). Nous questionnons ce mécanisme en investiguant la relation inverse des effets structurants des infrastructures de transport : la construction du réseau ferroviaire français répond-elle à des mécanismes d'efficacité basée sur la desserte des centres et sur la polarisation de l'espace ? Nous nous sommes efforcés ici de traduire les mécanismes du modèle proposé à travers les grands principes édictés par la théorie des réseaux d'une part, que nous adaptons au contexte géohistorique d'autre part. Classiquement, la démarche démarre d'une situation initiale, à partir de laquelle le modèle est exécuté en renforçant la vitesse des tronçons jugés les plus rentables à l'issue d'une évaluation économique, basée sur un modèle gravitaire et l'interaction spatiale. Le mécanisme consiste à générer un réseau de transport qui minimise la construction de l'infrastructure en maximisant les potentiels flux sur laquelle ils peuvent transiter, répondant à la question posée dans le chapitre.

La dernière section propose une exploration théorique du modèle à partir de trois distributions stylisées d'un système de peuplement. Dans une démarche instrumentale, la première évaluation permet de voir dans quelle mesure le réseau généré est cohérent avec le service qui peut y être proposé. Ensuite, une étude de la distribution de la hiérarchie des flux montre comment le processus de concentration inhérent aux itérations du modèle participe à la sélection de tracés optimaux pour la construction de la nouvelle infrastructure. A partir d'un réseau homogène, on voit comment la configuration spatiale influence les tracés et comment le temps long peut effacer l'opposition entre deux axes *à priori* concurrents. A l'issue de cette démarche théorique, on constate que le modèle est capable de générer un réseau stabilisé, qui traduit une rupture entre les liens bénéficiant d'une infrastructure lourde, et ceux qui n'en bénéficient pas. De même, le modèle produit une hiérarchie dans les axes préférentiellement choisis pour la construction. Ces deux conclusions nous poussent à poursuivre nos investigations vers une démarche empirique, dont l'objectif est de simuler la génération du réseau de transport à un état initial qui caractérise l'hétérogénéité et l'anisotropie du territoire français, de sorte à comparer le réseau simulé avec le réseau réellement construit. Cette comparaison participe pleinement aux débats sur le sens de la causalité entre infrastructure de transport et hiérarchie du système du peuplement.

CONCLUSION DE LA DEUXIEME PARTIE

L'objectif de cette seconde partie était de s'appuyer sur l'*« invention de la vitesse »* (Studeny, 1995) et sur son évolution sur le long terme pour construire un instrument de recherche dans le débat sur les effets structurants des infrastructures de transport. Il s'agit là d'une étape nécessaire dans le *« challenge émergent »* que nous avons identifié dans les travaux géohistoriques, faisant appel à l'outil SIG en particulier. La seconde partie a donc participé à la construction et l'adaptation de dispositifs d'analyse en adéquation avec la question de recherche posée, qui mêle vitesse, accessibilité et longue durée.

Conscients que *« les nouveaux outils permettent, à partir d'un travail sur les parcelles, les pixels ou le comportement des individus, de développer une recherche sur l'émergence de structures et de dynamiques globales à partir d'interactions qui se produisent à un niveau local »* (Guermond, 2005), la diversification des outils géomatiques participe au dialogue nécessaire avec les sciences informatiques et plus largement expérimentales. L'outil SIG est un outil parmi d'autres, et à vocation à communiquer directement avec ceux qui sont mobilisés dans le reste de la thèse. Parmi eux, le formalisme des graphes a montré alors le meilleur potentiel pour modéliser l'accessibilité géohistorique au niveau de la structure élémentaire de la commune française, en cohérence avec les modèles comportementaux. Ce chapitre a permis enfin de questionner les exigences de la source historique d'une part, et leur confrontation avec la construction d'informations à de larges échelles d'autre part, ce qui a permis d'aboutir à la construction d'un multigraphe, modélisant les déplacements géohistoriques à partir d'une triangulation du semis de communes.

Le cinquième chapitre a permis de poser les bases d'un corpus pour étudier l'effet du réseau sur les dynamiques démographiques. La dotation en réseau des communes est alors qualifiée de manière inédite, à partir de la construction de multiples indicateurs d'accessibilité. Pour autant, la prudence méthodologique nous fait avoir recours à la modélisation économétrique, dont la structuration en panel nous conduit à mettre l'accent sur les dimensions individuelles et temporelles. Pour cela, nous avons fait le choix de la diversification de l'information géohistorique pour affiner les estimations et réduire le risque important d'un biais d'endogénéité. Une première stratégie consiste à prendre en compte des variables environnantes susceptibles d'avoir un effet sur le contexte local, qu'elles soient variantes ou invariantes. La seconde stratégie suppose l'utilisation de variables instrumentales, testant par la même occasion l'adaptabilité de la démarche développée dans le *Chapitre 4* dans la construction du multigraphe.

Le sixième chapitre a permis de complexifier encore l'approche, en investiguant une autre partie des débats sur l'effet structurant des transports. L'étude de la relation inverse, misant sur le primat d'un système de peuplement préexistant ayant régi la construction du réseau ferroviaire français, permet d'éprouver les potentialités de la modélisation dynamique, courant dans lequel la

géographie connaît un vif succès depuis 20 ans, sans que l'évolution propre des réseaux de transport n'y occupe une place importante. C'est donc par une approche endogène, souvent peu privilégiée, des configurations d'un réseau de transport que nous avons construit un modèle qui questionne deux processus – diffusion et hiérarchisation – dans l'application de la théorie des réseaux. L'approche par l'exploration du modèle dans trois espaces stylisés a permis de comprendre les mécanismes et l'impact du choix du modélisateur dans les paramètres utilisés dans la génération du réseau.

C'est à l'issue de ces trois chapitres que nous pouvons construire un plan d'expérimentation qui mobilise l'étendue de notre instrument de recherche au service de la question qui guide cette thèse. La *Figure 6. 20* permet de définir la stratégie d'analyse du lien entre réseau et territoire en montrant comment l'ensemble de l'instrument sera mobilisé dans la troisième partie de la thèse, quand, comment et pourquoi ? Elle montre comment les phases de de modélisation de la donnée et des relations sont étroitement imbriquées. Elle démontre le besoin d'une cohérence dans la réflexion afin de produire des connaissances géohistoriques, en favorisant le dialogue entre la question de recherche, les données mobilisées et les dispositifs d'analyse dont il a été question dans cette seconde partie. Ainsi, au cœur des interactions entre réseau et population, le multigraphe est le premier jalon nécessaire à la modélisation des données pour envisager l'effet du réseau. En enrichissant les analyses par de nouveaux indicateurs d'accessibilité, cette réflexion doit néanmoins d'appuyer sur la prise en compte d'un contexte local qui requiert la mobilisation de données extérieures. A l'inverse, l'étude du rôle d'une structure préexistante dans le façonnement des réseaux mobilise tout autant le graphe perçu comme la base de nos analyses. Il est celui qui permet l'exécution du modèle d'évolution d'un réseau de transport en permettant la modélisation des interactions spatiales d'une part et de la dynamique de renforcement d'autre part. Ce côté du plan d'expérimentation doit également se doter de références géohistoriques pour confronter réseau simulé et réseau réel et dégager les pistes d'explication quant à la stratégie d'implantation des tracés. Le plan d'expérimentation est conduit à partir de la structure élémentaire qu'est la commune, depuis la première mise en connexité du réseau jusqu'à son expansion maximum.

De la modélisation des données, jusqu'à la mise en croissance dynamique de la génération d'un réseau, en passant par la démarche analytique de l'effet de l'infrastructure, ce plan d'expérimentation participe à des exigences heuristiques, au cœur du couple transport et territoire qui revêt tant de complexité. Il participe tout autant à des exigences didactiques, tant le regard sur la situation passée peut apporter de nouvelles pierres à l'édifice de l'analyse de la situation présente. La troisième partie de la thèse se place ainsi dans le domaine empirique, en cherchant à répondre aux enjeux spatiaux et temporels posés depuis le début de la thèse. Le graphe devient alors non seulement l'objet des relations entre lien et nœud du territoire, mais celui de la génération des dynamiques circulatoires entre tous les nœuds mais seulement par certains liens.

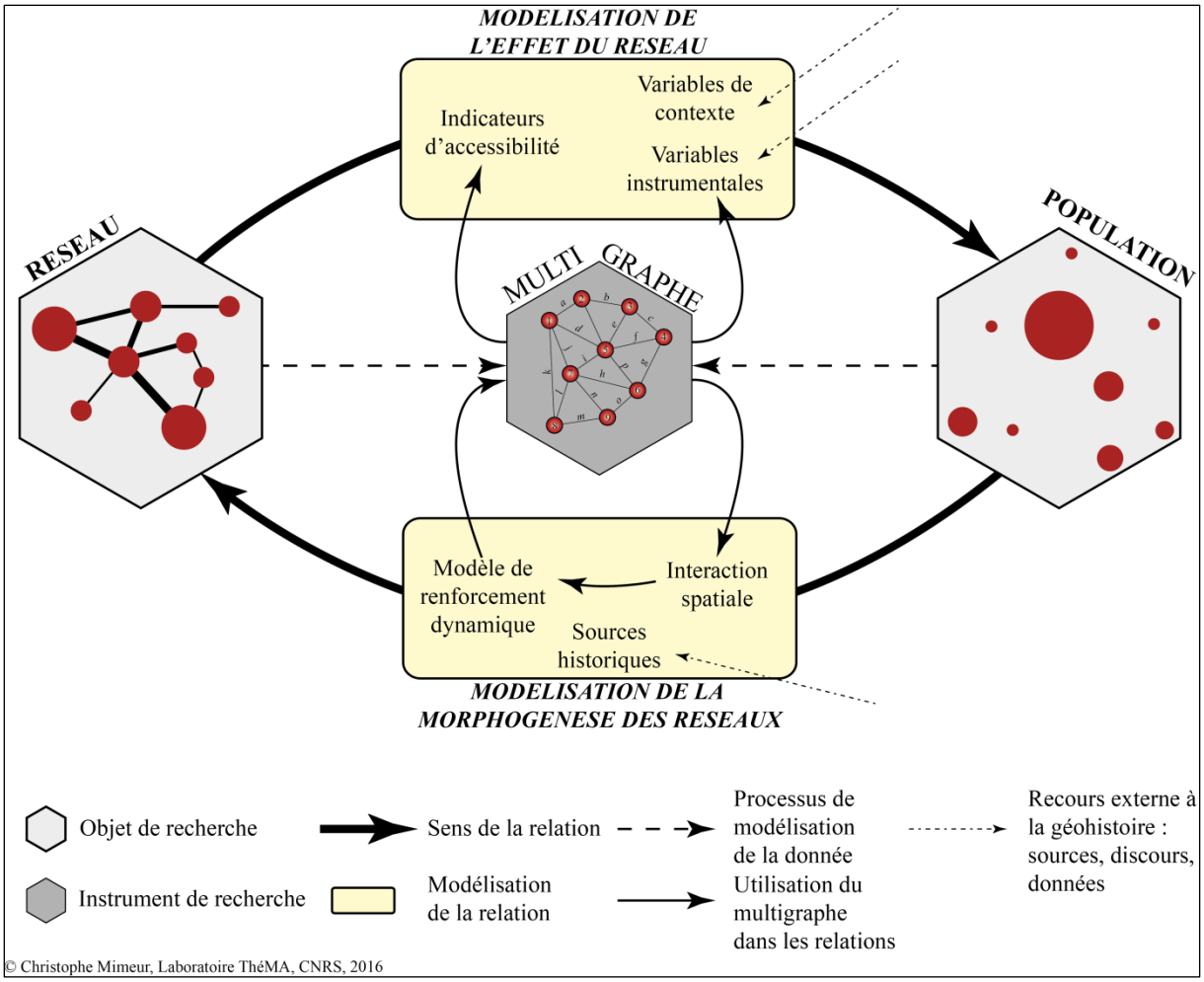


Figure 6. 20. Le plan d'expérimentation pour l'étude des interactions entre réseau et population

PARTIE 3 :

GEOHISTOIRE DE L'INTERACTION ENTRE RESEAU ET POPULATION :

DE LA MESURE DE L'EFFET DU RESEAU A LA SIMULATION DE SA CROISSANCE

INTRODUCTION DE LA TROISIEME PARTIE

« Les chemins mènent toujours quelque part.
L'essentiel n'est donc pas dans la direction qu'il suit,
mais dans le fait que la direction existe. »
Fiodor DOSTOÏEVSKI

La première partie de la thèse a permis de définir les objets de la recherche. La place des échelles est alors centrale et l'espace doit y jouer un rôle prééminent dans l'interrogation transcalaire des relations entre réseau et territoire. Dans ce cadre, le recours à la quantification s'est justifié à de larges échelles : dans notre projet, l'exploitation des données existantes l'emporte en revanche sur la phase de collecte de l'information. Par conséquent, la seconde partie a mis l'accent sur le processus de modélisation des relations entre nos deux objets, depuis la modélisation de la donnée géohistorique jusqu'aux dispositifs d'analyse qui prennent corps dans le plan d'expérimentation proposé à l'issue de ces réflexions. L'instrument de recherche ainsi constitué a pour vocation à être confronté à la phase de croissance du réseau ferroviaire français, en s'inscrivant dans une démarche constructiviste qui vise à dégager des structures explicatives en confrontant une expertise sur les sources historiques, des théories d'arrière-plan et les données issues de la base de données FRANcE.

Cette troisième partie permet désormais de mettre en œuvre le plan d'expérimentation en investiguant les interactions entre réseau et territoire à partir de la base de données FRANcE. Il s'agit donc d'envisager ces relations d'un point de vue non seulement spatial mais aussi temporel. En cela, ce travail s'inscrit non seulement dans la littérature des effets de long terme (Kasraian et al., 2016) mais aussi dans le challenge émergent des projets de SIG-H qui vise à valoriser la donnée pour en tirer la production de nouvelles connaissances en diversifiant les approches et en les confrontant avec le contexte dans lequel elles s'inscrivent (Gregory et Geddes, 2014). Appliquée au cas français, cette recherche s'inscrit désormais pleinement au cœur de la controverse scientifique portant sur les effets structurants des infrastructures de transport (Bretagnolle, 2014 ; Offner, 2014), dans la mesure où elle permet de réinterroger ses principaux vocables à de larges échelles, en faisant l'hypothèse que *« les effets structurants sont bien là »* dans la double démarche initiée par le plan d'expérimentation.

Il s'agit d'abord d'analyser les différents indicateurs d'accessibilité géohistorique construits auparavant afin de qualifier la phase d'expansion et d'accélération du réseau ferroviaire français dans les doubles dimensions spatiale et temporelle. Le septième chapitre propose alors d'investiguer les vocables d'homogénéité, de hiérarchie et d'hétérogénéité à partir de différentes approches. Les premières confrontations de ces indicateurs aux variables de population nous poussent à identifier des interrelations entre ces deux objets, qui permettent d'interroger les concepts de coévolution et

de centralité. Pour autant, les rigueur et prudence méthodologiques requises par l'étude des relations entre transport et territoire poussent à utiliser le cadre méthodologique de l'économétrie proposé dans le *Chapitre 5* afin de rechercher un rapport structurant du réseau sur les dynamiques démographiques. Cette étape participe à la mobilisation élargie du corpus rassemblé dans la base de données FRANcE, depuis la modélisation des effets directs et indirects jusqu'à l'utilisation exploratoire des variables instrumentales pour limiter le biais d'endogénéité (Chapitre 7).

La seconde partie du plan d'expérimentation nous incite à étudier l'autre sens des relations entre réseau et population. Pour cela, nous mobilisons de manière inductive les indicateurs locaux de graphe, aptes à nous renseigner sur les capacités circulatoires du réseau en fonction des choix infrastructurels effectués et des différentiels de vitesse appliqués. Désormais, ce sont les vocables de sélection, concentration et hiérarchisation qui sont interrogés dans le cadre de la controverse des effets structurants. Par la suite, la confrontation de ces nouveaux indicateurs aux données démographiques nous amène à interroger la mise en forme des réseaux à partir d'un ensemble de règles simples qui permet la mise en œuvre dynamique du modèle. Il s'agit d'utiliser le modèle d'évolution d'un réseau de transport imaginé dans le *Chapitre 6* dans une application à la base de données FRANcE afin de vérifier la citation de Fiodor Dostoïevski ci-dessus. Ainsi, dans un contexte dépourvu de considérations stratégiques et politiques, en quoi le modèle parvient-il à simuler et à reproduire la croissance du réseau ferroviaire français à partir des trois paramètres du modèle : cette croissance participe-t-elle à la prise en compte d'une structure complète ou non de la hiérarchie urbaine, participe-t-elle à un phénomène de diffusion de l'innovation et privilégie-t-elle la théorie de la *« branche commune »* dans la structuration du réseau ? La confrontation d'un réseau simulé et du réseau réel participe à la réponse à ces questions, à condition de joindre notre discours à d'autres sources et d'envisager des dispositifs d'analyse (Chapitre 8).

Finalement, la mise en œuvre du plan d'expérimentation montre dès maintenant comment le chercheur géohistorique ne peut s'isoler complètement à l'intérieur de son instrument de recherche. La mobilisation de données extérieures et le recours aux discours techniques, politiques et historiques sont indispensables à la contextualisation du processus de connaissances géohistoriques, et c'est bien la démarche entreprise lors de cette troisième partie de la thèse. Enfin, à partir des résultats produits, il s'agira d'envisager les perspectives de recherche à l'issue d'un projet auquel il est impossible de mettre un point final à l'issue de ce travail (Chapitre 9).

Chapitre 7. DE L'ACCELERATION DU RESEAU A LA DEFINITION DES EFFETS SUR LES DYNAMIQUES DEMOGRAPHIQUES

Introduction

Les premières exploitations de la base de données FRANcE ont fait l'objet de requêtes tantôt spatiales, tantôt temporelles. Ainsi, le *Chapitre 3* a permis d'étudier, dans le domaine empirique français, dans quelle mesure le développement du réseau ferroviaire français participe à un modèle de croissance proche de celui du cycle des produits. En historisant la phase de croissance, nous avons défini des périodes de balbutiement, de décollage et de maturité du réseau, dans son ensemble et selon sa hiérarchie. Ainsi, l'accent a davantage été porté jusqu'à maintenant sur l'analyse temporelle de la croissance. Cette troisième partie de la thèse a pour objectif de concilier les approches spatiales et temporelles, cette fois-ci en géohistorisant la phase de croissance. Ce chapitre participe à considérer cette phase dans un cadre spatial en utilisant la principale valeur ajoutée de la base, la vitesse et le calcul des accessibilités à de larges échelles.

Finalement, il s'agit de questionner les principaux facteurs qui font que le mythe des effets structurants demeure aujourd'hui d'actualité (Offner, 1993a, 2014 ; Plassard, 1977). On se propose alors de remettre en perspective les principaux vocables de la controverse scientifique, en les insérant dans un quadruple questionnement : où, quand, comment, pourquoi ? Pour cela, nous avons recours à l'explication spatio-temporelle, qui utilise l'instrument de recherche construit dans la seconde partie, et plus particulièrement le multigraphe assorti des modèles économétriques. Nous avons alors, dans la droite ligne de notre positionnement épistémologique, à effectuer des allers et retours entre la donnée contenue dans FRANcE, la théorie des réseaux et les contextes géohistoriques pour construire une modélisation des effets du réseau sur la population française.

Dans ce chapitre, nous nous insérons dans le cadre de la littérature des effets de long terme entre réseau de transport et territoire. Nous utilisons la complexification de l'information sur la dotation en réseau pour « géohistoriser » les relations entre transport et territoire dans la phase de croissance du réseau français jusqu'en 1930. Pour cela, nous verrons d'abord dans quelle mesure le mythe de l'association universelle souffre de disparités sur le territoire français (7.1.), complexifiant l'analyse des relations entre réseau et population. Ensuite, nous étudierons les relations structurantes entre les centralités créées par le réseau et la population (7.2.) avant d'analyser l'apport de l'économétrie de panel pour étudier les relations entre accessibilité et croissance démographique locale (7.3.).

7.1. Expansion et accélération : une réelle homogénéisation ?

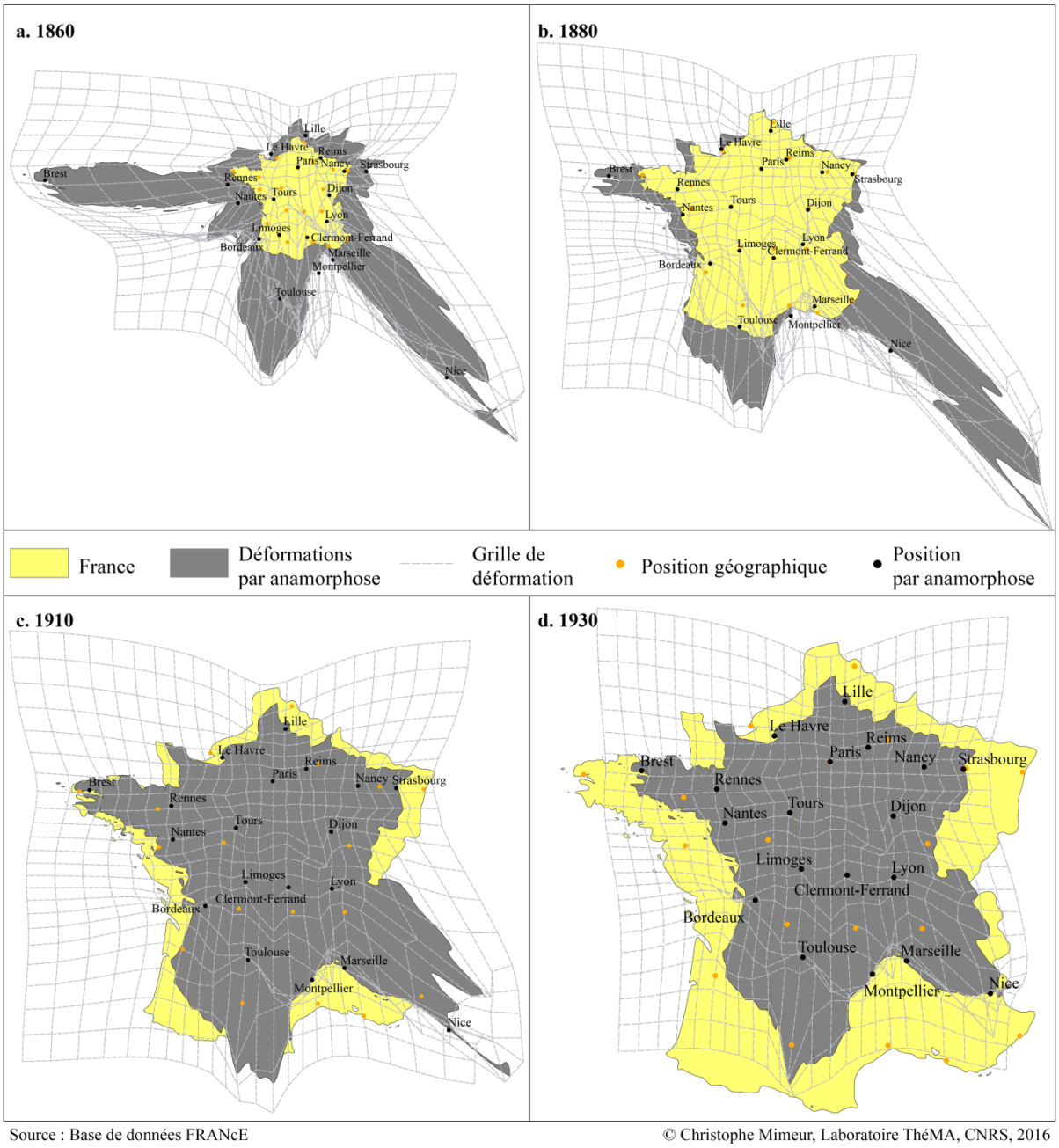
Dans le *Chapitre 3*, les premières exploitations du SIG ont permis de qualifier les différentes phases de la croissance du réseau ferroviaire selon la hiérarchie des voies *(cf. 3.2.2, p. 123)*. A partir de ces courbes tendancielles qui décrivent les phases de balbutiement, décollage et maturité du réseau, nous proposons ici d'investiguer une nouvelle fois ces trois moments à partir de la *« révolution de la vitesse »* qualifiée dans la seconde partie de la thèse. L'objectif est d'étudier la phase d'expansion en la complexifiant par la notion d'accélération, venant apporter de nouveaux éléments à la dotation en réseau, et d'étudier dans quelle mesure cette géohistoire du réseau répond au mythe de l'effet structurant, comme celui de l'association universelle d'une part, comme celui de l'initiateur d'une croissance démographique d'autre part, à différentes échelles, nous transportant du multiscalaire au transcalaire.

7.1.1. L'ouverture de l'horizon spatio-temporel : continuités et discontinuités des déformations de l'espace

Si la croissance du réseau ferroviaire français a permis de nouvelles connexions par un nouveau moyen de transport, alors *« l'essor de la rapidité différenciée a radicalement modifié ce rapport à l'espace-temps »* (Ollivro, 2000). Alors que jusqu'à maintenant, les moyens de transport ne provoquaient pas de bouleversements entre les configurations géographiques et la durée moyenne des déplacements, l'irruption ferroviaire provoque une rupture par l'instauration de l'hétérogénéité des vitesses. Il s'agit ici de qualifier l'ouverture de l'horizon spatio-temporel, telle qu'elle est vantée par le corps des ingénieurs des Ponts et Chaussées (Chevalier, 1832).

Dans la seconde moitié du XXème siècle, les rapides innovations informatiques ont permis de traduire les modifications de l'espace-temps par des représentations cartographiques déformées, qui rendent compte non seulement des espaces accessibles, mais aussi de ceux qui ne le sont que difficilement : on retrouve ici les métriques territoriales et réticulaires définies dans la première partie de la thèse (Lévy et Lussault, 2003). L'objectif n'est pas de substituer l'une à l'autre mais de voir dans quelle mesure elles se différencient et comment leur rapport évolue sur le temps long. L'une des principales méthodes utilisées est celle de la distorsion cartographique : alors qu'une simple carte de situation renseigne sur la localisation absolue d'entités spatiales, l'anamorphose permet l'étude des localisations relatives (Béguin et Pumain, 1994) parce qu'elle traduit la disparition de la proportionnalité entre l'espace et le temps. L'objectif est d'étudier dans quelle mesure le développement du réseau est homogène sur le territoire français, à quel moment influence-t-il le plus fortement ces distorsions, quels en sont les gagnants et les perdants ? Les premiers travaux ont été à l'initiative de Wentworth d'Arcy Thompson en 1948 et Stephen Jay Gould en 1971. Cette méthode a néanmoins connu un regain d'intérêt dans les années 1980 grâce aux travaux de Colette Cauvin (Thévenin, Schwartz et Sapet, 2013).

Dans un premier temps, nous étudions la modification de l'espace fonctionnel français à partir de l'accessibilité à la capitale, parce qu'elle est l'essence du programme de Legrand lancé en 1842. La *Figure 7. 1* montre l'évolution de cet espace entre 1860 et 1930.



Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Figure 7. 1. Evolution des distorsions spatiales par l'accessibilité à Paris entre 1860 et 1930

A partir des positions géographiques des lieux, et donc leur distance euclidienne à Paris, nous définissons tout d'abord les coordonnées de ces mêmes points à partir des distances temporelles, à l'aide d'un calcul vectoriel. Ce calcul est relatif à un vecteur qui fait le rapport entre la somme des distances métriques et la somme des distances temporelles. Pour permettre la

comparaison, nous choisissons un vecteur moyen, que l'on identifie pour la décennie 1880, deux ans après l'inflexion du réseau *(cf. Figure 3. 12, p. 124)*. La comparaison de ces deux types de position fait l'objet d'une régression multidimensionnelle, qui permet de mettre en évidence les déformations de l'espace[40]. Par interpolation de l'ensemble des positions, nous construisons alors l'espace temporel issu de la transformation.

La *Figure 7. 1* montre trois types d'informations. La première montre la position géographique initiale des villes dans l'espace euclidien. Ces mêmes villes sont aussi projetées dans l'espace temporel. La grille de déformation facilite la lecture du nouvel espace construit : alors que l'espace euclidien possède une grille régulière, la projection montre comment l'accessibilité à Paris modifie cette grille et déforme l'espace. La lecture de telles anamorphoses est facilitée quand la déformation de l'espace est continue, ce qui n'est pas forcément le cas ici. En 1860, on voit comment le réseau ferroviaire creuse finalement les inégalités, parce qu'il ne concerne qu'une partie du territoire *(Figure 7. 1. a.)*. La Bretagne se retrouve alors fortement éloignée. Plus naturellement, les reliefs des Alpes et des Pyrénées se retrouvent à l'écart eux aussi. La déformation montre comment le Massif Central est repoussé, aussi intensément que les Pyrénées. A l'inverse, la distorsion montre une rétractation de l'espace dans un couloir qui part du Nord de la France et qui s'étend du Havre à Nancy. En 1880, Nice, Brest et Toulouse bénéficient des rapprochements les plus forts *(Figure 7. 1. b.)*, traduisant un phénomène de rattrapage alors qu'au Nord, les positions des principales villes se rapprochent de leur position géographique, voire sont encore plus rapprochées de Paris, ce sont surtout les villes désignées par le plan Legrand, dans lesquelles on remarque que Bordeaux et Marseille bénéficient des radiales vers Paris. En 1910, les gains d'accessibilité profitent encore aux villes de la partie Nord du pays *(Figure 7. 1. c.)*, qui se rapprochent davantage de Paris, tandis que les villes de Brest et Nice poursuivent leur rattrapage. En 1930, la situation n'évolue que marginalement dans la partie Nord *(Figure 7. 1. d.)*, tandis que l'ultime phase de développement du réseau bénéficie aux reliefs les plus accidentés, même si la grille de déformation montre toujours l'effet barrière provoqué par le Massif Central.

Si la cartographie en anamorphose propose des clichés des espaces fonctionnels à certaines dates, nous souhaitons confirmer et mieux qualifier la phase d'accélération par l'étude de la trajectoire des villes dans les espaces nouvellement créés. Cette analyse est rendue possible par le choix d'un vecteur moyen, utilisé pour l'ensemble des quatre dates retenues. La *Figure 7. 2* reprend en pointillé la forme des espaces fonctionnels obtenus par la régression multidimensionnelle. A chaque fois, la position des principales villes françaises est reportée et permet de tracer les trajectoires dans l'espace fonctionnel distendu par la croissance du réseau ferroviaire français. Tout d'abord, toutes les trajectoires passent par la position géographique de ces mêmes villes. Le moment de ce passage par la position géographique permet de qualifier le moment à partir duquel le gain d'accessibilité est le plus sensible, où le rapport au temps semble se substituer au rapport à l'espace. Dès lors, on peut interpréter ces trajectoires en isolant trois groupes de villes, qui croisent espace et temps :

[40] Toutes ces opérations sont intégrées dans deux logiciels libres développés au sein du laboratoire ThéMA : IsoDistAngle et Darcy.

- Le premier groupe rassemble des villes dont l'apport en chemin de fer est significatif dès le déploiement du réseau entre 1860 et 1880 : il s'agit de Lille, de Reims, de Nancy et de Strasbourg. Ce sont elles qui bénéficient des grandes radiales décidées dans le programme de Legrand. Lyon, Bordeaux et Marseille complètent cette liste. On constate toutefois qu'un certain nombre d'autres sont absentes de ce premier groupe.
- Le second groupe regroupe celles qui font l'objet d'un rattrapage dès le début des années 1880. Avec Dijon, Tours, Limoges et Montpellier, ce sont d'abord des villes de passage le long des grandes radiales dessinées par Legrand. C'est même aussi le cas pour Clermont-Ferrand, Rennes, Toulouse de manière plus mesurée. D'autres villes terminus poursuivent un rattrapage important, comme Le Havre et Brest.
- Le troisième groupe ne comprend que la ville de Nice, qui bénéficie d'une position favorable qu'entre 1900 et 1930. La barrière alpestre est contournée par le segment Marseille-Nice, dont l'augmentation des performances semble tardive.

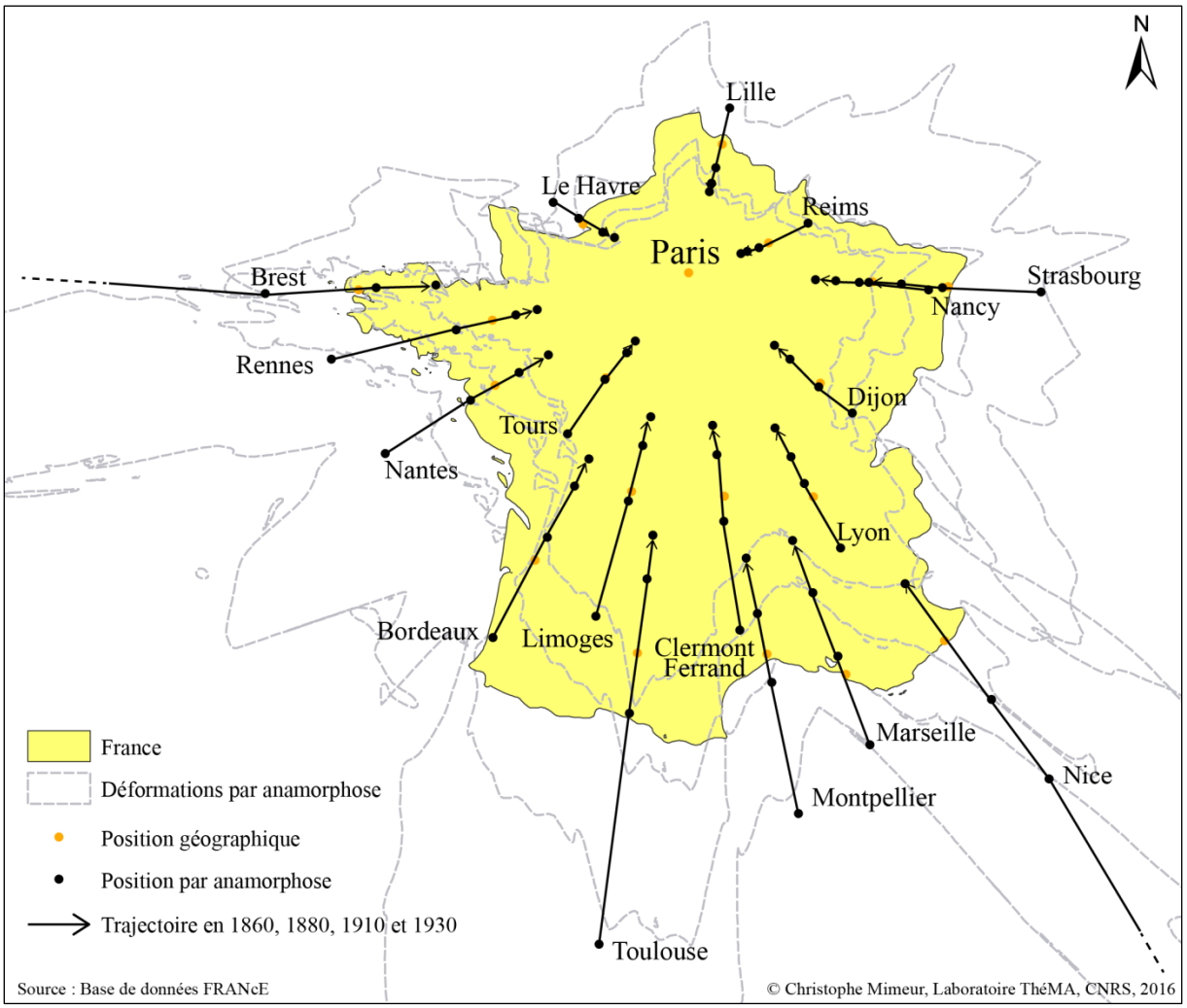


Figure 7. 2. La trajectoire des villes françaises dans la distorsion de l'espace fonctionnel pour rejoindre Paris entre 1860 et 1930

Au final, la *Figure 7. 2* traduit d'abord une accélération permanente vers Paris, qui se traduit

par une contraction croissante de l'espace. Cette phase n'est cependant pas homogène : la première phase est dans tous les cas la plus importante. Pour les villes dont la situation est la plus favorable – le premier groupe –, les améliorations ultérieures demeurent marginales. Pour les autres groupes, le phénomène de rattrapage est souvent fonction de la position géographique, ou de l'amélioration significative de l'infrastructure, comme c'est le cas dans le premier tiers du XXème siècle au sud-est de la France.

Pour autant, cette approche ne prend en compte que les principales villes, dans leur rapport unique à la capitale. On approche ici une limite de l'anamorphose unipolaire, parce qu'elle favorise une vision continue des distorsions alors qu'au contraire, *« la variété parfois quotidienne des appartenances et l'impossibilité de représenter avec fidélité le monde contemporain […] est au centre du processus »* (Ollivro, 2000). Dès lors, il semble intéressant de prendre en compte la position relative des villes, mais cette fois-ci pas seulement dans leur rapport à Paris. L'objectif est d'étudier les contrastes d'accessibilité entre ces différents points de l'espace, en mobilisant les nouveaux indicateurs construits dans la seconde partie de la thèse. La méthodologie est presque identique : elle oppose une matrice de distance euclidienne à une matrice de distance temporelle, où chaque couple de points est caractérisé par ces deux attributs. La comparabilité et la généralisation de l'information à l'ensemble du territoire requiert cette fois un processus d'ajustement par transformation euclidienne.

La *Figure 7. 3* répond aux mêmes règles de construction que la précédente. Une simple lecture des différentes déformations montre la relative homogénéisation de l'espace fonctionnel français, puisqu'elles se rapprochent progressivement de la forme géographique du pays, épousent sa forme, jusqu'à se rétracter à l'intérieur de celui-ci. Par ailleurs, la lecture des trajectoires renseigne non seulement sur le processus individuel de chaque ville mais aussi sur le processus de proximité relative entre ces villes. L'écart à la position géographique montre l'apport du chemin de fer et à quel moment il est le plus important.

Dans un premier temps, l'évolution du réseau entre 1860 et 1880 montre les transformations les plus importantes dans notre période d'étude, où il s'agit des premiers moments du décollage du réseau. Elles montrent aussi comment les inégalités étaient importantes au début de la période. Paris est la ville dont la trajectoire est la moins importante : son statut central est une constante dans notre période d'étude. La capitale semble alors fédérer autour d'elle les principales villes de l'Est, du Centre et du Sud. Ce sont elles qui vont se déplacer le plus significativement vers leur position géographique par la suite. Se rapprocher de celle-ci signifie finalement trouver sa place dans le réseau dans son ensemble, où la ville est à la fois une force d'émission et d'attraction.

Dans tous les cas, les changements sont ensuite plus marginaux. Mais la densification progressive du réseau participe à des processus de rapprochement ou d'éloignement entre les villes de province. Elle participe à l'identification de proximité laissant penser à l'émergence de systèmes régionaux. C'est le cas de la Bretagne où les trajectoires de Rennes, Nantes et Tours sont à peu près semblables et participent au rapprochement de ces trois villes. Brest s'y rattache au fur et à mesure que le temps avance. Nous avons le même ordre d'état de fait dans le Sud, entre les villes de Montpellier, Marseille et Nice, où l'évolution montre des processus successifs de rapprochement et de relatif éloignement entre Marseille et Montpellier.

Limoges semble faire figure d'exception, dans la mesure où la première période d'étude est

l'occasion d'un grand rapprochement vers Bordeaux, en même temps qu'un éloignement de la capitale : elle trouve une place progressive dans un axe qui se dessine entre Paris et Bordeaux, en passant par Tours par un itinéraire et donc par Limoges par un autre itinéraire. Dans le même temps, Toulouse semble se rapprocher de ce système qui rejoint la capitale, tout en participant à l'homogénéisation de l'espace dans son environnement plus immédiat, au premier rang duquel figure le relief pyrénéen.

Dans le Nord et l'Est de la France, les trajectoires répondent aux axes qui structurent cette partie du territoire, c'est-à-dire les radiales qui partent vers le Havre, Lille et Strasbourg. Reims et Nancy se replacent alors progressivement comme des villes intermédiaires le long de ce dernier axe. C'est aussi le cas de Dijon dans l'axe Paris-Lyon-Marseille en début de période. Pour autant, le développement d'un réseau franc-comtois et alsacien modifie sa trajectoire vers l'est du pays. Enfin, les positions relatives de Clermont-Ferrand et Lyon sont significatives d'un éloignement progressif des deux villes, dans le renforcement des radiales, alors même que les phases ultimes de développement du réseau participent finalement à leur rapprochement.

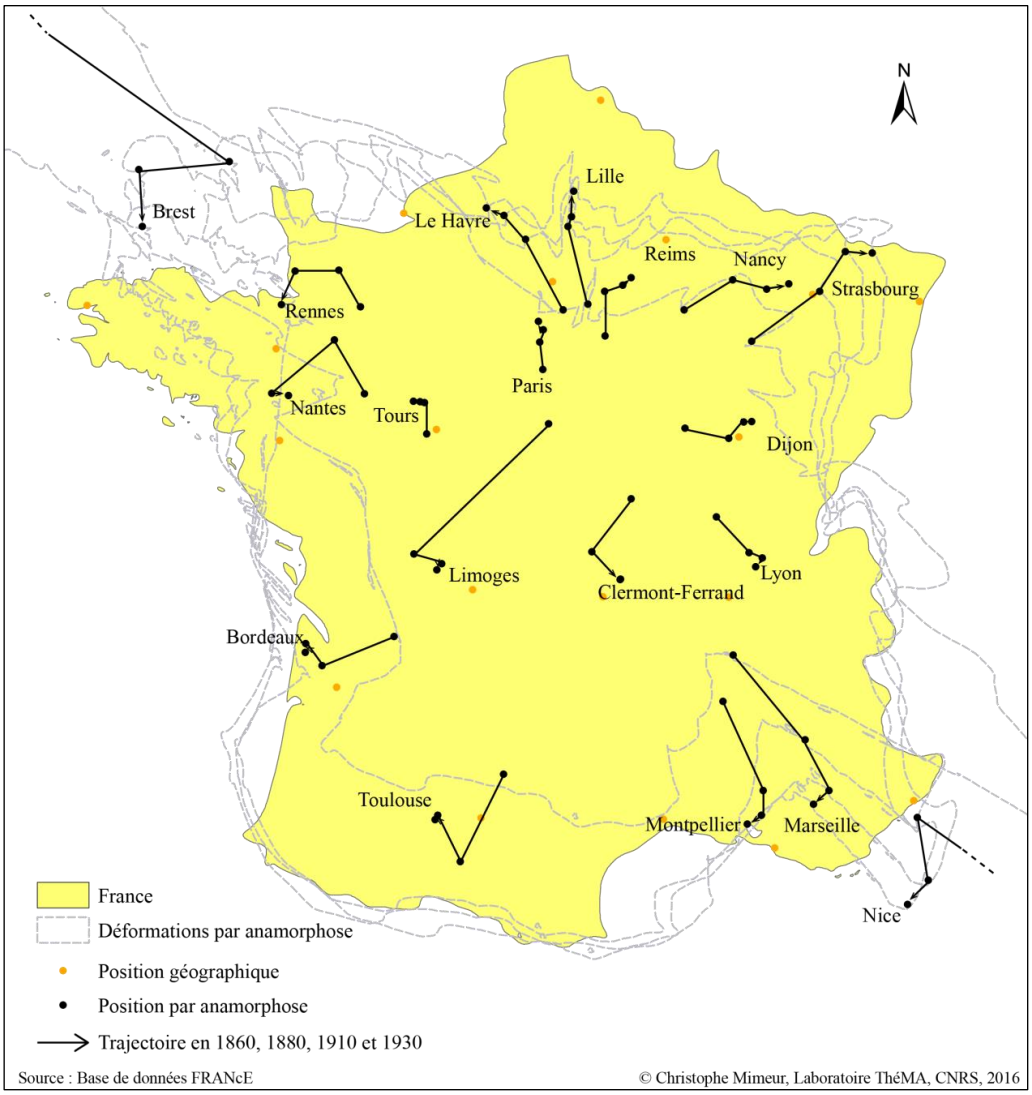


Figure 7. 3. La trajectoire des villes dans la distorsion multipolaire de l'espace fonctionnel entre 1860 et 1930

Ainsi, cette première approche de l'accélération propose une étude des trajectoires de villes dans la modification de l'espace fonctionnel français entre 1860 et 1930. La première phase de décollage, avant même le point d'inflexion défini en 1878 est l'objet des trajectoires les plus changeantes, comme si l'arrivée du chemin de fer avait bouleversé l'espace-temps : notre approche combine gains d'accessibilité et déformations de l'espace-temps. Les gains les plus significatifs sont alors à chercher du côté de l'ouest, pour rejoindre la capitale, parce qu'il s'agit des zones les plus éloignées. Les gains sont plus continus dans la partie est de la France, sauf pour la vallée du Rhône qui bénéficie de l'infrastructure la plus rapide. D'un point de vue multipolaire, la diffusion de chemin de fer participe à une dilatation progressive de l'espace pour se rapprocher de l'espace géographique, semblant rendre compte d'une homogénéisation de l'espace et de l'émergence de sous-systèmes régionaux.

Là où l'anamorphose unipolaire a démontré une forte contraction centrée sur Paris, l'anamorphose multipolaire laisse à croire à une homogénéisation de l'espace, où les radiales semblent toutefois conserver un rôle prééminent. Etienne Julliard, géographe français, affirmait avoir *« mal à [sa] France »* quand il voyait de telles anamorphoses, sûrement en partie parce qu'elles permettent davantage d'identifier des phénomènes continus que l'inverse. Nous proposons alors une vision intermédiaire, que nous qualifions de polynucléaire, pour identifier l'échelle des bouleversements, de même que leur rythme et leur contribution à la création de continuités et de discontinuités.

7.1.2. **Une accélération polynucléaire ?**

La déformation de l'espace reste encore utilisée aujourd'hui comme un formidable outil de communication, dont nous venons toutefois de pointer les limites. Nous visons ici une vision multiscalaire de l'accélération permise par l'infrastructure ferroviaire, qui passe par un autre mode d'analyse, puisque les procédés en anamorphose ne proposent pas de représentations intermédiaires entre l'accessibilité unipolaire et l'accessibilité multipolaire. L'objectif est d'étudier quelle est la portée du phénomène d'accélération dans la phase de croissance du réseau ferroviaire : dans quelle mesure la construction d'un réseau rapide, intermédiaire ou lent participe à l'amélioration de l'accessibilité à des échelles nationales, régionales ou départementales ? Dans notre logique géohistorique d'une part, en mobilisant le modèle conceptuel PONT d'autre part, cette partie mobilise la structure administrative du pays, en comparant les accessibilités sur le temps de la capitale Paris, des préfectures régionales de la France à 22 régions et des préfectures départementales. Est-ce que la vision centralisatrice du corps des Ponts et Chaussées participe à des gains préférentiels vers la capitale ? Est-ce que la mise en place du plan Freycinet n'a qu'une influence locale ?

Pour cela, nous utilisons une comparaison entre une situation contrefactuelle[41] et une situation avec le réseau ferroviaire. D'un point de vue absolu, cette comparaison permet de comprendre l'ampleur des progrès réalisés par le réseau ferroviaire. D'un point de vue relatif, sur

[41] Une situation contrefactuelle est une situation inobservée dans la réalité, c'est-à-dire dans bénéficier de l'innovation ferroviaire.

la longue durée, cette comparaison permet de voir à quel rythme le différentiel entre ces deux situations augmente, s'il y a des ruptures, ou une évolution continue. L'objectif est de comprendre l'évolution de ce différentiel aux trois échelles définies plus haut : nationale, régionale et locale. La *Figure 7. 4* expose la démarche pour qualifier l'accélération à l'échelle nationale. La première étape consiste à confronter des temps de parcours sur le graphe P – la situation contrefactuelle – et sur le multigraphe F – la situation avec le réseau ferroviaire *(Figure 7. 4. a.)*. Une régression linéaire permet d'apprécier la pente qui qualifie l'effet de l'accélération du réseau. Notons que le coefficient de détermination de ces régressions est toujours significatif. Alors on peut confronter les droites de régression à chaque décennie *(Figure 7. 4. b.)*. Ces deux étapes sont répétées pour les trois échelles de l'étude. Dans les trois cas, les pentes s'accentuent sur le temps long, attestant d'une accélération continue sur l'ensemble de la période d'étude.

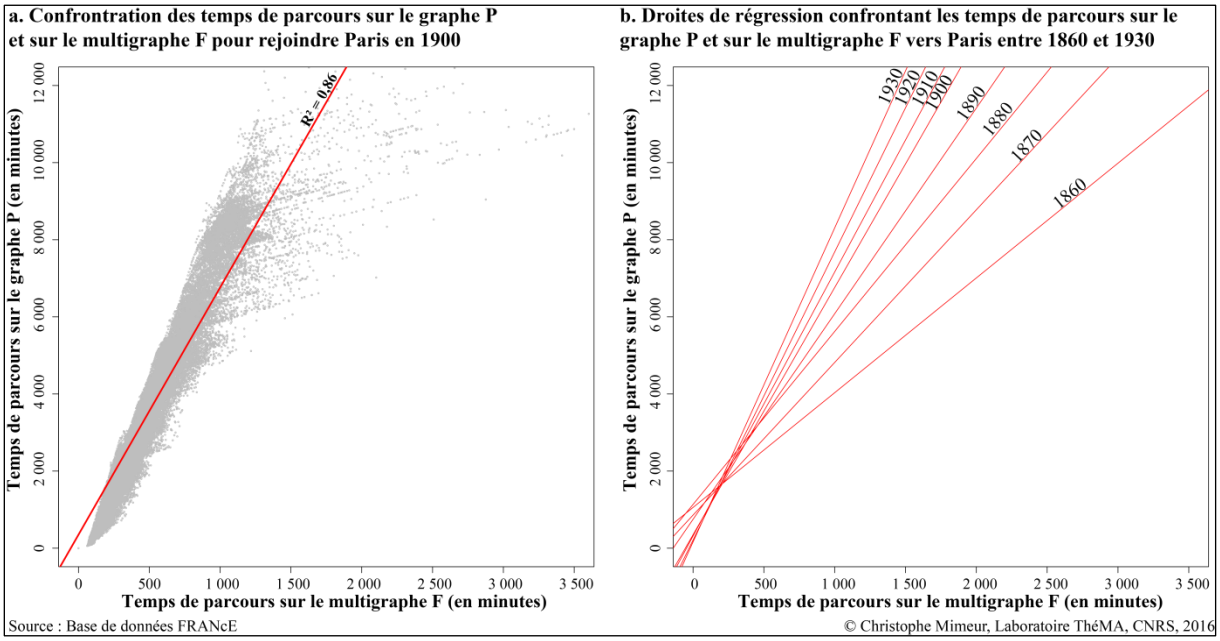


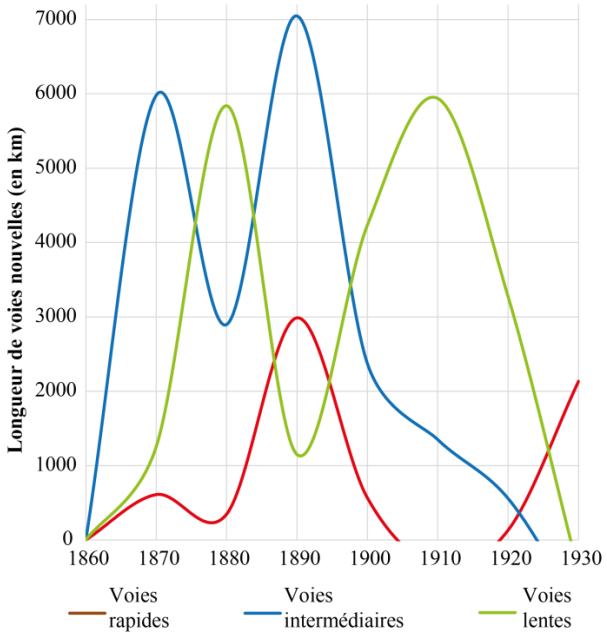
Figure 7. 4. Une comparaison à une situation contrefactuelle sur le temps et à trois échelles

Ce sont toutefois précisément les écarts entre chaque accentuation de pente qui nous intéressent et qui qualifient le rythme de l'accélération. La *Figure 7. 5. a.* permet de dresser l'évolution des différentiels entre chaque droite de régression, pour chaque décennie et pour les trois échelles. L'objectif est alors de les interpréter en les confrontant à notre propre connaissance géohistorique de la croissance du réseau, étudiée à partir des *S-Curve* qui décrivent les trois phases qui nous sont maintenant bien connues *(cf. Figure 3. 14, p. 127)*. Pour rappel, la *Figure 7. 5. b.* permet de voir les rythmes de construction des voies selon leur hiérarchie.

Par exemple, entre 1860 et 1870, c'est la pente qui décrit l'accessibilité à la préfecture régionale qui a connu l'accentuation la plus forte. Les gains d'accessibilité sont très forts, bien qu'ils soient suivis de près par l'accessibilité à la capitale, alors que l'amélioration de l'accessibilité est bien moins importante pour rejoindre la préfecture départementale. Ainsi, dès la mise en connexité du réseau, l'accessibilité à Paris est encore améliorée, mais la croissance du réseau intermédiaire influence fortement l'accessibilité à l'échelon régional.

En %	a. Evolution des pentes des droites de régressions entre le temps de parcours sur le graphe P (situation contrefactuelle) et sur le multigraphe F (situation réelle)		
	Accessibilité à la capitale Paris	*Accessibilité à la préfecture régionale*	*Accessibilité à la préfecture départementale*
1860-1870	46,49	**57,73**	11,82
1870-1880	17,89	**30,25**	11,75
1880-1890	14,71	**25,73**	20,92
1890-1900	12,25	**14,47**	13,98
1900-1910	9,02	13,31	**13,51**
1910-1920	**6,71**	4,38	3,34
1920-1930	**9,33**	1,84	-1,18

b. Création de voies nouvelles selon le type de voies



Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Figure 7. 5. Evolution des rythmes d'accélération selon l'échelle et le temps face à la construction de voies nouvelles selon le type de voies

Par la suite, l'évolution des gains d'accessibilité est relativement moins forte, même si les écarts inter-échelles restent importants. Jusqu'en 1900, l'amélioration de l'accessibilité à la préfecture régionale conserve le haut du pavé, alors que les gains pour rejoindre Paris deviennent de moins en moins importants, en dépit de la construction de nouvelles infrastructures rapides entre 1880 et 1890. Ainsi la phase d'inflexion des réseaux – 1878 pour l'ensemble des réseaux, entre 1878 et 1883 quand on regarde chaque type de réseau – est propice à l'accessibilité à l'échelle régionale. Ce n'est que dans la seconde phase de croissance que les gains semblent s'homogénéiser à toutes les échelles : la mise en œuvre tardive du plan Freycinet fait que l'échelle la plus locale ne bénéficie des gains d'accessibilité les plus forts qu'à l'aube du XXème siècle, alors que le plan a été décidé 20 ans plus tôt. Et encore, les gains à l'échelle régionale restent très proches, ce qui la place décidément comme la grande gagnante de la phase de croissance du réseau.

La phase de stabilisation du réseau, à partir de 1910, fait considérablement chuter les gains d'accessibilité, et ceux qui repartent le plus à la hausse sont ceux qui vont vers la capitale, profitant du regain d'intérêt pour les voies rapides, alors même que la phase très éphémère de maturité du réseau provoque des premières pertes à partir de 1920 à l'échelle locale. Alors, si l'on admet que la croissance de voies lentes n'a qu'un impact modéré sur de longues distances, il est plus contre-intuitif de voir que la croissance de voies rapides n'a qu'un effet limité sur de courtes distances. Cela nous pousse désormais à étudier la contribution de chaque type de réseau dans ces rythmes d'accélération aux trois échelles décrites.

Pour cela, nous nous appuyons sur le nuage de points qui compare la situation contrefactuelle et la situation ferroviaire. La *Figure 7. 6* montre l'exemple de 1910 pour les trois échelles d'étude. Si le nuage reprend toutes les communes françaises, nous les différencions par leur connexion à des voies rapides, intermédiaires, lentes, ou ne bénéficiant d'aucun réseau.

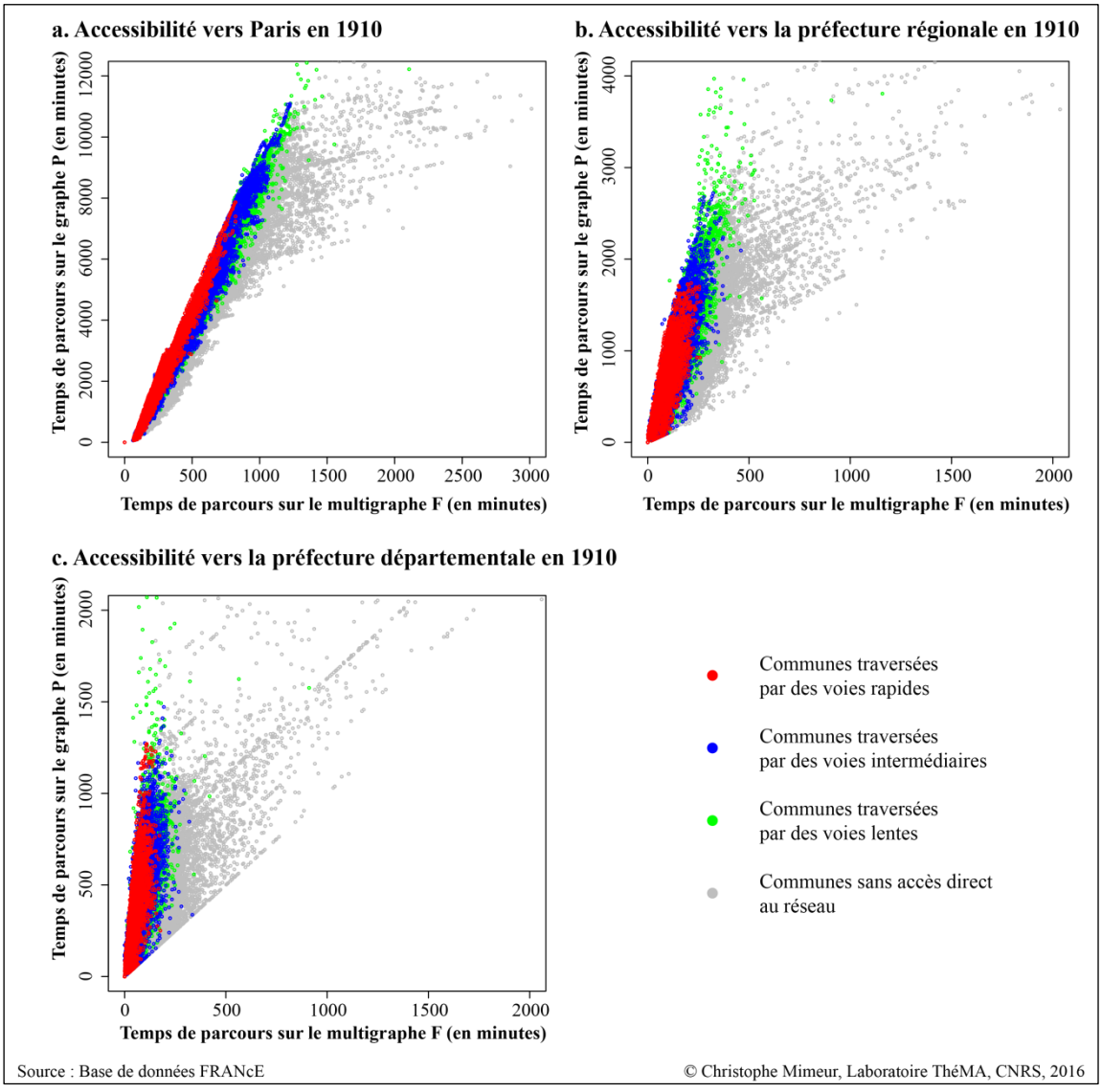


Figure 7. 6. La contribution de la hiérarchie du réseau aux gains d'accessibilité aux trois échelles en 1910

Il faut d'abord noter que si nous isolons ici la date de 1910, l'organisation est la même pour l'ensemble des autres décennies. Il s'agit ici de questionner la portée des réseaux, à la croisée de la situation contrefactuelle et de la situation ferroviaire. Concernant l'accessibilité à la capitale *(Figure 7. 6. a.)*, on remarque que la portée du réseau le plus rapide est large, et permet d'avoir naturellement les accessibilités les plus intéressantes. Logiquement, le réseau intermédiaire lui succède avec une portée encore plus forte, qui se traduit par une densité de points grandissante à des distances plus grandes. Il en est de même pour le réseau local, dont la portée est maximale.

Quand on descend d'échelle, on voit comment le réseau rapide n'a qu'un impact limité, à proximité des centres, à la fois dans la situation contrefactuelle et la situation ferroviaire *(Figure 7. 6. b.)*. Ce sont les communes à proximité de la préfecture régionale qui bénéficient du réseau rapide. La hiérarchie est encore une fois respectée, avec une forte densité du réseau intermédiaire dont la portée est presque doublée par rapport au premier. La portée du réseau local est encore forte, bénéficiant aux communes les plus éloignées, sans que les gains de temps soient forcément élevés.

A l'échelle départementale, la même hiérarchie est respectée *(Figure 7. 6. c.)*. La densité des points selon le type de voies atteste des gains permis par les voies rapides à proximité de la préfecture, alors que les voies intermédiaires ont une plus grande portée, bien qu'elles soient elles aussi très présentes à proximité du centre. La densité des communes bénéficiant d'un réseau lent est quasi identique, bien que leur portée soit encore une fois plus large.

Ainsi, on retient de cette organisation un phénomène hiérarchique qui s'applique à toutes les échelles. Ici, dans le cas de 1910, on voit comment le réseau rapide bénéficie d'abord aux communes les plus proches du centre étudié, attestant d'un gradient centre/périphérie dans la distribution territoriale de la hiérarchie du réseau. On peut donc ici conclure à une hiérarchie du réseau qui suit une logique centre/périphérie, ce phénomène étant de plus transcalaire, puisqu'il s'applique aux trois échelons que nous avons étudiés.

Cette organisation est en revanche plus ou moins marquée. Pour 1910, c'est bien à l'échelon régional que la succession de ces trois types de réseaux est la plus perceptible. Sans montrer l'ensemble des nuages de points pour l'ensemble de notre période, notre étude permet de voir que cette organisation hiérarchique et transcalaire est visible dès 1860. Les différents rythmes de construction du réseau, comme ceux des gains d'accessibilité selon l'échelle, n'ont pas d'impact sur la hiérarchie centre/périphérie du réseau. C'est la croissance de la portée des différents réseaux qui est la plus significative. A tous les niveaux, la distribution initiale du réseau bénéficie d'abord aux communes les plus proches des centres étudiés. C'est à partir de la date d'inflexion des différents réseaux que la hiérarchisation s'accentue en couvrant davantage de communes. Cet effet hiérarchique est prégnant pour l'accès à la capitale, puis l'expansion progressive des réseaux augmente leur portée, surtout intermédiaires et lents. Le fort ascendant hiérarchique des voies rapides tend à s'estomper en fin de période. Pourtant, il reste important aux échelons inférieurs, et c'est décidément à l'échelle régionale que la dimension hiérarchique est la plus perceptible, et qui demeure jusqu'en 1930. Cet état de fait est donc permanent et est même renforcé dans notre période d'étude, alors même que des notables locaux s'érigeaient dès le milieu des années 1860 contre une conception du réseau *« en fonction d'une logique des coûts qui aboutissait à construire un réseau fortement hiérarchisé, organisé autour de lignes principales sur le tronc desquelles venaient se*

brancher des voies secondaires, ou embranchements, mais aussi d'une logique de la centralité qui faisait converger ces axes principaux vers Paris » (Caron, 1997).

Ainsi notre étude polynucléaire participe ici à l'atténuation du mythe de l'homogénéisation pourtant perceptible par les déformations de l'espace. Si un phénomène hiérarchique était visible à l'échelle nationale, cette approche montre la permanence de la hiérarchie du réseau à tous les étages. Fortement centralisée depuis la capitale, la croissance du réseau ferroviaire français obéit toutefois à des gradients de centre/périphérie aux échelles nationales, régionales et locales. Ce phénomène transcalaire est particulièrement important à l'échelle régionale, laissant à penser à l'existence de systèmes « régionaux ». Sous couvert d'une visée nationale, l'idée d'une convergence en termes d'infrastructures sur l'ensemble du territoire participe alors à une hiérarchisation des échelons inférieurs. L'anamorphose multipolaire montre la convergence inter-régionale. Pour autant, la seconde partie de l'étude conclut au maintien – sinon au renforcement – des disparités intra-régionales, dans la mesure où la hiérarchisation demeure à tous les échelons analysés. Dès lors, il s'agit d'étudier à quels types de règles obéissent ces organisations hiérarchiques ? Sont-elles les mêmes à toutes les échelles ? Nous abordons les difficultés à lier la dotation en réseau avec la situation démographique locale.

7.2. Les effets structurants du réseau sur la structure démographique

Alors que la phase de croissance avait été définie de manière temporelle jusqu'à maintenant, la première section a mis l'accent sur les aspects spatiaux de l'accélération du réseau ferroviaire, qui a favorisé une convergence inter-régionale mais qui a maintenu des différences intra-régionales. Il s'agit alors d'étudier ces différences en les reliant à la structure démographique du territoire français. Ainsi, dans la seconde moitié du XIXème siècle et le premier tiers du XXème siècle, cela nous permet d'interroger l'affirmation de François Caron, stipulant que *« la détermination des tracés était devenue un enjeu politique majeur. Tous les responsables locaux avaient la certitude que l'avenir des communautés dont ils avaient la charge en dépendait »* (Caron, 1997). Au-delà des relations avec la proximité au centre fonctionnel, la construction d'une hiérarchie progressive du réseau à toutes les échelles répond-elle à une structure de population ? Celle-ci est-elle préexistante ou provoquée par le réseau ? Pour cela, nous voyons d'abord les liens que l'on peut établir entre les niveaux de population et les performances du réseau, en termes d'accélération à partir des indicateurs déjà construits dans la seconde partie et de structures du réseau à partir des propriétés du graphe.

7.2.1. **Accessibilité et niveaux de population : une coévolution**

Alors que les analyses précédentes font fi de la donnée démographique, nous confrontons ici la diffusion du réseau ferroviaire avec la desserte des populations. La *Figure 7. 7* confronte l'accumulation des populations desservies avec les indicateurs d'accessibilité aux centres administratifs, puis en dernier lieu en fonction du temps d'accès moyen à l'ensemble du territoire entre 1860 et 1930.

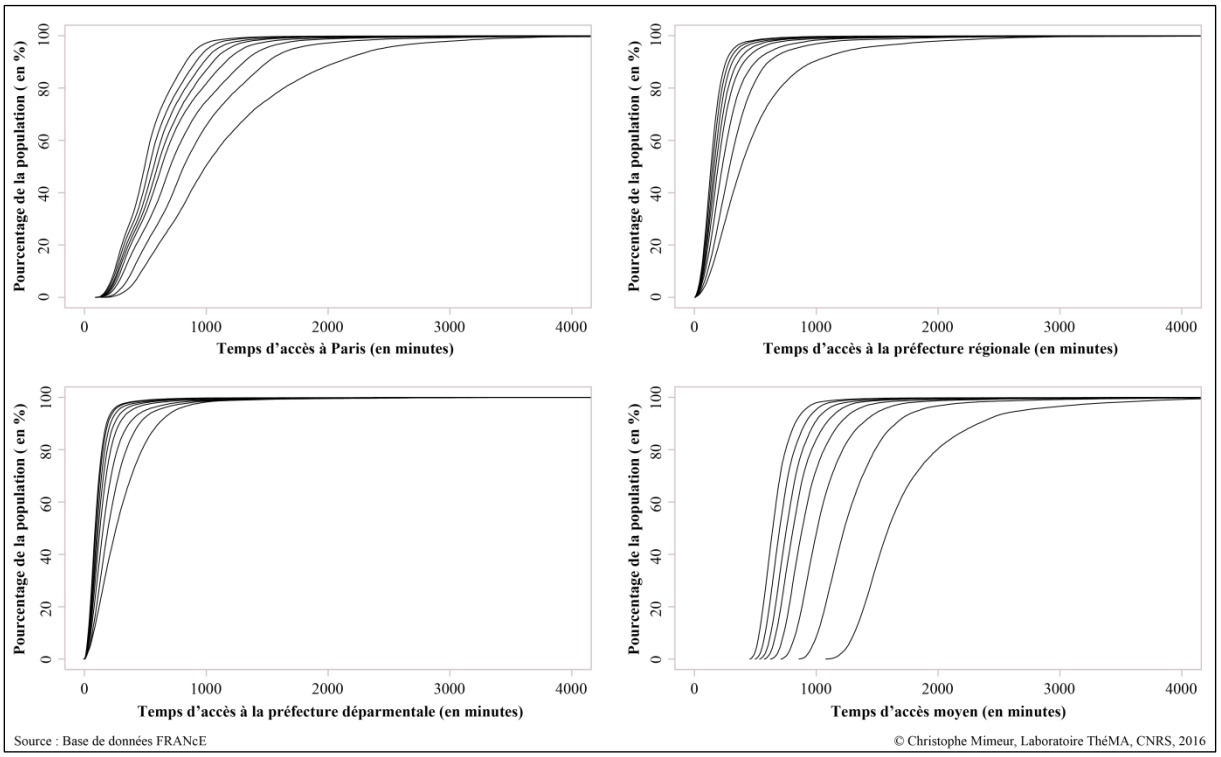


Figure 7. 7. La diffusion de la desserte des populations selon quatre indicateurs d'accessibilité entre 1860 et 1930

Ces quatre graphiques viennent confirmer les résultats de la section précédente, croisés cette fois avec la population. La phase d'accélération montre une succession des courbes de droite à gauche décennie après décennie. La première phase de chaque courbe montre une certaine horizontalité, qui traduit une concentration des populations bénéficiant des accessibilités les plus intéressantes. Ainsi, même si les résultats ne prennent ici pas en compte l'Île de France, parce que le poids de Paris est décidément trop important, alors les courbes confirment ici une diffusion centre/périphérie de l'accessibilité. En 1860, seulement 4 % de la population est située à moins de 6 heures de Paris, alors qu'un Français sur 6 peut rejoindre Paris en moins de 6 heures en 1930. L'inflexion de la phase de diffusion, traduite par la pente de la courbe, montre qu'elle est de plus en plus rapide, et que le seuil est de plus en plus proche du centre, traduisant une convergence de l'accessibilité. Cet état de fait est de plus confirmé par la diffusion de l'accessibilité moyenne. Le graphique montre que la concentration des meilleurs indices pour une petite part de la population

est de moins en moins importante, et que la phase de décollage est de plus en plus rapide dans la distribution des accessibilités. Par ailleurs, alors que nous avons défini des rythmes dans les gains d'accessibilité en fonction des centres, ces graphiques montrent quoiqu'il en soit, que les progrès les plus forts pour joindre davantage de population sont surtout le fait des premières décennies de l'étude, alors que les changements les plus tardifs restent marginaux, surtout à partir de 1910. Cela dit, il s'agit désormais d'étudier quelles populations bénéficient le plus de la vitesse et du réseau ? Est-ce que cela concerne plutôt les villes au détriment des campagnes ? Est-ce que cela répond à des choix d'aménagement, pris à quelles échelles ?

Pour cela, nous proposons de confronter les différents indicateurs d'accessibilité construits lors de la complexification de la dotation en réseau. Notons d'ores et déjà que cette confrontation conserve les mêmes caractéristiques d'une décennie à l'autre, c'est pourquoi nous prenons ici l'exemple isolé d'une décennie pour comprendre comment ils nous renseignent davantage sur la desserte des populations. La *Figure 7. 8* confronte le temps d'accès moyen de chaque commune à deux complexifications de l'indicateur. Dans les nuages de points obtenus, on distingue les communes qui bénéficient d'une gare de celles qui n'en possèdent pas.

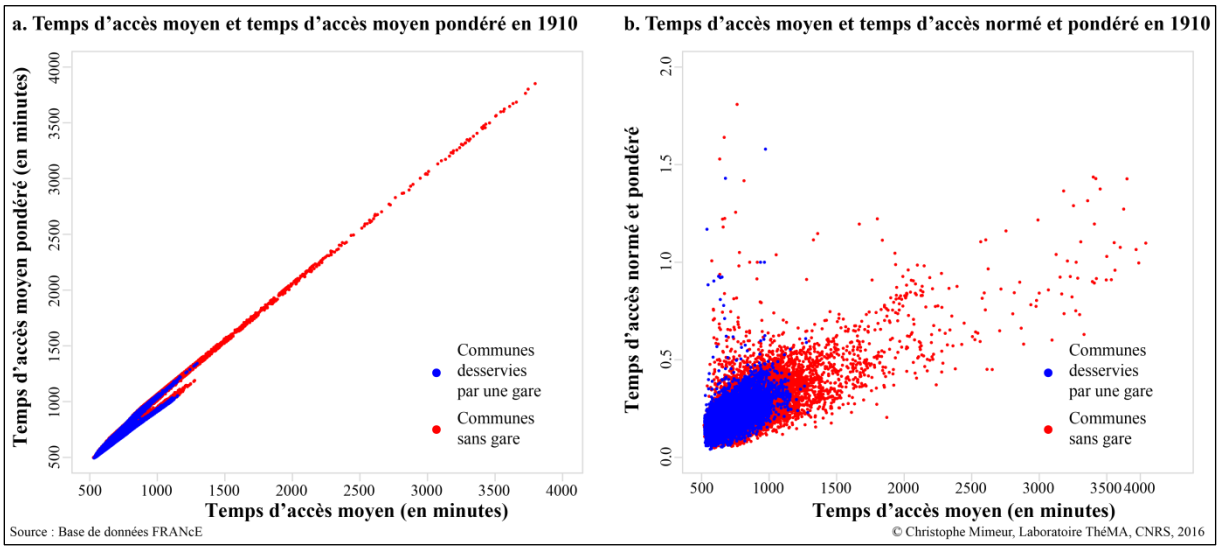


Figure 7. 8. La confrontation des indicateurs d'accessibilité en 1910

D'abord, la *Figure 7. 8. a.* confronte le temps d'accès moyen au temps d'accès moyen pondéré par la population. Théoriquement, il s'agit du temps de parcours moyen d'un Français pour rejoindre un autre Français. Il n'est pas surprenant de voir une forte corrélation, supérieure à 99 %, entre les deux. Si les communes possédant une gare ont une meilleure accessibilité, certaines autres communes, très proches, bénéficient du réseau. Cet effet est conservé voire renforcé par la pondération par la population : à proximité de noyaux de population, on va retrouver une plus forte densité de communes avec un haut niveau de population. Ainsi, l'indicateur pondéré permet de traduire l'effet polarisant des foyers de population, rendant compte d'une bonne accessibilité avec ou sans la présence d'un réseau, à condition d'être à proximité d'un centre.

Le nuage de points obtenu dans la *Figure 7. 8. b.* est en revanche plus difficile à saisir. L'axe des ordonnées représente le temps d'accès pondéré par la population et normé par une situation contrefactuelle. Le seul indicateur normé a fait l'objet d'une analyse dans la seconde partie *(cf. 5.1.2, p. 183)*. Nous ne l'avons pas représenté ici, mais il existe une forte corrélation entre le seul indicateur normé et cet indicateur normé et pondéré. Cela montre que les communes qui bénéficient d'un effet vitesse important sont celles qui permettent de rejoindre davantage de population. Ainsi, ces communes sont souvent celles qui sont proches de foyers de population intermédiaires. Dans la *Figure 7. 8. b.*, on voit donc que par rapport au seul temps d'accès moyen, cet indicateur complexifié permet davantage de faire une distinction entre les noyaux de la population, et ceux qui bénéficient de leur proximité, indépendamment qu'ils aient une gare ou non. Ainsi, la forme conique qui s'étire vers le haut définit des communes qui bénéficient d'une accessibilité, parce qu'elles sont proches d'un foyer de population, même sans desserte directe. Nous aurons donc à mesurer les effets directs et indirects du réseau dans la suite de notre raisonnement[42].

Ainsi, cette première étape montre que l'apport des indicateurs complexifiés est limité quand on les confronte un à un. Pour définir l'accessibilité de noyaux de population, la corrélation semble élevée avec la possession d'une gare, que l'indicateur soit ou non pondéré par la population. En revanche, pour qualifier l'accessibilité dans les campagnes, le dernier indicateur semble le plus évocateur des situations diverses, parmi lesquelles la proximité d'un noyau de population tient le haut du pavé, à défaut de posséder une gare, même si aucune forme caractéristique des nuages de points ne peut se dégager. Pour autant, il s'agit de voir si ces performances, pour les villes et les campagnes, se traduisent dans les niveaux de population des communes desservies, en interrogeant les caractéristiques de la desserte. En effet, les principales limites aux indicateurs multipolaires, assortis de la complexité qu'on a introduit, sont leur forte dépendance aux effets de bord d'une part, à la signification et l'interprétabilité géohistorique d'autre part. Pour autant, ils permettent une approche non-paramétrique de l'accessibilité : il s'agira alors de les confronter à la recherche d'une explication causale entre infrastructure et croissance démographique.

Pour conduire ces investigations, nous choisissons de classer l'ensemble des communes par quartiles d'accessibilité, traduite par le temps de parcours moyen vers toutes les autres communes : des 25 % les plus accessibles jusqu'à la totalité des communes. Dans quelle mesure peut-on affirmer qu'une catégorisation des niveaux de population est sensible à cette distribution de l'accessibilité ? En prenant l'ensemble des communes françaises, la forte hétérogénéité déjà constatée dans les premières analyses du chapitre ne fait ressortir aucune structure, bien que les valeurs extrêmes aient été retirées de l'analyse pour chaque décennie.

Nous choisissons alors de distinguer villes et campagnes, sans pour autant présager d'un seuil les séparant *a priori*. Dans la démarche abductive du travail, nous déterminons ce seuil comme un paramètre de l'analyse, permettant d'interroger les logiques d'expansion du réseau alors que la pression de l'Etat est grandissante envers les compagnies à partir des conventions de 1865 pour étendre coûte que coûte le réseau ferroviaire (Caron, 1997). Ainsi, des notables locaux dénoncent rapidement la logique de coûts dont l'objectif est la desserte de populations denses au détriment des

[42] Ces indicateurs ont davantage de pouvoir descriptif qu'explicatif : une autre approche sera envisagée dans la prochaine section pour qualifier les effets directs et indirects du réseau.

campagnes. A partir de tests de plusieurs seuils[43], c'est dans les communes au-delà de 2 500 habitants qu'une structuration hiérarchique en fonction de niveaux de population se dessine pour le niveau supérieur : ce seuil est valable pour toutes les décennies de l'étude, ce pourquoi il est conservé dans la *Figure 7. 9*. Il met en avant une hétérogénéité dans la relation entre transport et population.

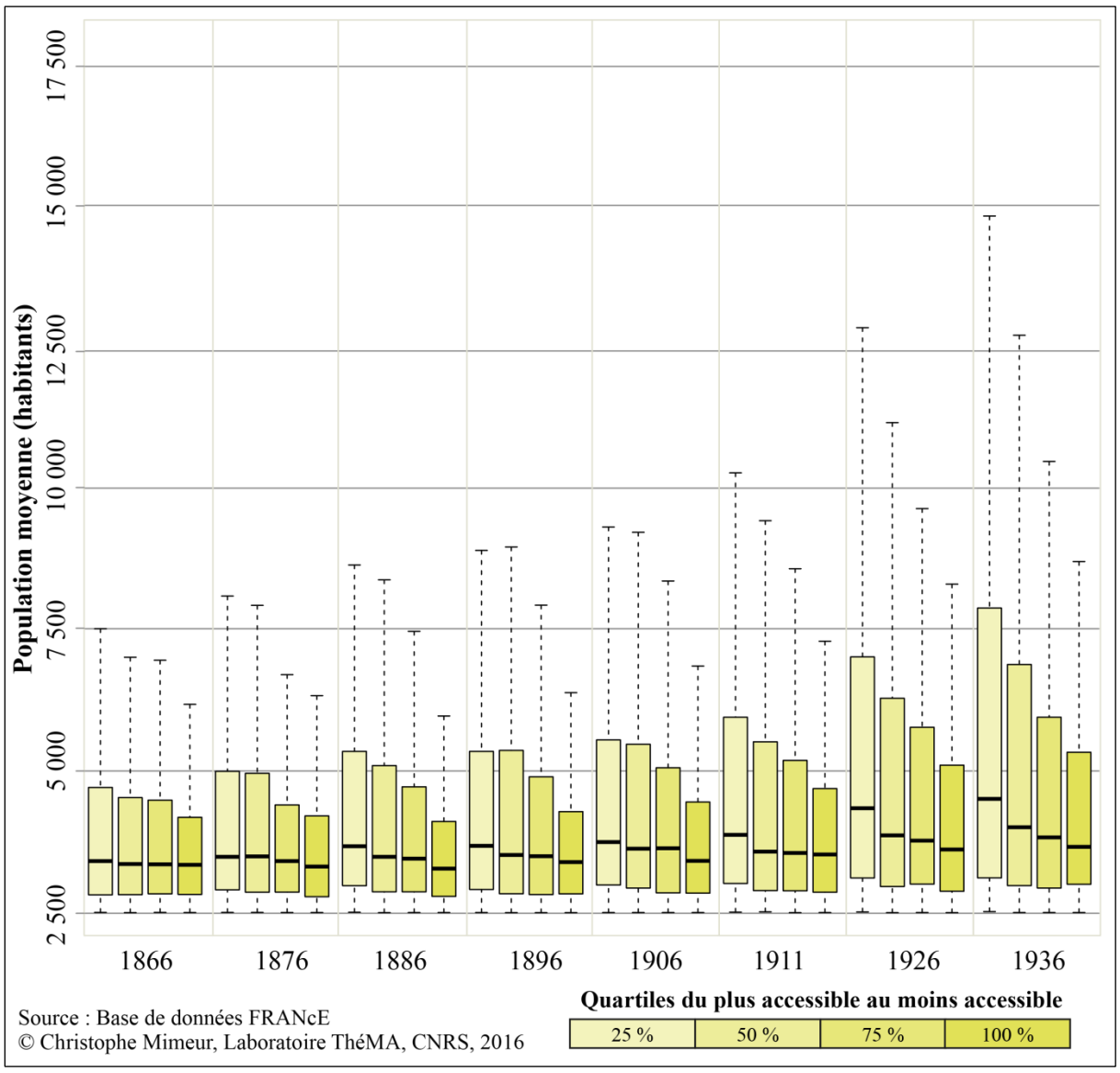


Figure 7. 9. Les niveaux de population de villes supérieures à 2 500 habitants en fonction des quartiles d'accessibilité

[43] Les tests de chi-deux permettent de mesurer la significativité d'une relation entre deux paramètres à un seuil de confiance de 95 % : en-dessous du seuil de 2 500 habitants, aucun des deux sous-échantillons ne présente des résultats significatifs.

La distribution des accessibilités semble suivre la hiérarchie urbaine préexistante (Pumain, 1982), conformément aux conclusions tirées par Denise Pumain, qui utilise d'ailleurs le même seuil pour définir les individus de son analyse. Pour autant, on voit que cette hiérarchisation semble croissante, dans la mesure où les différentiels sont de plus en plus forts au cours du temps, parallèlement à l'urbanisation croissante et donc à l'augmentation des taille de villes. Ainsi, parmi le premier quart des communes les plus accessibles, on va retrouver forcément les villes plus peuplées, quelle que soit la décennie. Pourtant, les valeurs moyennes restent proches. La rupture avec les deux premiers quartiles n'est pas évidente au début de la période mais le devient à mesure que le réseau s'étend, attestant de la reproduction de la hiérarchisation de la structure urbaine dans la hiérarchisation du réseau, et surtout à partir de la stabilisation du réseau en 1910. A ce stade, on pose alors l'hypothèse que le rattachement précoce au réseau participe à l'accroissement des villes, ce qu'il faudra vérifier avec les taux de croissance démographique, tout en tenant compte de la croissance démographique de fond. Pour autant, il faut garder à l'esprit que parmi le premier quartile, l'amplitude de la taille des villes reste la plus importante, attestant que des petites villes bénéficient elles aussi de ces bons résultats : l'adaptation à la structure urbaine semble attestée, mais semble aussi profiter à d'autres de manière collatérale, ce qui nous amène à questionner ce même type de distribution dans les campagnes.

La *Figure 7. 10* présente les niveaux de population selon les quartiles d'accessibilité multipolaire[44] pour les communes inférieures à 2 500 habitants. On remarque alors que leur nombre important participe à l'absence de structures quand on analyse les quelques 36 000 communes dans leur ensemble. En effet, dans les campagnes françaises entre 1860 et 1930, il est difficile d'établir une hiérarchie de tailles de communes conformément à leur accessibilité, même en prenant en compte leur proximité à des foyers de population. L'hétérogénéité n'est plus seulement à saisir dans la hiérarchie urbaine mais aussi dans la diversité des effets, dont il est difficile de faire ressortir des généralités. Certains résultats semblent même contre-intuitifs dans la mesure où certaines communes mal desservies jusqu'en 1896, ont une taille moyenne plus élevée que les communes rurales les mieux desservies. L'ensemble de ces résultats doit être considéré toutefois avec beaucoup de prudence dans la mesure où l'amplitude de chaque classe est similaire et très élevée, même si un semblant de hiérarchie semble se dégager une fois que le réseau est stabilisé après la première guerre mondiale. Alors on pourrait poser l'hypothèse que la desserte des campagnes françaises pourrait être frappée de contingence, en position intermédiaire entre deux villes à relier, même si l'expansion du réseau participe à des tracés réfléchis de plus en plus localement : ainsi, quand il est planifié davantage, par le plan Freycinet ou par des initiatives locales, le réseau participerait à un phénomène de hiérarchisation. Pour autant, la contingence ne saurait dicter un raisonnement géographique, et même géohistorique.

[44] Les mêmes analyses ont été conduites à partir des indicateurs normés et pondérés et présentent le même type de résultats.

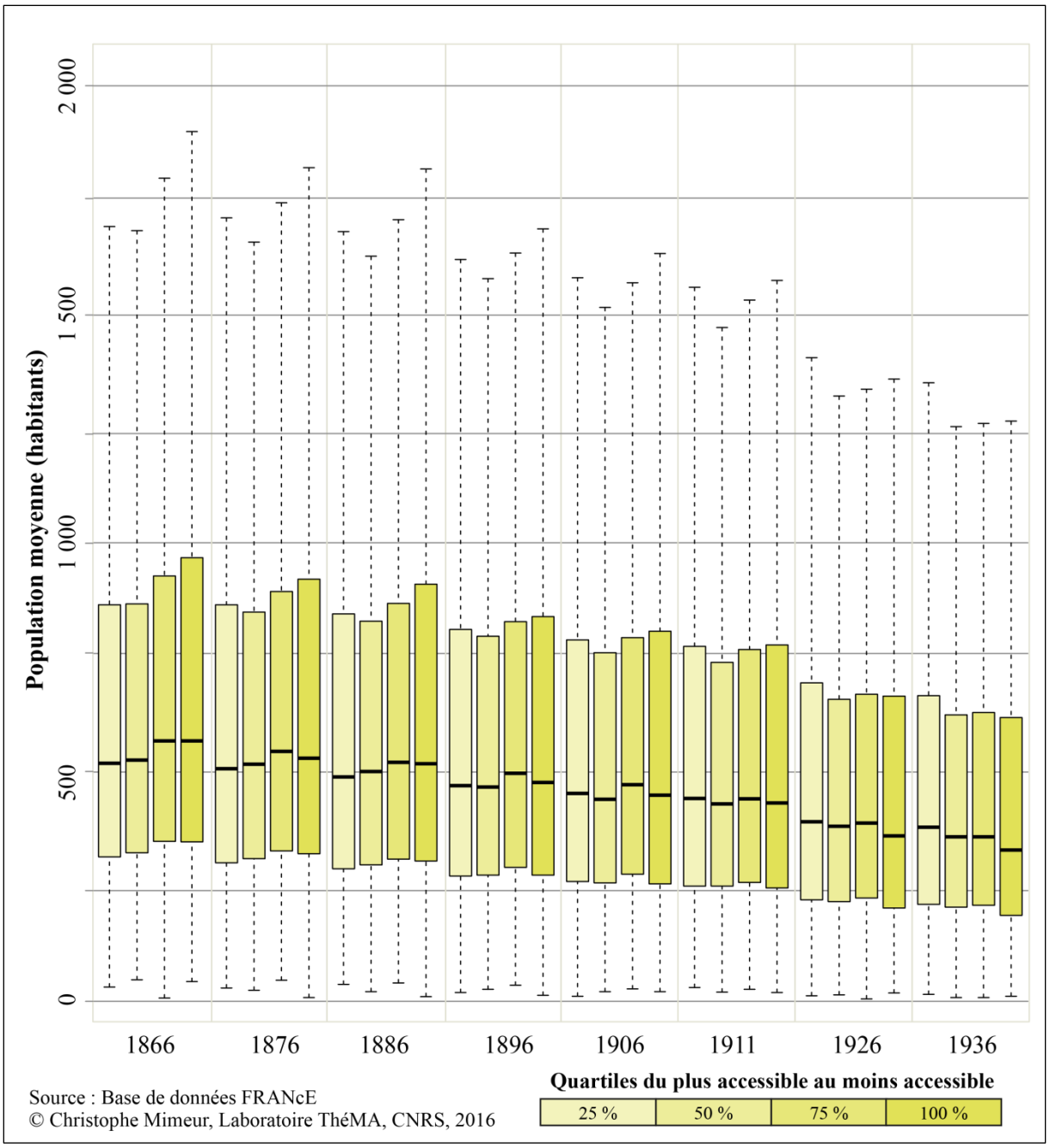


Figure 7. 10. Les niveaux de population des communes inférieures à 2 500 habitants en fonction des quartiles d'accessibilité

Parce que la contingence est un raccourci facile dans l'explication géohistorique, nous nous efforçons à ce stade à continuer nos investigations dans la recherche d'effets structurants de l'infrastructure de transport, en introduisant progressivement des éléments de contextes locaux, endogènes au réseau pour l'instant. Puisque la distribution des accessibilités propose de larges amplitudes dans la taille des communes, nous proposons dans la suite de la section de nous intéresser davantage à deux autres types d'informations : une qui concerne le réseau, à partir d'une étude de la centralité des gares. L'autre propose de compléter les niveaux de population par la croissance démographique locale.

7.2.2. **Centralité et niveau de population : une double coévolution**

Nous venons de voir à quel point il est difficile d'établir une relation entre le niveau de population et l'accessibilité des communes inférieures à 2 500 habitants, alors qu'au-delà, une adaptation mutuelle entre population et accessibilité semble patente, en dépit de fortes amplitudes dans chacune des classes. Nous proposons ici de nous intéresser de manière plus qualitative aux liens entre réseau et niveau de population, pour qualifier autrement l'arbitrage entre homogénéité, hiérarchie et hétérogénéité. Cela répond à une autre complexification de la dotation en réseau, par ailleurs déjà étudiée dans la littérature traitant de la croissance urbaine et des réseaux de transport (Bretagnolle, 2003, 2009 ; Pumain, 1982). Ainsi, la construction d'indicateurs de desserte répond au principe du *« comptage, ville par ville, du nombre d'axes desservant le lieu et de la qualité de l'infrastructure, aboutissant à une typologie de l'ensemble des villes en différents niveaux d'accessibilité »* (Bretagnolle, 2003). L'étude est toutefois étendue à l'ensemble des communes françaises qui, s'appuyant sur les conclusions des paragraphes précédents, reprend la distinction entre villes supérieures à 2 500 habitants et les autres, de sorte de traiter au mieux l'hétérogénéité de chacune des catégories. En revanche, seules les communes desservies sont prises en compte dans cette sous-section.

Dès lors, nous proposons d'abord de confronter le nombre de lignes divergentes aux niveaux de population. Pour cela, nous utilisons une autre discrétisation de la population que la distinction entre petites, moyennes et grandes villes proposée dans des travaux antérieurs (Bretagnolle, 2009). Alors que la distribution de la population conserve sur la longue durée une forme dissymétrique, nous utilisons une discrétisation de la population par écart-type, après transformation en logarithme, de sorte d'avoir une distribution log-normale : la lecture des boîtes à moustaches se fait de gauche à droite, des communes les moins peuplées aux communes les plus peuplées. Comme pour les données d'accessibilité en distance-temps, la prise en compte globale de l'échantillon des 36 000 communes ne propose pas la moindre structure pour comprendre les relations. La *Figure 7. 11* présente les relations pour les villes supérieures à 2 500 habitants entre 1860 et 1930.

Avoir une ligne divergente signifie avoir le statut de terminus d'une ligne, tandis qu'en avoir deux signifie être une simple étape. Au-delà nous avons des carrefours qui deviennent de plus en plus complexes. Les points sur la *Figure 7. 11* sont des valeurs extrêmes dans la distribution des lignes divergentes. L'indicateur de desserte présente ici davantage de contraste qu'avec les indicateurs multipolaires. Entre 1866 et 1886, les deux premières classes de villes présentent sont surtout des étapes, alors que quelques-unes sont des terminus ou des carrefours à 3 branches. Dès 1866, les villes les plus peuplées sont celles qui ont des carrefours complexes, où 75 % de ces villes-là ont au moins des carrefours à 3 branches. Cette hiérarchisation est continue tout au long de notre période. Après 1886, on observe une distinction nette entre les villes les plus peuplées qui ont en moyenne un carrefour à 4 branches, et les villes les moins peuplées qui restent de simples étapes. Entre elles, une classe intermédiaire se constitue avec des amplitudes plus grandes, où certaines villes moyennes deviennent des carrefours importants. La corrélation entre ces deux indicateurs est significative avec un seuil de confiance de 99 %.

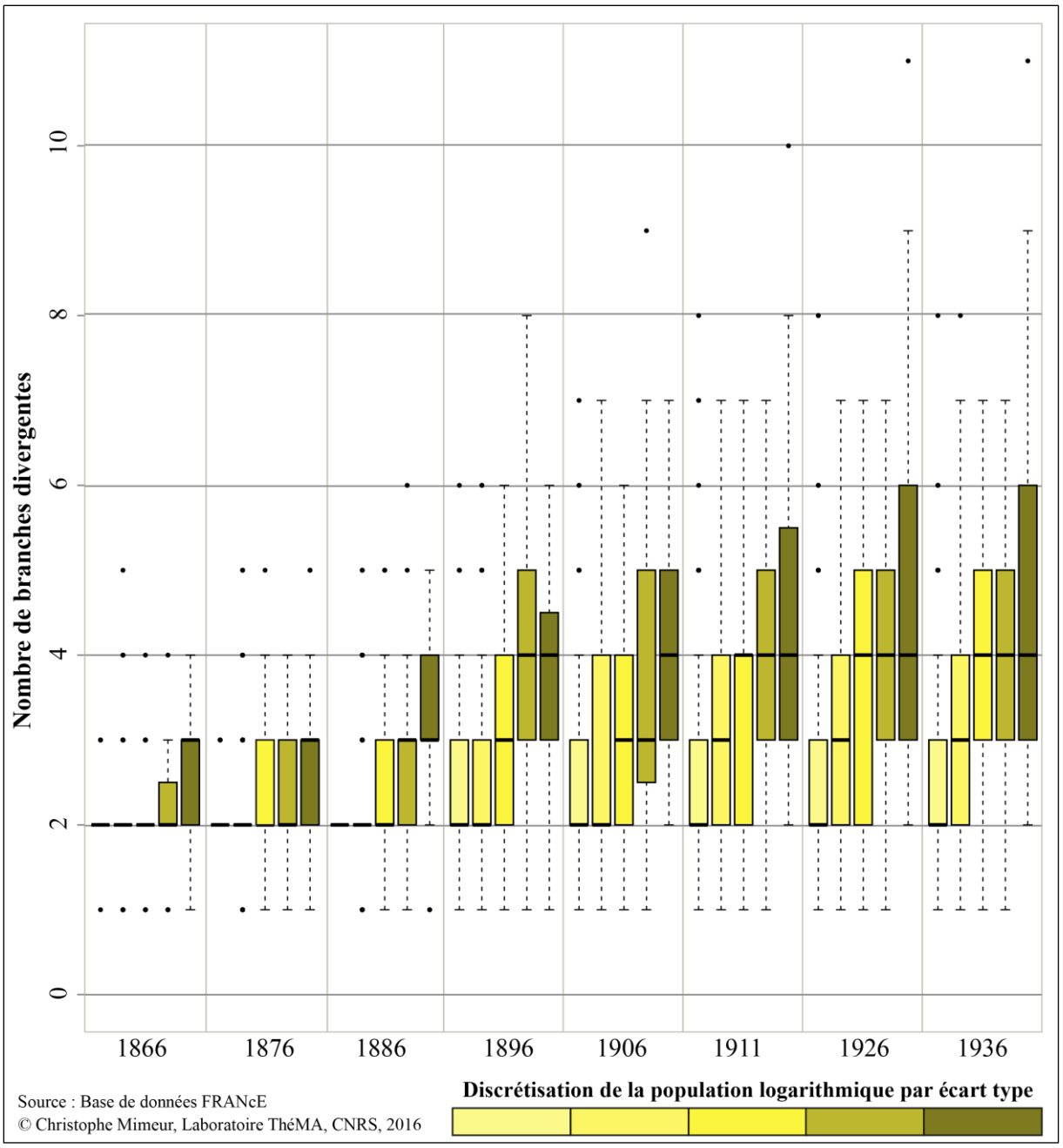


Figure 7. 11. Carrefours et niveaux de population pour les villes ayant plus de 2 500 habitants entre 1860 et 1930

Ainsi, cette figure montre deux phénomènes concomitants : une relation parallèle entre la complexité du carrefour et le niveau de population et une relation plus étonnante dans laquelle des villes sont à une position stratégique sur le réseau. Il s'agira de voir si ces villes-là bénéficient plus spécifiquement de l'effet du réseau sur sa croissance locale, et sur sa position ultérieure dans le réseau.

En-dessous de 2 500 habitants, la distribution entre branches divergentes et niveau de population des campagnes participe à un phénomène hiérarchique, bien que parmi celles qui sont desservies par le réseau, la grande majorité des communes rurales ne sont que des étapes et seules

quelques situations exceptionnelles en font des carrefours. Il n'a pas été choisi de les représenter, parce que seuls des points extrêmes répondent à une hiérarchie.

Ainsi, cette section portant sur la centralité des réseaux apporte davantage de significativité quant aux niveaux de population que l'approche par l'accessibilité multipolaire. L'effet direct de l'infrastructure semble jouer davantage que les effets indirects de l'accessibilité, élément qui sera testé par la modélisation économétrique plus loin dans nos propos. Les résultats semblent attester que la desserte d'une commune par un réseau est un facteur très discriminant, plus que les temps de parcours, quand on confronte simplement accessibilité et population, davantage pour les campagnes que pour les villes. Pour autant, les résultats semblent complémentaires : on observe un phénomène de coévolution entre croissance du réseau de transport ferroviaire entre 1860 et 1930 et structure urbaine existante. Décennie après décennie, cette structure semble évoluer de manière conjointe, alors que la différenciation entre les différentes classes d'accessibilité ou de population semble s'accentuer. Si, à la manière du réseau postal du XVIIIème siècle, on constate une *« imbrication de plus en plus étroite entre le réseau et la trame urbaine, selon une sorte de processus d'adaptation mutuelle des deux formes »* (Bretagnolle, 2009), le réseau ferroviaire français participe quant à lui à une double hiérarchie, en réponse à la hiérarchie urbaine préexistante :

- Une hiérarchisation du réseau, qui reproduit une logique centre/périphérie, mais dont la caractérisation la plus forte est sensible au niveau régional ;
- Une hiérarchisation des centres du réseau, qui semble correspondre à une structure urbaine, perturbée par la création de carrefours stratégiques dans des villes moyennes.

Alors, les processus qui mêlent ces deux types de hiérarchisation appellent à se pencher sur le lien entre la croissance du réseau et la croissance démographique locale, en introduisant progressivement des aspects liés aux contextes locaux, de sorte de réintroduire progressivement l'ensemble des communes françaises, sans distinction.

7.2.3. La double hiérarchisation du réseau face à la croissance démographique locale

Afin d'appréhender la relation entre la centralité du réseau, l'accessibilité en distance-temps et la croissance démographique locale, nous nous inspirons de la typologie des centralités définie par Anne Bretagnolle, que nous adaptons à notre base de données, pour proposer trois aspects complémentaires :

- L'étude est étendue à l'ensemble du réseau ferroviaire entre 1860 et 1930 qui prend ainsi en compte une hiérarchie évolutive des vitesses le long du réseau ;
- L'étude est étendue à l'ensemble des communes françaises, villes d'un côté et communes ayant une population inférieure à 2 500 habitants d'un autre ainsi que

celles qui ne sont pas desservies par un réseau ;

- L'étude permet un croisement des résultats obtenus avec les indicateurs d'accessibilité purement quantitatifs, en termes de distances-temps, qui permet d'étudier les configurations spatiales des centralités du réseau, en lien avec la croissance démographique.

La *Figure 7. 12* montre comment nous avons adapté la typologie de la littérature (Bretagnolle, 2003) avec la base de données FRANcE. Nous distinguons ainsi les communes qui ne sont pas desservies par le réseau, celles qui constituent un terminus, celles qui sont une étape le long d'une ligne et celles qui sont situées sur un carrefour de lignes. L'information sur la vitesse nous permet par ailleurs de différencier les écarts et les carrefours simples selon qu'ils sont sur un réseau rapide, intermédiaire ou lent.

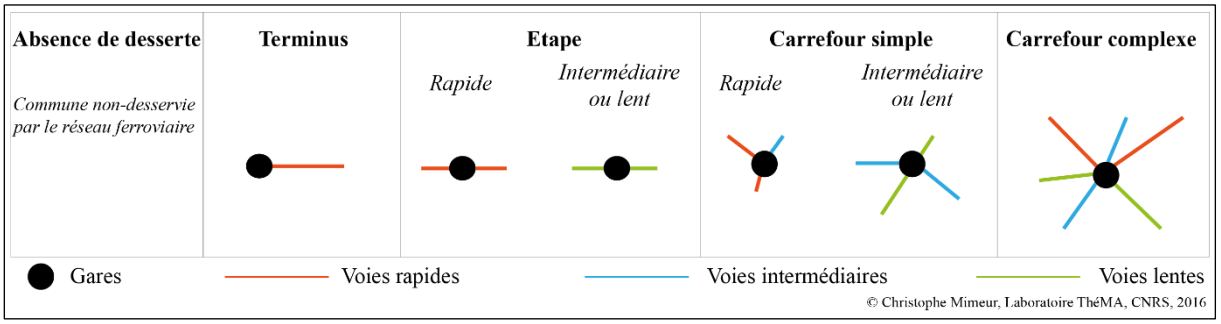


Figure 7. 12. Caractériser la desserte par la vitesse et la centralité

Puisque cela a déjà été évoqué souvent dans la littérature, nous débutons l'analyse par les effets conjoints de la centralité des gares sur la croissance démographique et sur les indicateurs d'accessibilité. Afin d'analyser les effets cumulatifs du réseau, nous étudions la croissance démographique entre deux recensements à partir de la situation du réseau au début de chaque période intercensitaire, limitant l'endogénéité. La croissance s'entend ici par le taux de croissance annuel moyen. Nous nous limitons de plus à la croissance jusqu'au recensement de 1911, parce que les pertes dues à la Première Guerre Mondiale sont un facteur exogène trop important pour être ignoré dans le couple réseau/territoire. Nous définissons trois décennies à comparer :

- Entre 1866 et 1876, quand le réseau d'intérêt national est connecté ;
- Entre 1886 et 1896, quand la date d'inflexion du réseau a effectivement eu lieu ;
- Entre 1906 et 1911, quand le réseau du plan Freycinet connaît son apogée.

La *Figure 7. 13* confronte dans un même nuage de points le taux de croissance annuel moyen pour chacune des trois périodes avec le temps de parcours moyen vers toutes les autres communes. Par ailleurs, chaque commune est caractérisée par son niveau de population en début de période ainsi que par sa place dans la structure dans le réseau.

On constate d'abord un fort différentiel entre les villes qui possèdent une gare et celles qui n'en ont pas. C'est spécialement le cas entre 1886 et 1896 où ces villes connaissent un déclin de 0,1 % par an en moyenne. Pour les deux autres décennies, la différence entre les deux cas est au moins de 0,3 points entre la croissance d'une commune desservie d'une autre qui ne l'est pas. Par ailleurs, les taux de croissance annuel moyen semblent suivre la taille des villes. Pour les trois décennies de l'étude, les villes qui ont les plus hauts niveaux de population bénéficient du phénomène d'urbanisation tandis que, comparativement, ce sont les plus petites villes qui perdent de leur population. L'analyse successive des nuages de points suggèrent aussi que le réseau influence la croissance démographique à travers deux phénomènes qui coexistent.

Le premier est constaté par la forme générale de la distribution du nuage de points. Il suggère un rôle de la vitesse le long du réseau. Les villes qui bénéficient d'une accessibilité efficace sont celles qui ont le niveau le plus important, conformément aux conclusions que nous avons déjà faites. Il semble que ces mêmes villes bénéficient d'une croissance démographique plus forte et positive. En revanche, à mesure que l'on avance dans le temps, on n'observe pas forcément un rapprochement des villes les plus peuplées. Il semblerait qu'on assiste au maintien d'un espacement des grands pôles, participant à une décentralisation du réseau : une étude de l'évolution de l'accessibilité polynucléaire propose pour ces villes-là un temps de parcours moyen vers Paris qui ne diminue plus. Cela vient renforcer la reproduction de la hiérarchie à des échelles inférieures, à partir même de ces centres : alors que la place dans le réseau ne semblait pas déterminante en début de période dans une aire relativement réduite, alors elle le devient peu à peu. Cette étude nous amène à réfléchir à l'effet du temps long sur la trajectoire démographique, en lien avec la dotation en infrastructures. Par la modélisation économétrique que nous avons entreprise dans la seconde partie de la thèse, nous serons en mesure d'investiguer la force de ces effets.

Le second est l'effet de la centralité des gares : les villes aux carrefours les plus complexes ont en moyenne une croissance supérieure aux autres, indépendamment de leur niveau de population. Les communes qui sont une simple étape sur une ligne connaissent une meilleure croissance si elles se trouvent sur un réseau rapide. Cette même hiérarchie est conservée sur les trois périodes. En revanche, la position de terminus a un comportement plus particulier : ces communes-là présentent alors une croissance forte, dont l'ampleur ne s'amenuise pas dans les deux décennies d'étude suivantes, elles correspondent à de grandes villes dont la connexion a été décidée dans les plans dessinant le réseau d'intérêt national. Le renforcement de la centralité que l'on constate sur le nuage de points suggère un renforcement de son rôle dans la croissance démographique locale. Là encore, la modélisation économétrique qui suit dans ce chapitre viendra corroborer ces premières constatations.

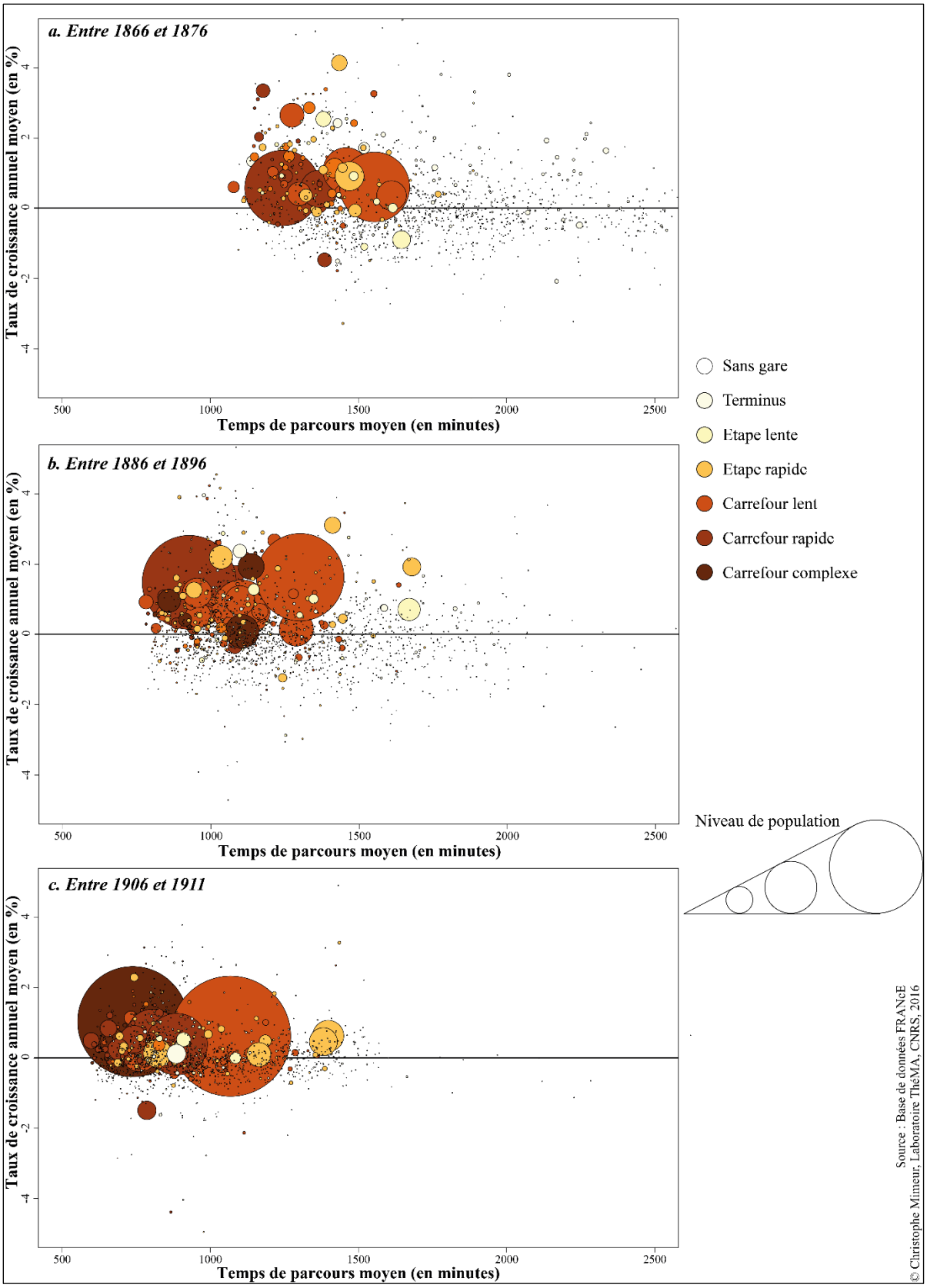


Figure 7. 13. Centralité, niveau de population et croissance démographique des communes de plus de 2 500 habitants entre 1866 et 1911

		TCAM (en %)	*Ecart-type*	Accès à la préfecture régionale (en minutes)
1866-1876	Sans gare	-0,37	*1,10*	498,20
	Terminus	1,39	*3,04*	200,26
	Etape lente	-0,04	*1,29*	150,71
	Etape rapide	0,05	*1,49*	112,56
	Carrefour lent	0,27	*1,48*	129,63
	Carrefour rapide	0,89	*1,98*	91,62
	Carrefour complexe	-		-

		TCAM (en %)	*Ecart-type*	Accès à la préfecture régionale (en minutes)
1886-1896	Sans gare	-0,61	*1,06*	297,84
	Terminus	0,03	*1,29*	159,96
	Etape lente	-0,33	*1,04*	139,62
	Etape rapide	-0,12	*1,60*	93,30
	Carrefour lent	-0,15	*1,16*	124,28
	Carrefour rapide	0,19	*1,41*	70,25
	Carrefour complexe	0,53	*1,15*	66,14

		TCAM (en %)	*Ecart-type*	Accès à la préfecture régionale (en minutes)
1906-1911	Sans gare	-0,37	*0,85*	232,27
	Terminus	-0,01	*0,81*	154,23
	Etape lente	-0,24	*0,80*	132,67
	Etape rapide	-0,09	*0,77*	88,23
	Carrefour lent	-0,08	*1,01*	119,94
	Carrefour rapide	0,10	*1,09*	77,25
	Carrefour complexe	0,85	*2,78*	71,21

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 1. Centralité dans le réseau et croissance démographique locale des communes de moins de 2 500 habitants

L'étude de la *Table 7. 1* interroge en revanche le phénomène d'exode rural, pour les communes de moins de 2 500 habitants. La même représentation graphique que précédemment n'est pas efficiente : les niveaux de population ne sont pas assez discriminants et aucune forte structure ne ressort encore. Si la présence même d'une gare est l'élément déterminant de la perte de population entre 1876 et 1886, les deux périodes suivantes montrent que même des communes desservies par le réseau sont touchées. Cet état de fait vient d'abord confirmer le besoin de complexifier l'approche du réseau par d'autres variables, tant on voit que les écarts-types présentent des valeurs élevées. Ainsi, globalement, la moyenne des pertes de population tend à s'amenuiser à mesure que la gare communale monte dans la hiérarchie de la centralité. Pourtant, à partir de 1886, où le phénomène s'intensifie, la tendance montre qu'être au cœur d'un carrefour avec des voies intermédiaires ou lentes ne signifie pas être mis à l'écart de l'Exode Rural. De manière parallèle, on lit un impact de l'impact de l'accessibilité à l'échelle régionale, dans la mesure où l'évolution des taux de croissance annuel moyen suit parallèlement l'évolution des temps de parcours en

fonction de la centralité dans le réseau.

Encore une fois, l'étude des centralités est venue confirmer davantage encore la coévolution avec les niveaux de population. Ici, l'analyse différenciée des villes et campagnes montre que la structure du réseau semble influencer de la même manière les deux grandes familles de communes, même si la tendance démographique de fond devra être intégrée pour donner plus de véracité à ces analyses.

Les deux premières sections ont permis de remettre au goût du jour deux termes importants de la controverse et même de les croiser, à partir d'une étude inédite qui s'appuie sur la base de données FRANcE sur un échantillon de 36 000 communes. Coévolution et hiérarchisation sont deux éléments de la controverse des effets structurants que nous venons valider face à l'accélération des vitesses et qui font osciller la phase de croissance du réseau de l'homogénéisation à l'hétérogénéisation :

- Les débuts du réseau montrent les écarts les plus importants, alors que la hiérarchisation progressive des vitesses profite davantage à une convergence inter-régionale sur le territoire français et des changements plus marginaux ensuite ;
- La coévolution entre accessibilité et niveau de population est attestée ici à partir des nouveaux indicateurs produits à de larges échelles, où l'écart entre les niveaux de population croît à mesure que le réseau se développe et participe au maintien de disparités intra-régionales. En revanche, aucune structure de la sorte n'apparaît pour les communes inférieures à 2 500 habitants.
- La coévolution entre centralité et niveau de population est confirmée à partir des nœuds du multigraphe constitué dans le cadre de la thèse. Ce phénomène est perceptible pour l'ensemble des communes, indépendamment des villes et campagnes. Même, l'ampleur du recul des populations rurales est fonction du niveau de centralité des communes dans le réseau.

Les approches que nous avons développées jusqu'ici confrontent deux à deux les descripteurs démographiques et les descripteurs réticulaires. Par ailleurs, ils sont confrontés décennie après décennie. Ainsi, la comparaison des formes atteste d'une coévolution du réseau ferroviaire français avec la structure démographique par la desserte et par l'accessibilité.

Des tendances quant à leur influence sur la croissance démographique locale ressortent également du dernier nuage de points, sans que l'on puisse juger du caractère déterminant du réseau. La forte hétérogénéité des données implique des résultats où les amplitudes demeurent importantes. Mais l'approche développée et défendue dans la seconde partie de la thèse nous autorise à ce stade à considérer l'ensemble des communes françaises – large échelle spatiale –, la longue durée de notre travail cette fois – large échelle temporelle –, ainsi que la prise en compte d'un contexte à la fois local et global – les tendances de fond de la croissance démographique, les tendances socio-économiques des territoires. Le recours à l'économétrie de panel permet de produire un modèle analytique pour étudier l'effet de l'infrastructure de transport sur la population.

7.3. Portée spatiale et temporelle de l'effet du réseau

Au contraire de la section précédente, celle-ci s'attache à mesurer l'effet propre du réseau dans la croissance démographique locale, alors que l'on sait déjà que les formes de peuplement et de réseau évoluent de manière parallèle, et que l'on a identifié des clés pour l'appréhension des rapports causaux. Si la littérature de long terme s'y attache depuis longtemps, nous enrichissons cette approche par la complexification de la dotation en réseau : les temps de parcours, l'effet de la vitesse, l'ancienneté de la connexion ou encore la portée des réseaux. On vise alors à prendre en compte la variabilité sur l'ensemble du territoire et sur le long terme. Alors même que le biais d'endogénéité demeure important, nous proposons deux approches qui permettent de le limiter : cela passe par des choix dans l'échantillon à intégrer dans la modélisation, de même que le recours à l'économétrie de panel et à l'instrumentation. Cette section reprend cette montée en complexité, en accompagnant chaque étape des tests qui permettent de valider ou d'évaluer dans quelle mesure les modèles doivent être considérés avec prudence. Nous sélectionnons ici les résultats les plus significatifs et évocateurs et repoussons en annexes les détails nécessaires à la robustesse des résultats. La première sous-section détaille la spécification économétrique jusqu'à la prise en compte d'effets directs et indirects. La modélisation en panel permet de prendre en compte plusieurs dimensions dans l'analyse : la portée spatiale de l'effet du réseau, la portée temporelle de la connexion, la force de la centralité dans le réseau. Enfin, la méthode par les variables instrumentales vise à prendre en compte l'endogénéité dans la coprésence des effets directs et indirects du réseau.

7.3.1. Effets directs et indirects sur la croissance de long terme : une homogénéisation ?

L'objectif de cette sous-section est de définir quel échantillon nous allons retenir dans les analyses. Il s'agit de prendre en compte l'hétérogénéité des données pour évaluer d'un seul tenant ce qui se passe dans les villes et dans les campagnes, en contrôlant l'instabilité des paramètres. Pour autant, à la mesure des analyses précédentes, nous devons limiter le biais de simultanéité que nous supposons pour certains individus : l'objectif de la modélisation est de voir dans quelle mesure la connexion influence la croissance démographique. Ainsi, à la manière de certains autres auteurs (Hornung, 2012a), nous choisissons de retirer de notre échantillon les villes plus grandes : telles qu'elles sont mentionnées dans la loi de 1842, les plus grandes villes ont été connectées le plus tôt au réseau initié par Legrand. Dès lors, la présence du réseau dans ces villes est précisément liée à la construction du réseau d'intérêt national, de manière endogène. Pour les mêmes raisons, nous avons choisi d'éliminer de l'échantillon les terminus : ils sont la plupart du temps ces mêmes grandes villes, et présentent de plus des taux de croissance largement plus forts, d'après ce que nous avons vu dans la section précédente. En ne conservant que les communes dont nous disposons de toutes les données, notre échantillon s'élève à 30 450 communes.

Afin d'évaluer la cohérence de notre stratégie de modélisation, de même que les biais économétriques qu'une telle entreprise engendre, nous proposons d'abord d'étudier les relations entre croissance et territoire à partir des effets de très long terme. De la même manière

qu'auparavant, nous limitons nos analyses à la période 1860-1911, parce que les estimations seraient fortement biaisées par les pertes de la Première Guerre Mondiale si on poursuit nos analyses au-delà. L'objectif est de mesurer la stabilité des paramètres dans le temps, et comment les différents indicateurs de la dotation en réseau nous renseignent sur son effet.

Les tableaux de résultats sont présentés de manière analogue dans l'ensemble de la section. La variable dépendante est toujours l'information démographique. Chaque ligne correspond aux variables qui la déterminent. En gras sont présentées les variables d'intérêt qui caractérisent le réseau. Suivent ensuite les variables de contrôle en italique[45]. La dernière partie de chaque colonne rapporte le nombre d'observations, le pouvoir explicatif du modèle. Pour chaque variable, l'estimation du coefficient donne la force et le sens de la relation, sa significativité - * à 10 %, ** à 5 % et *** à 1 % de marge d'erreur, et son écart-type. En résumé, la couleur verte montre la force d'un facteur. La couleur orange traduit sa force relative. La couleur rouge traduit en revanche l'absence de significativité de la relation. Les numéros de colonnes sont appelés au fil du commentaire.

Ainsi, lorsqu'on effectue une régression des Moindres Carrés Ordinaires, l'obtention d'une gare sur le temps long a un effet positif sur la croissance de la population, à la hauteur de 0,68 %, même si le pouvoir d'un tel modèle est très faible avec un coefficient de détermination inférieur à 4 % *(Table 7. 2, Colonne 1)*.

Variable dépendante : Croissance de la population 1866-1911	**(1)** *Modèle simple*			**(2)** *Prise en compte des variables de contrôle*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accès au réseau en 1860	**0,0068**	***	**0,0003**	**0,0038**	***	**0,0003**
Croissance 1856-1866				0,1293	***	0,0110
Accès aux voies navigables				0,0028	***	0,0002
Ressources minières				0,0007	***	0,0002
Accès à la route				0,0005	***	0,0001
Hiérarchie administrative				0,0043	***	0,0002
Spécialisation agricole				-0,0014	***	0,0003
Spécialisation industrielle				-0,0012	***	0,0002
Diversité sectorielle				0,0005	***	0,0001
Constante	-0,0043	***	0,0000	-0,0045	***	0,0002
Observations	30450			30450		
R^2	0,0371			0,1547		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 2. Croissance de la population entre 1866 et 1911 et accès au réseau

La première complexification consiste à introduire l'ensemble des variables qui ont été décrites dans le *Chapitre 5*, et notamment toutes celles qui peuvent influencer le contexte local *(cf. 5.3, p. 201)*. Nous présentons en *Annexe 1.1* la démarche de complexification croissante du modèle. On introduit tout d'abord des effets de tendance par rapport à la croissance de la période précédente *(Colonne 2)*. On remarque alors que la croissance précédente a une influence sur la trajectoire de croissance des communes. Il est important de la prendre en compte. Puis, il s'agit de voir l'influence des réseaux concurrents au chemin de fer dans la seconde moitié du XIXème siècle : on utilise les variables qui décrivent la présence d'une voie navigable ou d'une voie routière, définie dans la carte

[45] Les variables instrumentales ne seront insérées que dans la dernière sous-section.

de Cassini. Nous saisissons également la structure socio-économique saisie à l'échelle départementale, que nous avons enrichie par des variables de spécialisation et de diversité économiques. Dès lors, on peut contrôler l'effet du réseau sur la croissance démographique par d'autres variables exogènes, telles que la spécialisation agricole, l'industrialisation progressive ou la diversification de l'économie (Combes et al., 2011). La prime à la croissance démographique est alors donnée aux territoires bénéficiant d'une diversité sectorielle. Si l'accès aux réseaux concurrents influence positivement la croissance, l'obtention d'un accès au réseau permet d'accroître la croissance de la population de l'ordre de 0,38 %, avec un pouvoir explicatif augmenté à 15 %. D'une part, cette nouvelle spécification renforce le pouvoir explicatif, qui démontre que la première souffrait d'un manque de variables. D'autre part, elle rend compte d'un coefficient qui reste significatif, toujours positif mais dont la force est plus faible : pris isolément, le modèle tendait à surestimer l'effet du réseau. Cela démontre encore une fois la nécessité de prendre en compte des éléments de contexte local. On peut désormais envisager de complexifier la dotation en réseau par les mesures d'accessibilité.

Nous reprenons dans la *Table 7. 3* les mêmes variables de contrôle pour la spécification du modèle en prenant désormais en compte les temps de parcours moyens calculés grâce au multigraphe. Nous les renforçons en prenant en compte des effets fixes pour chaque commune saisis à l'échelle départementale : ils permettent de prendre en compte un effet que l'on suppose inobservé par les autres variables, qui viendrait biaiser les estimations. Dans notre cas, il permet surtout de contrôler les effets de bord de l'accessibilité multipolaire. On remarque que le pouvoir explicatif avec ces variables augmentent à hauteur de 31 %[46], ce qui est un résultat satisfaisant quand notre échantillon comporte plus de 30 000 individus *(Colonne 1)*.

Variable dépendante : Croissance de la population 1866-1911	**(1)** *Effet des temps de parcours moyens*			**(2)** *Effet direct et indirect du réseau*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	**-0,0028**	***	**0,0003**	**-0,00171**	***	**0,0003**
Effet marginal "gare"				**0,00045**	***	**0,0000**
Croissance 1856-1866	0,0760	***	0,0097	0,07307	***	0,0096
Accès aux voies navigables	0,0026	***	0,0002	0,00224	***	0,0002
Ressources minières	0,0006	***	0,0002	0,00059	***	0,0002
Accès à la route	0,0007	***	0,0001	0,00061	***	0,0001
Hiérarchie administrative	0,0039	***	0,0002	0,00365	***	0,0002
Spécialisation agricole	-0,0001		0,0003	-0,00010		0,0003
Spécialisation industrielle	-0,0001		0,0003	-0,00010		0,0003
Diversité sectorielle	0,0010	***	0,0002	0,00112	***	0,0002
Constante	0,0172	***	0,0022	0,00937	***	0,0022
Observations	30450			30450		
R^2	**0,3148**			**0,3228**		
Effets fixes	Département			Département		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 3. Croissance de la population entre 1866 et 1911 et temps de parcours

L'effet des temps de parcours est significatif : des temps de parcours réduits contribuent à une croissance démographique plus forte. L'expression des variables avec des logarithmes nous

[46] L'augmentation du pouvoir explicatif est surtout à imputer à l'augmentation du nombre de variables engendré par la prise en compte d'effet fixes. Il s'agit d'un coefficient souvent supérieur à ceux rencontrés dans la littérature.

permet une interprétation par les élasticités. Ainsi, lorsque les temps de parcours moyens diminuent de 1 %, alors le niveau de population augmente de 0,28 % *(Colonne 1)*. Nous avons vu dans les sections précédentes les difficultés à interpréter de manière aisée les autres indicateurs d'accessibilité construit. Ainsi, nous choisissons de nous borner à l'indicateur multipolaire dans le cadre de ces analyses[47].

La position relative de la commune grâce au réseau influence sa croissance démographique : en revanche, à l'intérieur même de cet effet, cela ne nous permet pas de saisir l'effet déterminant de posséder une gare. C'est pourquoi nous construisons un indicateur d'interaction entre le temps d'accès moyen et le fait de posséder une gare. De cette sorte, quand on combine les variables explicatives du temps d'accès moyen et de l'indicateur d'interaction, on mesure l'effet marginal de la possession d'une gare. La *Colonne 2* montre un pouvoir explicatif légèrement supérieur : les résultats montrent que lorsque le temps de parcours moyen diminue de 1 %, le niveau de population augmente de 0,17 %. Si ce gain de temps est lié à l'obtention d'une gare, alors le niveau s'accroît encore de 0,04 %[48]. Cette analyse permet de différencier l'effet du réseau direct et indirect : la croissance du réseau français a un impact dans les communes qu'il dessert directement mais a un effet indirect sur les communes voisines. La valeur plus faible des estimations démontre qu'une part des coefficients jusqu'alors obtenus est porteuse d'une endogénéité.

Désormais, nous souhaitons étudier dans quelle mesure l'échelle d'analyse de l'accessibilité peut-être un facteur explicatif plus ou moins important pour la croissance démographique locale de long terme. Nous utilisons pour cela l'indicateur d'accessibilité potentiel *(cf. 5.1.2, p. 183)* à deux échelles différentes *(Table 7. 4)*. Parce qu'il a été construit à l'échelle régionale, nous devons remplacer les effets fixes non plus à l'échelle du département mais de la région : passant de plus de 90 effets à 22 effets, le R^2 est de fait diminué, mais dans une mesure acceptable parce qu'il demeure à environ 24 % *(Colonne 1)*. On voit que l'accessibilité potentielle à une aire locale a un impact positif et significatif, de l'ordre de 0,03 % *(Colonne 1)*. L'accessibilité mesurée à l'échelle nationale n'a qu'un effet trois fois inférieur et n'est pas significatif dans la spécification du modèle. On en conclut à ce stade que les performances du réseau à l'échelle régionale ont un impact significatif, plus qu'à l'échelle nationale, sur la croissance démographique locale, même si les valeurs des coefficients restent faibles, que l'on impute à la grande hétérogénéité des situations et à la force des disparités intra-régionales.

[47] Des tests ont été effectués à partir des autres indicateurs, dont le pouvoir explicatif reste le même ou devient insignifiant.

[48] Le coefficient est ici bien positif : quand on dérive la croissance par rapport aux temps de parcours, alors l'effet marginal de l'obtention d'une gare doit être positif pour bénéficier à la croissance de la population.

Variable dépendante : Croissance de la population 1866-1911	(1) Echelle régionale			(2) Echelle nationale		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accessibilité potentielle à l'échelle régionale	**0,00038**	***	**0,00006**			
Accessibilité potentielle à l'échelle nationale				**0,00010**		**0,00019**
Croissance 1856-1866	0,09643	***	0,01039	0,10122	***	0,01027
Accès aux voies navigables	0,00251	***	0,00019	0,00268	***	0,00019
Ressources minières	0,00046	**	0,00021	0,00052	**	0,00022
Accès à la route	0,00047	***	0,00008	0,00058	***	0,00007
Hiérarchie administrative	0,00402	***	0,00020	0,00447	***	0,00019
Spécialisation agricole	-0,00244	***	0,00036	-0,00264	***	0,00036
Spécialisation industrielle	-0,00100	***	0,00023	-0,00091	***	0,00024
Diversité sectorielle	0,00061	***	0,00012	0,00055	***	0,00012
Constante	-0,00277	***	0,00056	-0,00111		0,00195
Observations	30450			30450		
R^2	**0,249**			**0,247**		
Effets fixes	Région			Région		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 4. Croissance de la population entre 1866 et 1911 et accessibilité potentielle

Ainsi, alors que de nombreuses études de long terme s'attachent au simple accès au réseau (Atack et al., 2010 ; Hornung, 2012a), cette première étape montre qu'au-delà de la discrimination par la présence d'une gare, les temps d'accès moyen ont un impact sur la croissance démographique locale : le réseau a donc un effet direct mais aussi indirect, dont la portée à l'échelle régionale semble plus importante qu'à l'échelle nationale, montrant le rôle de la diffusion du réseau, dont il conviendra de définir la portée.

Par ailleurs, nous suspectons différents obstacles à la validation statistique de ce modèle. Le premier est le risque évident d'endogénéité des variables, tant nous avons vu le parallélisme dans l'évolution des formes des deux objets de notre recherche. Aussi, plusieurs tests ont été conduits à partir des estimations ci-dessus : dans tous les cas, le test de Ramsey Reset démontre l'omission de variables dans les spécifications, malgré la prise en compte croissante d'un contexte local. Par ailleurs, le test de Chow a permis de mettre en évidence une instabilité des coefficients à l'intérieur des deux sous-échantillons, dont la définition reprend le seuil des 2 500 habitants déjà utilisé plus haut. Ces biais croisés d'endogénéité et d'hétérogénéité viennent confirmer les aspirations développées dans le *Chapitre 5* pour la construction d'un modèle faisant appel à l'économétrie de panel *(cf. 5.2.3, p. 197)*. L'endogénéité, à partir du biais de simultanéité et de causalité inverse, doit aussi être pris en compte par une stratégie instrumentale.

7.3.2. **Une portée hétérogène des réseaux dans l'espace et dans le temps**

Pour affiner la dimension temporelle, une solution aurait été de découper les régressions précédentes en sous-périodes. Deux obstacles s'érigent alors : les biais d'endogénéité et

d'hétérogénéité auraient demeuré, l'effet temporel sur une période longue ne pourrait pas être saisi dans son ensemble[49]. Alors les estimations en données de panel permettent de mesurer l'effet du réseau entre 1866 et 1911, en prenant en compte chaque évolution dans les sous-périodes. La démarche classique consiste à étudier l'ensemble des informations indépendamment du temps dans une première étape puis d'introduire des effets dans le temps et entre les individus. Ces effets rendent compte de l'hétérogénité inobservée en prenant en compte l'omission de variables. Pour chaque sous-période et afin de limiter la causalité inverse, on mesure les niveaux de population d'une période intercensitaire par rapport à la dotation en réseau dans la période précédente. Comme précédemment, les effets sur les individus contrôlent la trajectoire des communes tandis que les effets sur les périodes prennent en compte les tendances nationales de la croissance démographique. Alors, les coefficients des variables de dotation en réseau traduisent la croissance additionnelle pour un même niveau de population sur une période intercensitaire.

En supprimant les individus pour lesquels ils manquent des informations, nous obtenons un panel cylindré de 30 450 individus, ce qui rend plus robustes les estimations obtenues. Les variables de contrôle doivent en revanche être variantes dans le temps : nécessairement, les informations sur les ressources minières et la proximité au réseau dessiné par Cassini disparaissent. Ces informations sont en revanche modélisées dans les effets attribués à chaque individu. La variable d'accès aux voies navigables demeure pour contrôler l'apparition de nouveaux canaux dans la période d'étude, qui consiste en une répétition de six sous-périodes. A l'intérieur de celles-ci, l'évolution de la structure économique est prise en compte à partir des données d'emplois mesurées en 1860 et 1896. Comparativement à d'autres études sur le très long terme, nous soulignons que notre modèle bénéficie d'informations variantes dans le temps. Alors, sur le long terme et en prenant l'ensemble de la variabilité temporelle et individuelle, ce n'est plus la diversité sectorielle mais la spécialisation industrielle qui influe davantage sur la croissance démographique locale depuis la seconde moitié du XIXème siècle jusqu'à la veille de la Première Guerre Mondiale. Cet aspect nous conforte dans la spécification du modèle proposé.

Nous reprenons la même montée en complexité dans la spécification de la dotation en réseau. La variable binaire de la présence d'une gare nous sert à évaluer le modèle dans son ensemble, puis avec des effets fixes, des effets aléatoires et d'arbitrer quelle spécification est la plus appropriée pour la suite de nos investigations *(Annexe 1.2)*. Le modèle par effets fixes est appliqué aux variables présentées dans la sous-section précédente (*Annexe 1.3*). Nous ne présentons pas ici les résultats dans la mesure où ils démontrent encore une fois l'existence d'effets directs et indirects du réseau : ils viennent confirmer la présence d'effets quand on prend en compte l'hétérogénéité inobservée, et en intégrant une granularité temporelle. Alors que la possession d'une gare a un impact sur la croissance de 0,6 % par an en moyenne, l'effet des temps de parcours moyen est du même ordre, avec un pouvoir marginal de 0,1 % par an en moyenne pour une commune bénéficiant d'un accès direct au réseau. L'effet marginal est donc supérieur avec cette spécification.

Désormais, l'intérêt de cette sous-section est d'exploiter la valeur ajoutée de la structuration

[49] De telles analyses ont été conduites mais ne sont pas présentées ici compte tenu de leur faible pouvoir explicatif et de nombreuses variables dont la significativité est nulle.

en panel, en diversifiant les dimensions dans lesquelles nous poursuivons les analyses. Comment les doubles dimensions temporelles et individuelles peuvent encore davantage participer à la mesure d'un effet du réseau ? Les tableaux sont à lire de la même manière. En revanche, ils spécifient bien la prise en compte d'effets temporels et individuels dans chaque spécification en panel[50].

Pour étudier dans quelle mesure le temps de parcours moyen influence les évolutions dans les niveaux de population, nous mobilisons à nouveau la notion d'accessibilité potentielle à plusieurs échelles *(Table 7. 5)*. En mobilisant l'accessibilité potentielle, on s'attend ainsi à ce que plus une commune a potentiellement la possibilité de se déplacer vers les autres forces en présence, plus son dynamisme s'en trouvera amélioré. Contrairement aux premières analyses, la spécification du modèle en panel permet de mesurer successivement l'effet propre de chaque échelle. Dans les deux cas, nous retrouvons les mêmes caractéristiques générales du modèle, avec un pouvoir explicatif qui est stabilisé, toutefois légèrement supérieur qu'avec les seuls temps de parcours, laissant à penser que l'analyse par les niveaux d'échelle se justifie. La significativité des variables d'intérêt est toujours attestée. Dans le premier cas, une augmentation de 1 % de l'accessibilité potentielle se traduit par une augmentation de 1,2 % par an en moyenne[51]. L'effet d'une meilleure accessibilité potentielle à l'échelle régionale permet alors une croissance marginale de 0,6 % par an en moyenne. Dans l'autre cas, l'accessibilité potentielle a un impact de 1,4 % par an en moyenne, mais l'effet marginal de l'accessibilité à l'échelle nationale n'est que de 0,2 % par an en moyenne. Ainsi, on peut conclure qu'avoir une meilleure accessibilité à l'échelle régionale a un impact 3 fois supérieur que celle d'avoir une meilleure accessibilité à l'échelle nationale.

Variable dépendante : Niveau de population entre 1866-1911	**(1)** *Effet régional*			**(2)** *Effet national*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accessibilité potentielle	**0,135**	***	**0,001**	**0,162**	***	**0,001**
Effet régional	**0,062**	***	**0,001**			
Effet national				**0,023**	***	**0,001**
Accès aux voies navigables	0,022	***	0,008	0,023	***	0,009
Spécialisation agricole	-0,202	***	0,006	-0,221	***	0,006
Spécialisation industrielle	0,106	***	0,005	0,114	***	0,006
Diversité sectorielle	-0,095	***	0,003	-0,109	***	0,003
Constante	6,219	***	0,004	6,213	***	0,004
Observations	30450			30450		
R^2	**0,3337**			**0,2995**		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 5. Effet de l'accessibilité potentielle entre 1866 et 1911

[50] Des effets croisés démontrent la multiplicité des dimensions dans lesquelles les estimations peuvent être conduites. Leur utilisation n'apporte toutefois pas ici une meilleure compréhension du modèle.
[51] Les coefficients d'estimation estiment des évolutions sur la période intercensitaire que nous transformons en taux de croissance annuel moyen au fil du texte.

Si l'effet marginal de l'accès direct au réseau n'est plus à démontrer, il reste en revanche des interrogations quant à sa portée. Pour cela, nous étudions l'interaction entre la distance à la gare la plus proche et les temps de parcours moyens. Comme auparavant, il s'agit de qualifier l'effet marginal de la distance au réseau dans les temps de parcours moyen, en opérant une classification des communes par couronne concentrique dont l'échelle mesure le temps nécessaire pour rejoindre une gare. Par exemple, parmi toutes communes non-desservies directement par une gare, l'objectif est de tester l'effet marginal qui consiste à être à proximité d'une gare à moins de 30 minutes. Par la suite, parmi celles qui sont à plus de 30 minutes d'une gare, nous testons l'effet marginal qui consiste à être à proximité d'une gare à moins de 40 minutes. L'opération est répétée autant de fois que l'effet marginal est significatif *(Table 7. 6)*. Notons que les effets de l'environnement sont pris en compte dans les effets individuels. Le point positif dans ces spécifications est la croissance de l'effet indirect au détriment de l'effet direct à mesure que l'on s'éloigne de la gare. C'est cependant à partir du seuil de 60 minutes que l'effet marginal de la proximité de la gare devient de moins en moins significatif *(Colonne 3)* pour être nul à partir du seuil de 70 minutes *(Colonne 4)*.

Variable dépendante : Niveau de population entre 1866-1911	(1) Seuil de 30 minutes			(2) Seuil de 50 minutes			(3) Seuil de 60 minutes			(4) Seuil de 70 minutes		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	**-0,055**	***	**0,006**	**-0,066**	***	**0,005**	**-0,068**	***	**0,005**	**-0,069**	***	**0,006**
Effet marginal "proximité"	**0,006**	***	**0,000**	**0,003**	***	**0,000**	**0,001**	*	**0,000**	**0,000**		**0,000**
Accès aux voies navigables	0,035	***	0,009	0,031	***	0,010	0,034	***	0,010	0,019	*	0,011
Spécialisation agricole	-0,216	***	0,012	-0,175	***	0,013	-0,177	***	0,013	-0,195	***	0,014
Spécialisation industrielle	0,012		0,008	0,031	***	0,007	0,032	***	0,008	0,026	***	0,008
Diversité sectorielle	-0,054	***	0,003	-0,057	***	0,003	-0,056	***	0,003	-0,053	***	0,003
Constante	6,620	***	0,042	6,708	***	0,040	6,731	***	0,041	6,752	***	0,042
Observations	29268			28075			27733			26609		
R^2	**0,3885**			**0,4588**			**0,4768**			**0,4827**		
Effets temporels	Oui			Oui			Oui			Oui		
Effets individuels	Oui			Oui			Oui			Oui		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 6. Effet de la proximité de la gare entre 1866 et 1911

Partant de ce constat, il nous a semblé intéressant de regarder si l'impact du réseau est analogue quant à l'étude la centralité. Pour cela, nous reprenons l'indicateur portant sur le nombre de lignes divergentes utilisé dans la section précédente *(Table 7. 7)*. Ainsi, par rapport à une commune qui n'a pas d'accès au réseau, la croissance additionnelle pour un même niveau de population est de 0,5 % par an en moyenne *(Colonne 1)*. Si l'effet reste modéré quand une commune ne se trouve que le long d'une ligne, il se renforce avec la centralité dans le réseau. Quand la commune se situe dans un carrefour à 3 branches, la croissance additionnelle est de 1,5 % par an en moyenne. Le coefficient croît de manière linéaire *(Figure 7. 14. a.)* à mesure que le carrefour est complexifié, et l'effet reste toujours significatif. Pour vérifier que cet effet n'est pas capté par les plus grandes villes, bien que contrôlé par le choix de l'échantillon, on mesure l'effet pour les communes de moins de 2 500 habitants *(Colonne 2)* : le pouvoir explicatif ne s'en trouve pas pénalisé. On constate une stabilité des coefficients, voire même, ils sont supérieurs quand le carrefour est doté de cinq branches ou plus, attestant du caractère déterminant de la place de la commune dans le réseau situé en zones rurales.

Variable dépendante : Niveau de population entre 1866-1911	(1) Centralité dans le réseau			(2) Centralité et campagnes		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Etape	**0,053**	***	**0,002**	**0,051**	***	**0,002**
Carrefour à 3 branches	**0,165**	***	**0,004**	**0,159**	***	**0,004**
Carrefour à 4 branches	**0,314**	***	**0,006**	**0,305**	***	**0,008**
Carrefour à 5 branches	**0,418**	***	**0,013**	**0,440**	***	**0,018**
Carrefour à plus de 5 branches	**0,641**	***	**0,027**	**0,662**	***	**0,042**
Accès aux voies navigables	0,007		0,009	0,004		0,009
Spécialisation agricole	-0,212	***	0,006	-0,207	***	0,007
Spécialisation industrielle	0,107	***	0,006	0,100	***	0,006
Diversité sectorielle	-0,085	***	0,003	-0,079	***	0,003
Constante	6,260	***	0,004	6,176	***	0,004
Observations	30450			28991		
R^2	**0,2579**			**0,2766**		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 7. Centralité dans le réseau entre 1866 et 1911

Après avoir caractérisé des effets d'échelles et de centralité, l'intérêt d'une telle modélisation cylindrée est de pouvoir investiguer plus largement les effets temporels. Pour analyser des effets de long terme et de très long terme, on se propose de mesurer l'impact de l'ancienneté de la connexion sur la trajectoire de population des communes. De cette manière, la structuration en panel nous permet de comparer la trajectoire des villes pour une même ancienneté de connexion par rapport à toutes les autres. L'objectif est alors de montrer si les effets temporels sont linéaires ou non. La *Table 7. 8* propose cette spécification pour l'ensemble du cylindre d'une part *(Colonne 1)*, et les communes avec moins de 2 500 habitants au début de la période d'étude *(Colonne 2)*. Alors, l'effet individuel du panel permet de stabiliser les coefficients qui ne l'étaient pas par de simples régressions. Le pouvoir explicatif du modèle est le même que celui que l'on constate pour l'effet de la centralité du réseau.

On constate des effets positifs de l'accès au réseau pour toutes les périodes, où l'augmentation progressive des estimations démontre l'effet cumulatif et retardé de la connexion au réseau sur l'augmentation des niveaux de population. Si l'obtention d'une nouvelle connexion permet une croissance additionnelle de 0,5 % par an en moyenne, l'effet au bout de 10 ans atteint 0,7 % par an en moyenne. En revanche, le phénomène s'accentue par la suite et est non-linéaire à mesure que la connexion est ancienne *(Figure 7. 14. b.)*. Entre échantillon complet et campagnes, la stabilité des paramètres montre que ce phénomène touche à la fois les villes et les campagnes, avec un pouvoir explicatif qui est lui aussi stabilisé. Ainsi, sur une longue période, l'expérience de la connexion est le gage d'une plus grande augmentation de la population pour un même niveau. Une courbe de tendance exponentielle est significative à plus de 99 %.

Variable dépendante : Niveau de population entre 1866-1911	(1) *Mémoire de la connexion*			(2) *Connexion et campagnes*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Nouvelle connexion	0,046	***	0,002	0,043	***	0,002
Connexion depuis 1 période	0,076	***	0,002	0,070	***	0,002
Connexion depuis 2 périodes	0,118	***	0,002	0,107	***	0,003
Connexion depuis 3 périodes	0,172	***	0,003	0,155	***	0,003
Connexion depuis 4 périodes	0,239	***	0,004	0,218	***	0,004
Connexion depuis 5 périodes	0,317	***	0,005	0,293	***	0,005
Connexion depuis 6 périodes	0,451	***	0,008	0,416	***	0,009
Accès aux voies navigables	-0,003		0,009	-0,007		0,009
Spécialisation agricole	-0,203	***	0,006	-0,201	***	0,006
Spécialisation industrielle	0,106	***	0,006	0,100	***	0,006
Diversité sectorielle	-0,078	***	0,003	-0,074	***	0,003
Constante	6,274	***	0,004	6,185	***	0,004
Observations	30450			28991		
R^2	0,2734			0,2884		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 8. Mémoire de la connexion entre 1866 et 1911

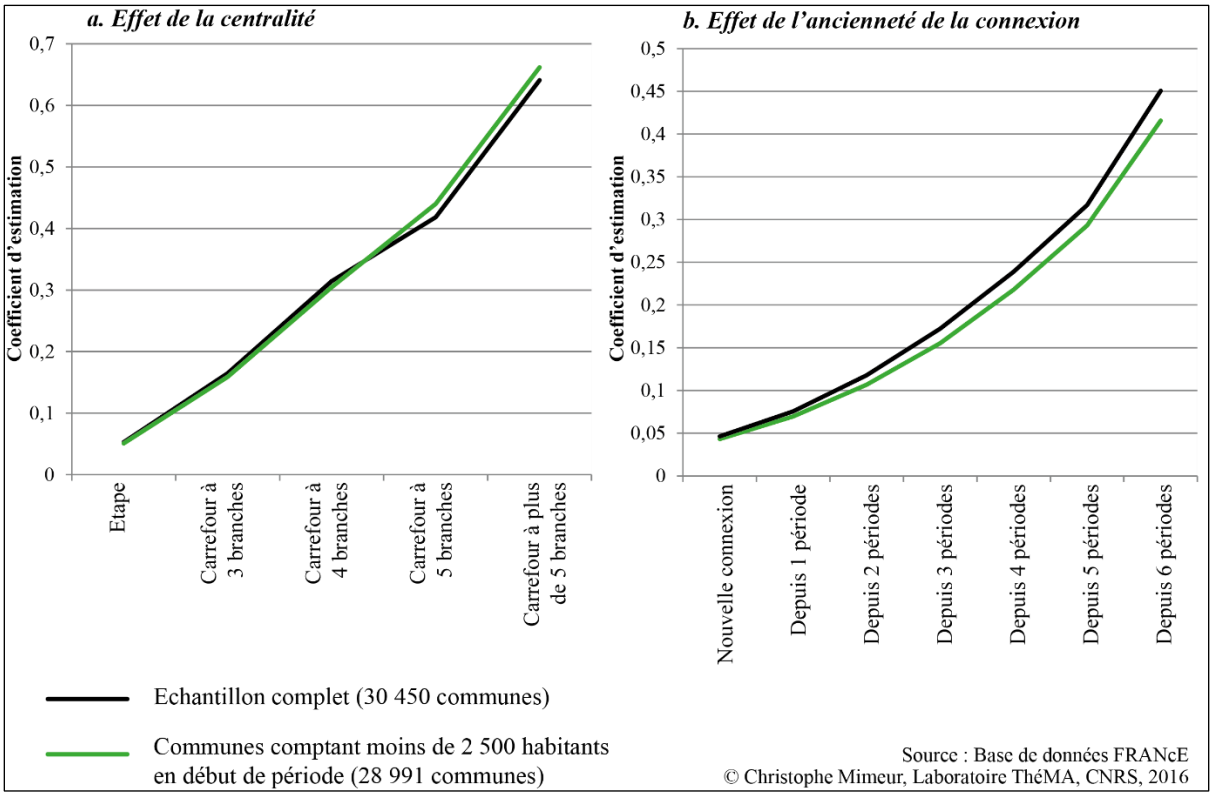


Figure 7. 14. Effets de la mémoire de la connexion et de la centralité dans le réseau dans la modélisation en panel entre 1866 et 1911

L'utilisation des temps de parcours moyens démontre comment le réseau a un effet direct et indirect sur la démographie française : la complexification des indicateurs montre que l'accessibilité potentielle à une méso-échelle a un impact marginal et significatif plus fort que l'accessibilité potentielle à une macro-échelle. La diversité des dimensions d'analyse du panel nous permet également de montrer comment la mémoire de la connexion effective au réseau participe à un phénomène non-linéaire et fortement cumulatif. Contre intuitivement, en revanche, l'effet de la centralité des réseaux suit un phénomène linéaire, sans distinction entre villes et campagnes : l'effet temporel a un impact cumulatif plus fort que l'effet du réseau lui-même.

Cette sous-section montre ainsi la richesse des dimensions dans lesquelles nous pouvons envisager l'analyse du lien entre réseau et territoire. Par ailleurs, la dimension de panel permet de limiter l'omission de variables – omniprésente dans un contexte géohistorique – et à la stabilisation des coefficients des estimateurs, en conjuguant les effets sur les individus et sur les périodes. Toutefois, nos analyses doivent être assorties de solutions quant au biais de causalité et de relation inverse, tant on voit le caractère cumulatif des effets de très long terme. Pour cela nous proposons une instrumentation exploratoire de notre démarche.

7.3.3. **Vers l'instrumentation des effets directs et indirects du réseau**

L'objectif de cette section est d'utiliser la méthode des variables instrumentales afin de valider les résultats déjà produits en s'assurant qu'ils ne souffrent pas d'un fort biais d'endogénéité : nous vérifions la coprésence d'effets directs et indirects en limitant les biais statistiques, de sorte à asseoir les analyses qui ont été conduites concernant la portée de ces effets. Les principes de l'utilisation des variables instrumentales ont déjà été évoqués dans la seconde partie de la thèse *(cf. 5.3.2, p. 209)*. Les variables utilisées doivent déterminer la dotation en réseau, tout en étant indépendante de la situation démographique, de sorte à ne pas surestimer les effets et de contourner les biais de causalité inverse et de simultanéité. Nous relevons deux critères dans cette démarche exploratoire : le premier réside dans le raisonnement géohistorique, le second dans la validation statistique. Ainsi, s'il est toujours très difficile de qualifier et valider le caractère exogène des variables instrumentales (Duranton et Turner, 2012), le recours au raisonnement géohistorique permet de conduire les investigations jusqu'à leur terme, en les replaçant dans le strict contexte de leur application.

Nous appliquons cette démarche dans la dimension en panel, pour prendre en compte les effets temporels dans la relation entre réseau et population : cela nous contraint à utiliser des variables instrumentales variantes dans le temps, et à combiner plusieurs instruments pour une variable endogène. En revanche, cela nous permet de conserver toutes les variables de contrôle variantes dans le temps utilisées jusqu'à présent. Dès lors, la démarche consiste à estimer la corrélation entre les instruments et la variable endogène, puis leur exogénéité par rapport à la

variable dépendante pour les utiliser dans les spécifications du modèle. En reprenant les principales clés de l'analyse proposée ci-dessus, nous choisissons de déterminer des instruments pour la variable d'accès au réseau et pour le temps de parcours moyen. Le premier cas a été souvent traité dans la littérature (Atack et al., 2010 ; Duranton et Turner, 2012 ; Hornung, 2012b). Le second ne l'a pas été à notre connaissance dans la longue durée. Nous vérifions dans cette sous-section la coprésence d'effets directs et indirects. Ainsi, nous proposons une approche qui combine la prise en compte d'une dimension d'aménagement antérieure à notre période d'étude et celle de la situation géographique. Aussi, pour limiter les biais de simultanéité, la prise en compte d'une telle situation sera reculée de deux périodes.

Pour instrumenter l'accès au réseau, nous utilisons les deux grands plans d'aménagement décisifs dans la phase de croissance du réseau (Caron, 1997) : la loi relative à l'établissement des grandes lignes de chemin de fer en France en 1842 et le plan Freycinet lancé en 1878. Dans les deux textes de lois, on retrouve les toponymes des localités à desservir, sans préjuger du tracé par ailleurs. Reprenant les principaux vocables de la théorie des réseaux, on peut lire dans ces textes les réseaux de projets transactionnels pour le ferroviaire français : l'objectif est de relier alors le plus rapidement ces villes. Alors même que la *« détermination des tracés de toutes les lignes principales est hors notre portée »* (Caron, 1997), nous utilisons les techniques du SIG pour tracer les lignes droites entre les localités citées dans les deux plans *(Figure 7. 15)*. Cette démarche est utilisée dans la littérature : l'instrument construit consiste en une variable binaire qui prend la valeur 1 si la commune est traversée par cette ligne droite, avec une tolérance de deux kilomètres[52]. Il modélise l'adoption potentielle du réseau ferroviaire, avec l'idée derrière laquelle ces communes-là sont desservies pour des raisons exogènes, celles de relier les principales villes, et non endogènes à leur situation démographique. Ainsi, nous utilisons la variable décrivant les corridors issus du plan de Legrand pour les trois premières périodes de notre cylindre, et celles décrivant les corridors issus du plan de Freycinet pour les trois dernières périodes. De cette sorte, nous respectons un décalage de 12 à 18 ans suivant l'adoption des plans. Rappelons à ce stade que nous contrôlons l'endogénéité propre à ces plans parce que nous avons ôté de notre échantillon ces grandes villes ainsi que les terminus.

[52] Cette tolérance permet de prendre en compte certaines déviations dues à des obstacles topographiques, qui sont par ailleurs des éléments pris en compte dans les effets individuels modélisant l'hétérogénéité inobservée.

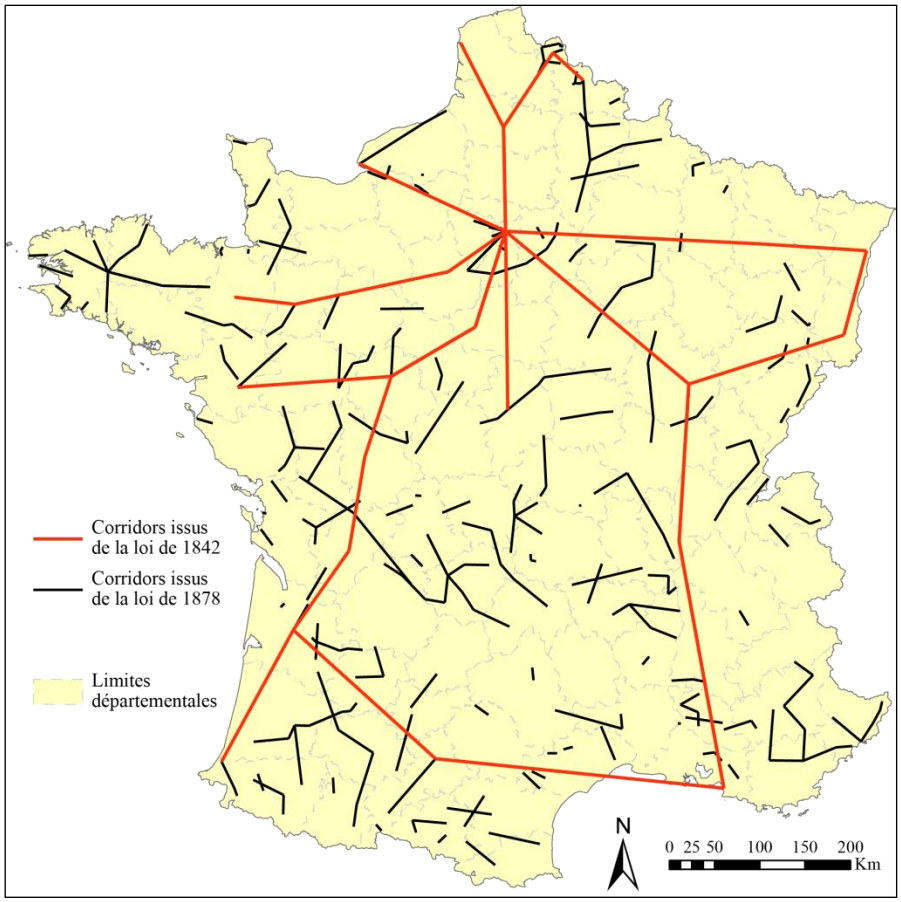


Figure 7. 15. Les corridors issus des lois de1842 et 1878

L'instrumentation de chacune des variables est proposée en *Annexe 1.4*. Les tests conduits poussent à les utiliser pour caractériser les effets directs et indirects. Il s'agit alors ici de croiser les différents instruments utilisés pour vérifier la coprésence d'effets directs et indirects du réseau dans l'évolution démographique entre 1866 et 1911. La spécification économétrique suggère l'interaction de chaque variable endogène par chacun de leur instrument. Ainsi, pour les deux variables suspectées d'endogénéité – temps de parcours et effet marginal –, leur croisement requiert quatre variables instrumentales. La validation de cette méthode est proposée en *Annexe 1.4*.

La *Table 7. 9* montre l'effet des temps de parcours et de l'infrastructure à partir de l'instrumentation. Le pouvoir explicatif de ce modèle est de 30 %, et ne souffre pas d'un déficit par rapport aux premières estimations[53]. Ainsi, lorsqu'on contourne les potentiels biais d'endogénéité dans la relation, l'effet des temps de parcours moyens reste significatif et a un fort impact sur la croissance des niveaux de population. La présence de l'infrastructure la renforce nettement, avec une significativité à hauteur de 95 % de confiance. Les coefficients renforcés montrent que les spécifications précédentes souffraient d'une endogénéité qui avait tendance à sous-estimer l'effet

[53] L'information supplémentaire dans le tableau – Over-id test p-value – consiste à vérifier que les instruments utilisés ne sur-identifient pas la variable qui décrit le réseau. La valeur élevée de 0,17 nous permet de rejeter cette hypothèse et de valider la méthode d'instrumentation.

du réseau mais dont l'instrumentation peut identifier en revanche des effets d'anticipation, qu'il faut donc considérer avec prudence.

Variable dépendante : Niveau de population entre 1866-1911	*Instrumentation par des variables retardées*		
	Coef,	*Signif,*	*Ecart-type*
Temps de parcours moyen	**-0,242**	***	**0,02732**
Effet marginal "gare"	**0,010**	**	**0,00499**
Accès aux voies navigables	0,015	***	0,01162
Spécialisation agricole	-0,15	***	0,00974
Spécialisation industrielle	0,135	***	0,00708
Diversité sectorielle	-0,1	***	0,00347
Constante	6,534	***	0,003
Observations	30450		
R^2	**0,3089**		
Over-id test p-value	**0,17**		
Effets temporels	Oui		
Effets individuels	Oui		

Source : Base de données FRANcE
© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 9. Instrumentation des effets directs et indirects du réseau entre 1866 et 1911

Nous voyons que malgré la richesse de la base de données FRANcE, l'une des principales difficultés est d'utiliser des instruments dont le caractère est variant dans le temps. C'est ce qui nous contraint à utiliser parfois des variables retardées, dont nous nous efforçons toutefois de les assortir d'un raisonnement géohistorique et d'un recul certain pour se dégager des biais de causalité inverse (Le Gallo, 2002). Mais, les effets toujours probables d'anticipation nous conduisent à poursuivre notre démarche exploratoire dans le cadre géohistorique, en mobilisant les réseaux historiques, qui nous ont permis de mettre à l'épreuve notre méthodologie dans la seconde partie de la thèse *(cf. 5.3.3, p. 211)*. La contrainte liée à la collecte des données, à leur disponibilité et leur fiabilité, nous conduit à réduire dans cette démarche le cylindre du panel jusqu'à présent utilisé. Ayant à notre disposition deux informations sur les réseaux historiques, nous réduisons notre panel à deux niveaux. Pour des raisons de commodité et pour avoir une vision symétrique, nous séparons le cylindre en son milieu, ce qui correspond par ailleurs à la phase qui suit l'inflexion du réseau suivant la classe de vitesses entre 1878 et 1883 *(cf. 3.2.2, p. 123)*. Par ailleurs, notre raisonnement géohistorique nous amène à penser raisonnablement que la structure du réseau antique et du réseau postal de 1833 n'a pas d'incidence directe sur l'évolution démographique entre 1866 et 1911. Cet état de fait ne nous dédouane pas de vérifier la véracité statistique de nos instruments.

Nous reprenons exactement la même démarche que pour le panel à 6 niveaux, en cherchant à instrumenter l'accès au réseau d'une part, les temps de parcours moyens d'autre part. C'est pour instrumenter la variable d'accès au réseau que le raisonnement géohistorique est le plus difficile. En effet, entre les réseaux antiques et postaux, on retrouve une densité qui est proche et qui ne traduit pas assez la phase d'expansion du réseau, qui est en revanche bien traduite dans la succession

des grands plans d'aménagement. Ainsi, pour limiter le biais d'anticipation, nous proposons d'assortir la variable des corridors utilisée auparavant avec la difficulté à rejoindre un réseau historique. Il s'agit d'une variable qui traduit le cheminement sur le graphe P nécessaire pour rejoindre le réseau antique ou postal. Alors, les tests statistiques qui nous permettent de valider les instruments sont présentés en *Annexe 1.5*. En revanche, il semble plus aisé d'instrumenter les temps de parcours moyens par les temps de parcours calculés sur les réseaux historiques : ils permettent d'expliquer plus de 90 % des temps de parcours sur les réseaux ferroviaires.

L'ultime étape de notre démarche nous pousse alors à étudier les effets d'interaction pour distinguer les effets directs et indirects du réseau à partir des variables instrumentales *(Table 7. 10)*. Cette fois, l'utilisation de deux instruments pour chacune des variables endogènes implique l'utilisation de six variables pour la démarche instrumentale[54]. Les résultats démontrent encore une fois le caractère significatif de l'effet indirect du réseau, par l'intermédiaire des temps de parcours moyen. Cet effet est renforcé quand une commune gagne une connexion. Il ne faut pas ici chercher à comparer les coefficients dans la mesure où les ordres de grandeur diffèrent. En revanche, on constate une stabilité dans les écarts entre effet direct et effet indirect. Ainsi, quelle que soit la méthode d'instrumentation, on voit que le rapport causal entre réseau et territoire est le fait de la coprésence d'effets directs et indirects, qui sont significatifs, mais dont le pouvoir explicatif est diminué. Nous posons l'hypothèse que les effets d'anticipation avaient tendance à surestimer le pouvoir explicatif de la précédente instrumentation et démontrent que la dimension temporelle est centrale dans l'étude des rapports causaux.

Variable dépendante : Niveau de population entre 1876-1906	*Instrumentation par les réseaux historiques*		
	Coef,	*Signif,*	*Ecart-type*
Temps de parcours moyen	**-0,496**	***	**0,100**
Effet marginal "gare"	**0,024**	***	**0,009**
Accès aux voies navigables	0,002	*	0,03104
Spécialisation agricole	-0,165	**	0,01651
Spécialisation industrielle	0,168	***	0,00173
Diversité sectorielle	-0,128	***	0,00122
Constante	6,241	***	0,003
Observations	34624		
R^2	**0,1792**		
Over-id test p-value	**0,0033**		
Effets temporels	Oui		
Effets individuels	Oui		

Source : Base de données FRANcE
© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 7. 10. Instrumentation des effets directs et indirects du réseau par les réseaux historiques entre 1876 et 1906

[54] On peut étudier les ultimes résultats malgré un test de sur-identification faible mais qui passe la valeur critique. Pour autant, avec autant d'instruments dans les interactions, les degrés de liberté sont au nombre de 4 : on arrive donc à rejeter l'hypothèse nulle de sur-identification.

On pointe toutefois quelques limites à notre analyse, qui est à un stade exploratoire, dans la mesure où de telles investigations sur le temps long, à partir de réseaux historiques, n'ont jamais été menées. Le principal apport porte sur la prise en compte des temps de parcours. Alors, l'une des principales limites réside dans la réduction du panel inhérente à la structuration des données. Par ailleurs, si la prise en compte de deux réseaux historiques, en modulant les vitesses de déplacement permet d'instrumenter des temps de parcours sur deux périodes, alors nous posons l'hypothèse que l'évolution d'un seul et même système à plusieurs dates permettrait encore davantage d'affiner les estimations, en conservant l'ensemble des niveaux du panel, et en étudiant l'évolution interne d'un même système. Cela renforcerait ainsi le raisonnement géohistorique : nous pointons la prudence avec laquelle nous devons envisager l'évolution des temps de parcours entre deux réseaux et deux modèles comportementaux différents, au-delà de la seule validation statistique, qui nous a permis toutefois dans cette démarche exploratoire de montrer les potentialités de telles entreprises.

A ce stade, nous retenons que la montée en complexité progressive de la dotation en réseau, assortie de méthodologies permettant de limiter les biais d'hétérogénéité, d'omission de variables et de causalité inverse, permet de dégager des effets du réseau sur les dynamiques de population *(Figure 7. 16)*. Cette figure caractérise un côté de la boucle définie dans le plan d'expérimentation *(cf. Figure 6. 20, p. 259)*. Compte tenu du caractère significatif de la relation dans nos estimations, jusqu'à l'épreuve de l'instrumentation, on peut raisonnablement ne pas rejeter les effets conjugués de l'accès direct au réseau et des effets indirects qu'il implique sur les temps de parcours moyens. Alors, nous avons mesuré un effet marginal de la gare, évalué entre 50 et 60 minutes pour notre période d'étude, dont nous postulons que l'effet non-linéaire est susceptible de se propager dans la limite de cette proximité, d'autant plus que le réseau améliore les performances de l'accessibilité à l'échelle régionale.

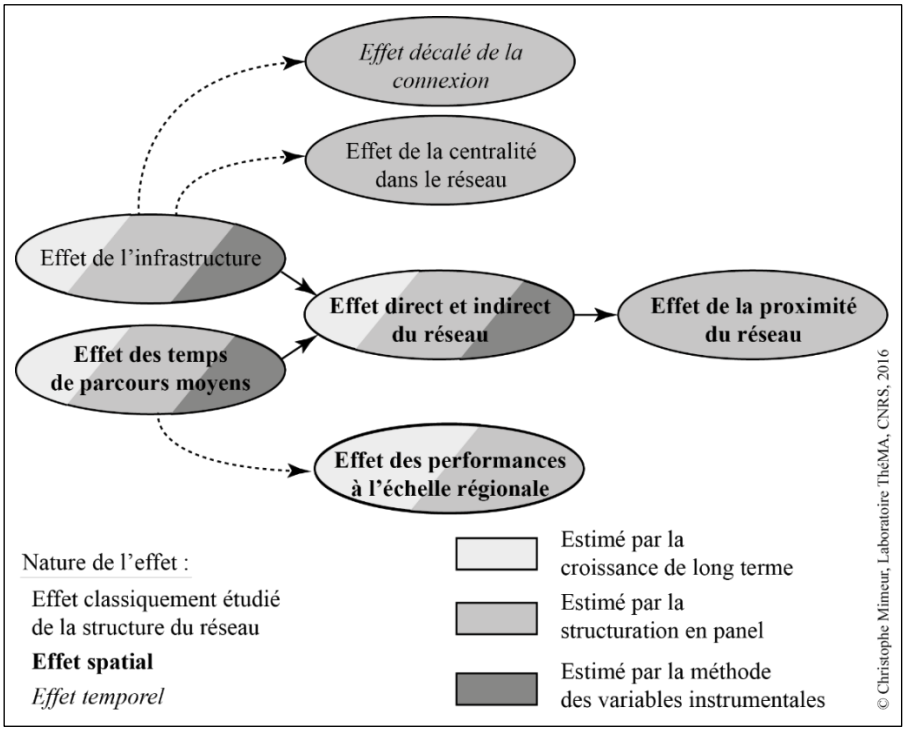


Figure 7. 16. La nature des effets du réseau par une complexification des estimations

Conclusion

L'arsenal d'indicateurs et d'outils mobilisés dans la seconde partie a été mis à l'épreuve des principaux vocables de la controverse, venant confirmer certains termes et en appuyer davantage d'autres. Alors que la phase d'accélération est souvent décrite d'abord par la mise en place d'un réseau d'intérêt national puis la mise en place d'un réseau local grâce au plan Freycinet, le phénomène de hiérarchisation des voies est transcalaire et traverse les échelles dès la mise en connexité du réseau en 1860. Alors que l'omnibus rejoignait les chefs-lieux de cantons, c'est toutefois à l'échelle régionale que la phase d'accélération semble la plus perceptible, et c'est notamment à partir de ce centre-là que la hiérarchie du réseau semble la plus significative. Ces conclusions remettent au goût du jour la notion de dualité des réseaux, où les plans largement décidés par le pouvoir central ont un impact à toutes les échelles, en dépit d'une forte armature nationale, préfigurant la constitution de sous-systèmes régionaux bien connu aujourd'hui pour les voyageurs à moyenne distance. Nous avons vu ensuite comment cette hiérarchisation du réseau épouse finalement la forme de l'armature urbaine, plus spécialement encore pour les villes, tandis que les situations sont plus confuses pour les communes rurales. Cette centralité n'est pas directement corrélée aux temps de parcours moyens, mais plutôt à une distribution polynucléaire, en en faisant des relais dans le système global.

Enfin, la dernière section a mis à profit la diversification géohistorique des données pour appréhender l'effet propre du réseau sur l'ensemble de la période 1866 à 1911. La dimension de panel permet d'affiner les estimations. Les temps de parcours moyen ont un effet dans l'évolution des niveaux de population – encore davantage à une échelle intermédiaire qu'à l'échelle nationale –, alors même que l'effet marginal de la gare demeure un élément important, qui se propage d'après les estimations dans un rayon d'une heure de marche à pied. Aussi, la mémoire de la connexion montre un effet non-linéaire qui démontre les effets de très long terme d'une connexion, dont les effets sont cumulatifs. Alors, la prise en compte de l'endogénéité par l'instrumentation des données est une voie dans laquelle nous insérons notre raisonnement géohistorique. Elle a permis de valider la présence d'effets directs et indirects dans les rapports causaux entre transport et territoire sur la longue durée, qui démontrent aussi une hétérogénéité de l'effet du réseau, parallèlement à sa hiérarchisation qualifiée de manière spatiale et temporelle.

Cette clé de lecture globale montre que l'effet structurant n'est pas à rejeter. Pour autant, nous ne pouvons pas rejeter non plus les nombreux biais, ceux déjà évoqués dans cette prise de conscience globale, de même que ceux qui ont plus particulièrement trait à l'espace. Nous identifions à ce stade des perspectives dans l'utilisation de modèles économétriques spatiaux pour encore affiner les estimations, à l'intérieur d'une structure en panel. Derrière cette spécification globale, l'objectif est d'établir des typologies de territoires dans lesquels l'effet direct, s'il existe et est significatif, est supérieur à l'effet indirect. On interroge alors la transposition de ces réflexions dans des régressions locales et non-paramétriques, qui ne doivent toutefois pas occulter les effets temporels identifiés dans ce chapitre. Ces réflexions trouvent leur place dans la littérature des effets de long terme et dont les spécifications peuvent trouver un écho dans des problématiques contemporaines, en démontrant l'intérêt de l'interdisciplinarité et du croisement des approches.

Chapitre 8. DES CAPACITES CIRCULATOIRES A LA SIMULATION DE LA CROISSANCE DU RESEAU FERROVIAIRE

Introduction

Les études portant sur les effets des structures préexistantes sur la planification des infrastructures sont surtout le fait d'échelles urbaines et métropolitaines (Ewing et Cervero, 2010), pour davantage étudier les choix modaux que la structure d'un réseau de transport (Commenges, 2013). Pour autant, certaines études empiriques ont conclu à l'adaptation du réseau ferroviaire français à une structure urbaine préexistante (Pumain, 1982). Nous avons nous-mêmes démontré la trajectoire proche que suivent la hiérarchie urbaine d'une part et la hiérarchie des centralités du réseau d'autre part. Mais, peu d'études se consacrent à l'étude des tracés des réseaux sur le temps long (Kasraian et al., 2016). De ce point de vue, même si la littérature et les débats politiques peuvent être utiles, *« l'analyse de la détermination des tracés de toutes les lignes principales est hors de notre portée »* (Caron, 1997) : une telle entreprise semble en effet vaine tant le recours aux archives semble inépuisable mais nous identifions dans les éléments d'un graphe, deux informations géohistoriques à exploiter à de larges échelles.

Ce chapitre s'inscrit alors dans un raisonnement abductif qui interroge le fonctionnement global du réseau et son application à l'échelle locale, à partir de situations contrefactuelles sans impact de la vitesse ferroviaire. A l'instar de la définition du réseau (Dupuy et Offner, 2005), l'approche circulatoire est l'entrée par laquelle nous envisageons ce côté de la controverse *(cf. Figure 6. 20, p. 259)*. Nous interrogeons alors les vocables de hiérarchisation, de sélection et finalement d'interactions entre une structure démographique préexistante et la phase de croissance du réseau ferroviaire. L'objectif n'est pas d'évaluer le rôle décisif du réseau dans les dynamiques territoriales mais d'étudier comment les tracés traduisent une forte inertie circulatoire dans la phase de croissance du réseau ferroviaire français. Ce chapitre démontre la diversité des approches, dans laquelle nous mettons à l'épreuve la théorie des graphes pour l'émergence de connaissances. Pour cela, le formalisme mathématique permet de faire ressortir des propriétés du réseau, encore inexploitées avec la généralisation des vitesses, et par le formalisme informatique, dans la mesure où l'établissement de règles simples nous a déjà permis de conceptualiser un modèle d'évolution des réseaux de transport, qu'il s'agit de mettre à l'épreuve de l'empirie.

Ce chapitre questionnera dans un premier temps l'impact de la structure et de la hiérarchisation des vitesses dans les capacités circulatoires du réseau ferroviaire français (8.1.), qu'il s'agira de confronter aux éléments de contexte de la croissance du réseau, pour en comprendre les mécanismes les plus élémentaires. Dès lors, il s'agira d'étudier dans quelle mesure ces simples règles permettent de reproduire la phase de croissance du réseau d'intérêt national (8.2.), à partir d'une variation progressive des performances et des paramètres du modèle, qu'il s'agira d'expliquer. La comparaison d'un réseau simulé et d'un réseau réel à plusieurs échelles permet d'éclairer les phénomènes de décalage et de dégager certaines pistes d'explication dans les tracés (8.3.), de même que des dispositifs pour adapter le modèle.

8.1. L'émergence géohistorique de corridors par les capacités circulatoires du réseau

La première phase de croissance du réseau ferroviaire français s'accompagne de la *« recherche de la bonne direction »* (Caron, 1997). Alors, le tracé général est décidé par la loi relative aux chemins de fer de 1842, qui vise le transit international ainsi que le renforcement de l'unité nationale à travers la centralisation des pouvoirs. La planification du réseau en étoile est décidée dès 1838, avec sept axes rayonnants et deux axes transversaux. Puis *« chacune de ces lignes devait être complétée par des embranchements afin de quadriller le mieux possible l'espace national »* (Caron, 1997) : l'objectif est ici de mieux saisir cette expression *« le mieux possible »* en définissant quels critères conduisent à la croissance du réseau. Le parti-pris de cette section, et plus globalement du chapitre, n'est pas d'interroger l'archive, ni même la commune qui l'a beaucoup été jusqu'à présent. Nous nous concentrons ici sur chaque lien du graphe, défini comme l'unité spatiale et l'individu statistique qui va guider notre réflexion. Ces propos s'inscrivent dans une analyse de la réticularité du réseau ferroviaire français, définie comme *« la capacité d'une structure à s'organiser et à fonctionner en réseau »* (Bavoux et al., 2005), à partir de trois vocables de la controverse portant sur les effets structurants : hiérarchisation, concentration et sélection.

8.1.1. Les capacités circulatoires sur le temps long : hiérarchisation et concentration

En faisant du lien du graphe la structure élémentaire de l'analyse qui va suivre, nous nous détachons dès maintenant des études qui ont porté sur l'analyse globale des graphes modélisant les réseaux de transport, ferroviaires (Dancoisne, 1984) ou autoroutiers (Dupuy et Stransky, 1996). En effet, s'ils permettent de catégoriser des régions plus ou moins avantagées par les structures du réseau, les indicateurs globaux des graphes ne permettent pas d'évaluer la contribution d'un lien en particulier dans la structure générale du réseau de transport. Tout au plus, l'étude globale et sur le temps long des graphes, nous permet de constater une évolution logistique du nombre de liens et de nœuds, de la même manière que nous avons pu historiser les phases de croissance du réseau à partir du simple classement des voies ferroviaires *(cf. 3.2.2, p. 123)*. L'objectif est ici différent : indépendamment des nœuds du réseau, et de leurs propriétés, il s'agit d'évaluer les capacités circulatoires du réseau ferroviaire français. Ainsi, alors que l'analyse des accessibilités a montré une hiérarchisation du réseau ferroviaire en termes de vitesse, nous mesurons ici les potentiels flux de déplacements générés par le réseau ferroviaire. Pour donner une signification géohistorique à une telle analyse, nous l'effectuons dans une double dimension temporelle, qui nous suit dans l'ensemble de nos réflexions :

- A partir d'une situation contrefactuelle : en quoi la différenciation par les vitesses participe à la hiérarchisation des flux ? Alors, contrairement aux études précédentes, la généralisation de l'information sur les vitesses permet d'étudier ces différentiels sur l'ensemble du territoire ;
- Sur le temps long : est-ce que l'expansion du réseau participe à la hiérarchisation de plus en plus intense des flux de circulation ?

A la manière de ce qui a été fait dans le chapitre précédent, il s'agit de décomposer les doubles effets de la structure du réseau sur sa hiérarchisation : ceux qui sont directement liés à la forme du réseau, ceux que l'on peut imputer à la discrimination par la vitesse. Nous avons pour l'instant une approche temporelle.

Une telle mesure s'effectue à partir de l'étude de l'intermédiarité, directement issue du formalisme des graphes *(cf. 4.2.3, p. 160)*. Cet indicateur permet de rendre compte de la répartition des plus courts chemins à l'intérieur du graphe : il est influencé par la structure du graphe quand il n'est pas valué, et aussi par les vitesses quand, en revanche, le graphe est valué. La comparaison de l'indicateur issue des deux cas permet finalement de rendre compte de l'effet marginal des vitesses. L'intérêt géohistorique réside dans le fait que le chemin de fer permet pour la première fois d'introduire des différentiels importants de vitesses (Ollivro, 2000). Plus la valeur de l'indicateur d'intermédiarité est élevée, plus le lien voit passer un flux potentiel important, que l'on traduit comme une capacité circulatoire du réseau. Il s'agit ici d'étudier la distribution et la dispersion statistiques des liens en fonction de ces indicateurs.

La *Figure 8. 1* confronte la distribution des capacités circulatoires en 1860 et en 1910, à partir de la seule structure du réseau, indépendamment des nœuds et de la qualité de l'infrastructure. Les données ont été standardisées pour pouvoir être comparées. Il s'agit de la situation contrefactuelle qui permet d'interroger l'évolution de la forme du réseau. On retrouve tout d'abord une distribution souvent rencontrée dans l'étude des réseaux de transports routiers et ferroviaires (Bavoux et al., 2005). Une telle distribution dissymétrique, en référence à la distribution de Poisson, se rapproche de la définition de réseaux dits aléatoires où le plus grand nombre de nœuds contient entre deux et quatre branches (Barabási et Albert, 1999). Surtout, on retrouve la même dispersion des flux que celle que l'on a constatée et expliquée dans le cadre de l'exploration théorique du modèle d'évolution d'un réseau de transport *(cf. 6.3, p. 240)*.

Ainsi, les liens possédant une capacité circulatoire faible sont très importants comparativement aux liens qui voient traverser un flux important. Ce différentiel est ici seulement la conséquence de la structure du réseau : la distribution des capacités démontre ainsi une structure en arbre du réseau ferroviaire français, conformément aux règles édictées dès 1838 par le corps des Ponts et Chaussées dans le dessein du réseau d'intérêt national. La permanence de cette structure en 1910 montre qu'en dépit des 19 000 kilomètres du plan Freycinet, la stratégie structurelle des initiateurs du réseau, à toutes les échelles, suit donc cette logique en arbre. L'expansion du réseau, avec la multiplication des niveaux dans les embranchements, aboutit même à une hiérarchisation exacerbée des liens du réseau sur le temps long *(Figure 8. 1. b.)*, où l'on observe une rupture entre les deux premières classes de la distribution qui est beaucoup plus forte qu'en 1860. Par ailleurs, la

comparaison des données standardisées[55] nous autorise à affirmer que sur le temps long, l'expansion de la structure du réseau participe à un phénomène de concentration, dans la mesure où certains liens concentrent un flux en 1910 bien plus important qu'en 1860.

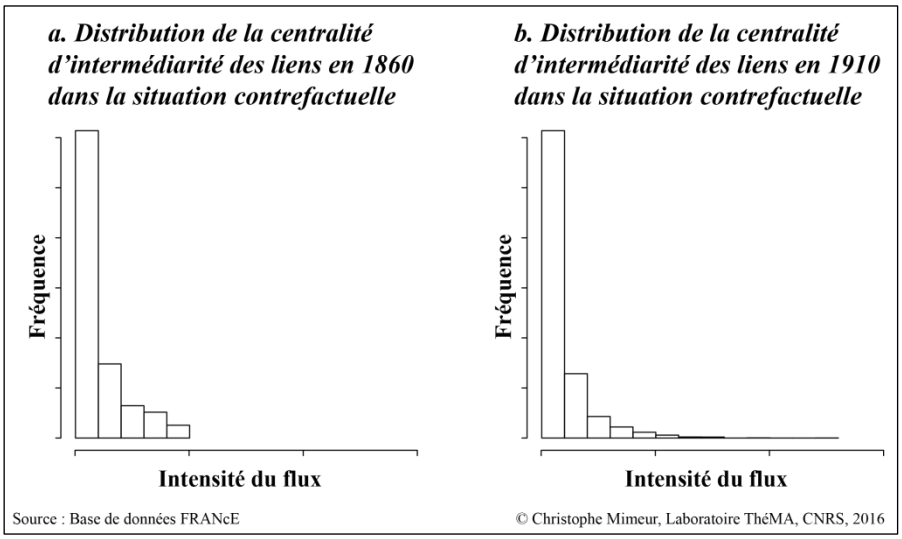


Figure 8. 1. La distribution des flux par la structure du réseau en 1860 et en 1910

Désormais, on peut alors étudier l'effet marginal de la différenciation par la vitesse dans ce processus de hiérarchisation. Nous prenons ici l'exemple du réseau ferroviaire en 1870 *(Figure 8. 2)*. Dans les deux cas, en prenant en compte ou non la vitesse sur le réseau, la structure en arbre permet une forte hiérarchisation des flux de déplacement, qui est encore une fois exacerbée par la différenciation par la vitesse. La comparaison des deux situations permet de voir comment la vitesse permet de concentrer les flux sur certains liens. Etonnamment, ce n'est pas sur la distribution des liens les plus faibles que l'incidence est plus grande. Le phénomène de concentration des flux opéré par la différenciation des vitesses ne suit pas le même processus que sur le temps long : la rupture entre les liens faibles et forts n'est pas significativement accentuée.

C'est dans un processus de sélection que la différenciation par la vitesse permet encore d'accentuer la concentration des flux parmi les liens déjà les plus forts dans la structure du réseau homogène. Pour autant, à partir de 1910, on remarque que la différenciation par les vitesses introduit elle aussi une rupture plus marquée avec les liens qui accueillent le flux le plus faible.

[55] Cela nous permet de contrôler l'effet de l'expansion du réseau qui fait augmenter de fait le nombre de plus courts chemins.

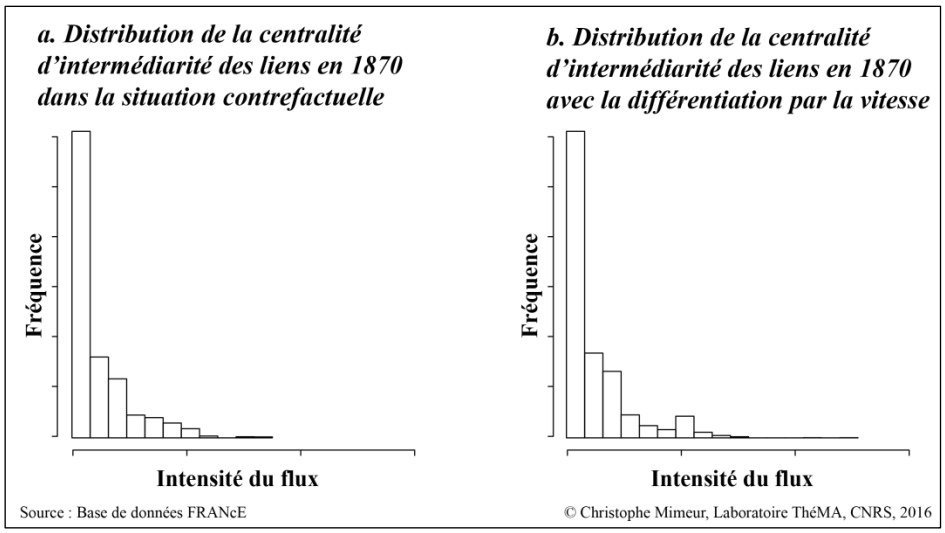


Figure 8. 2. La distribution des flux selon la structure du réseau et les vitesses en 1870

Ainsi, cette première approche de la distribution des flux a permis de constater un double effet dans la hiérarchisation progressive des liens du réseau : l'expansion du réseau fait perdurer la structure en arbre initiée dès la mise en place du réseau d'intérêt national, attestant d'un phénomène de concentration des flux dans les lignes principales. Par ailleurs, la différenciation par la vitesse vient accentuer la hiérarchisation à l'intérieur même du fonctionnement du système. Le phénomène de sélection se fait par extension des liens qui concentrent déjà des flux importants jusqu'en 1910, aux dépens des liens les moins fréquentés à partir de 1910.

Cette étude vient confirmer les hypothèses posées dans l'exploration du modèle théorique, parmi lesquelles le processus de hiérarchisation est à l'œuvre sur le temps long et dans lequel l'inertie de l'infrastructure lourde semble prégnante. Par ailleurs, le processus de sélection semble davantage l'apanage de la différenciation par la vitesse. Cette première analyse vient apporter des éléments de démonstration quant à la construction de règles simples pour modéliser la croissance du réseau de transport sur le temps long : le modèle théorique traduisait un parallèle entre hiérarchisation et sélection *(cf. Figure 6. 19, p. 255)*. L'analyse empirique de la distribution des flux vient corroborer cette importante dimension cinétique dans le temps, qu'il s'agit désormais d'étudier dans l'espace.

Nous nous déplaçons alors dans l'explication géohistorique, où nous confrontons éléments de contexte et analyse inductive pour comprendre les mécanismes techniques, politiques et économiques qui opèrent dans les processus de concentration et de sélection des liens les plus fondamentaux au fonctionnement du système ferroviaire français.

8.1.2. **Gains et pertes d'intermédiarité sur le temps long**

Cette sous-section vise à caractériser le processus de hiérarchisation et de concentration des flux sur le territoire français et sur le temps long : l'étude des tracés permet de comprendre en quoi *« elle ne réduisait pas à une confrontation d'intérêts entre les différentes localités susceptibles d'être desservies. Il s'agissait également d'une affaire de doctrine opposant les partisans des tracés directs à ceux de tracés indirects, desservant, au prix de détours et d'un allongement des distances, les centres d'activité situés sur le parcours de la direction principale »* (Caron, 1997). Cette analyse doit ainsi nous permettre d'étudier les différents facteurs qui ont influencé les tracés, pour mieux envisager plus tard leur reproductibilité par le modèle d'évolution des réseaux de transport. Trois facteurs ressortent dans les stratégies : la direction des tracés, les courbes selon leur rayon et leur pente ainsi que la concurrence avec les réseaux déjà existants. Ainsi, l'hypothèse qui sous-tend cette analyse est la dimension d'aménagement du territoire dans l'explication des tracés, les choix de hiérarchisation et de sélection, dans la mesure où le plus souvent, *« aucune ligne n'offrait un avantage matériel décisif »* (Bavoux, 1994) aux dépens d'autres.

En mobilisant l'intermédiarité des liens, nous posons la question de quels flux, quels liens, quels couloirs ont été pensés et privilégiés pour l'armature du réseau ferroviaire français ? Comment la différenciation des vitesses vient corriger les effets de la structure du réseau et au détriment de quels axes la sélection de certains corridors s'opère-t-elle ? Pour cela, nous comparons encore une fois la situation contrefactuelle issue de la seule structure du réseau à la situation historique discriminée par la qualité des infrastructures. Cette analyse s'appuie avant tout sur une comparaison cartographique et visuelle, dont l'œil du lecteur sera également guidé par le découpage des grandes compagnies de chemins de fer créées après 1857 et la taille des plus grandes villes françaises. Quatre clichés sont choisis pour conduire cette analyse : au lendemain de la mise en connexité du réseau en 1860, au moment de l'inflexion initiée vers 1880, au milieu de la phase d'expansion liée au plan Freycinet en 1900, et enfin en 1930 pour caractériser l'extension maximale du réseau. Avec cette double hiérarchisation, et *« dans une telle structure, certaines zones et certaines relations furent nécessairement marginalisées »* (Caron, 1997). C'est pourquoi l'étude porte sur la superposition des deux situations qui est exposée dans les deux sens : la première permet d'insister sur les axes où la vitesse renforce la position des liens dans le fonctionnement global du réseau, la deuxième traduit en revanche les liens et les axes sur lesquels la discrimination par la vitesse occasionne une perte de flux de circulation. Dans chacun des cas, la distribution statistique des flux observés grâce à l'indicateur d'intermédiarité est très dissymétrique. C'est pourquoi nous choisissons une discrétisation par intervalle géométrique (Béguin et Pumain, 1994), qui permet de porter notre regard sur les différentiels en termes de concentration des flux.

La *Figure 8. 3* montre la situation du réseau seulement cinq ans après sa mise en connexité complète. Le premier élément à souligner est l'impact fort de la structure du réseau dans les flux de circulation, tant les ajustements par les vitesses ne sont que marginaux *(Figure 8. 3. a.)*. Les flux sont alors guidés par les axes rayonnants décidés par la loi de 1842 qui instaure l'Etoile de Legrand. En revanche, les axes transversaux entre Atlantique et Méditerranée, ainsi qu'entre Saône et Rhin, ne concentrent pas les flux les plus hauts dans la hiérarchie.

La structure en arbre impose à elle seule la hiérarchie des flux, en distribuant les capacités *« selon une structure en arêtes de poisson »* (Caron, 1997). C'est d'autant plus vrai quand les axes rayonnants ont une grande portée lorsqu'ils ne rentrent pas en concurrence avec une branche voisine. Tout au plus, la redistribution des flux par l'organisation des vitesses donne une importance plus forte à l'axe Paris-Bâle. Alors, logiquement, on voit que c'est l'axe à l'est de Reims qui perd de son avantage *(Figure 8. 3. b.)* : cet effet de concurrence se fait à l'intérieur même de la Compagnie des Chemins de fer de l'Est. Ailleurs, la stratégie de chaque compagnie répond aux visées hiérarchiques, notamment pour la compagnie Paris-Lyon-Méditerranée et pour la Compagnie des Chemins de fer d'Orléans. A l'ouest, le développement est encore limité vers la Bretagne. La concurrence fait rage entre les branches vers Caen et Le Havre. On observe la même concurrence dans la Compagnie des Chemins de fer du Nord pour l'axe se dirigeant vers Lille, selon que l'on passe par Amiens ou non.

Quand on regarde la répartition des foyers de population, on voit bien la visée des ingénieurs et des politiques à rejoindre les villes de premier ordre (Caron, 1997). Il s'agira d'étudier dans quelle mesure les tracés répondent à une structure plus ou moins partielle de la hiérarchie urbaine quand on simulera la croissance du réseau de transport. Ainsi, la seconde phase de développement du réseau d'intérêt national est l'occasion de redistribuer les flux de circulation selon la structure du réseau et la différenciation par la vitesse.

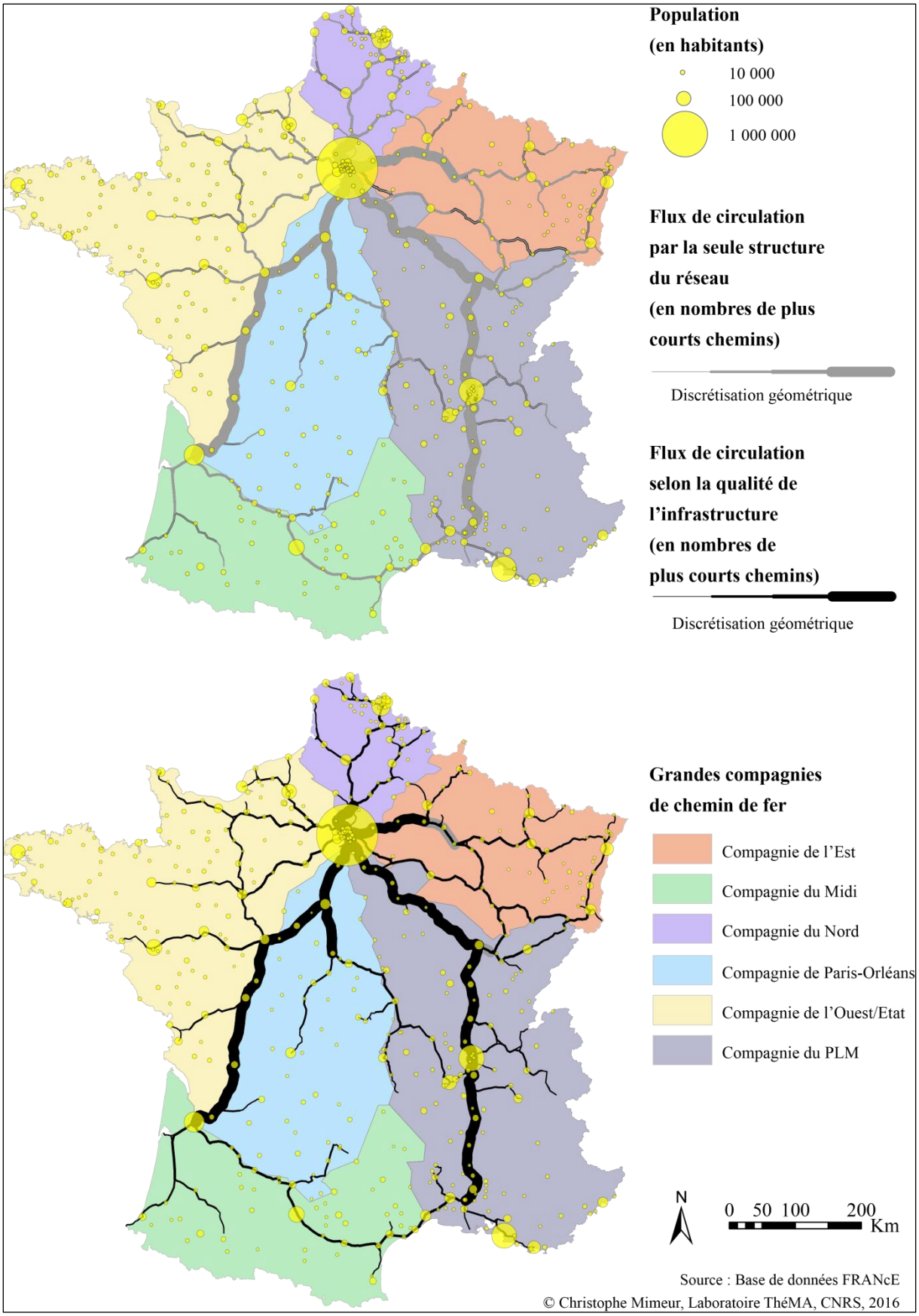


Figure 8. 3. Les capacités circulatoires du réseau ferroviaire français en 1860

La *Figure 8. 4* montre la situation en 1880. Un premier aperçu montre comment la situation diffère d'il y a vingt ans. La première phase d'expansion introduit une nouvelle distribution des flux qui sera encore davantage corrigée par la discrimination par la vitesse. La structure du réseau vient modifier la logique de radiales qui était pourtant bien à l'œuvre en 1860. Elles conservent toujours une concentration importante des flux de circulation, mais la rupture avec d'autres axes n'est plus aussi flagrante. Cette nouvelle étape voit alors l'apparition de nombreuses transversales dans le réseau, dont le maillage progressif laisse à penser à une homogénéisation des flux de circulation *(Figure 8. 4. a.)*. Deux catégories de transversales peuvent être dégagées : la première consiste en une couronne autour de Paris, qui passe par Amiens, Le Havre, Chartres, Orléans, Montereau puis Reims. Les secondes transversales sont encore davantage traversantes, comme les prémices d'un axe Centre-Europe-Alantique par la vallée de la Loire puis de la Saône. Et puis il y a aussi le contournement Nord du Massif Central pour rejoindre Bordeaux par Limoges.

Malgré tout, la discrimination par les vitesses continue de renforcer la concentration des flux sur les branches radiales, et surtout à proximité de Paris. Le renforcement des branches de l'Etoile de Legrand est alors toujours à l'œuvre près de 35 ans plus tard. Le caractère d'épine dorsale ou de colonne vertébrale devient incontestable quand l'aménageur discrimine son réseau par la vitesse. Cette plus forte concentration se fait surtout sentir quand deux axes concurrents sont proches : vers Bordeaux, la branche passant par Tours et Poitiers l'emporte sur celle qui passe par Limoges, tandis qu'à l'ouest, le passage par Le Mans et Chartres est renforcé par rapport à la ligne droite entre Brest et Paris. Enfin, à l'est, le chemin vers Nancy concentre davantage de flux que celui qui passe par Reims. En dernier lieu, on remarque le renforcement de quelques axes à proximité de carrefours, comme autour de Dijon, Strasbourg, Le Mans, Bordeaux, Toulouse et Montpellier, préfigurant la configuration de sous-systèmes régionaux, et laissant entrevoir la prise en compte d'une structure plus complète de la hiérarchie urbaine, que nous vérifierons également dans la suite du chapitre.

C'est ici que nous réinjectons l'ensemble des éléments de la définition d'un réseau (Curien, 2000 ; Dupuy et Offner, 2005) : faire réseau suppose non seulement une infrastructure support, mais aussi une confrontation entre la visée de l'ingénieur et de l'économiste. En la matière, si la concentration des grandes compagnies à la fin de la décennie 1850 a permis de mettre une fin aux troubles[56] qui faisaient que la fièvre du chemin de fer ne prenait pas, la logique réticulaire adoptée par chacune de ces compagnies tend à construire des sous-systèmes qui nécessitent des possibilités de rentabilité. Si la concurrence reste modérée quant aux politiques de tarification (Caron, 1997), le choix de la discrimination des vitesses montre comment les liaisons entre les territoires des différentes compagnies n'est pas leur priorité *(Figure 8. 4. b.)*. Ainsi, *« bien plus que la simple liaison ferroviaire (circulation d'un train entre deux points), c'est la qualité de la relation ferroviaire d'un lieu à un autre qui transforme le voyageur potentiel en utilisateur effectif du chemin de fer »* (Varlet, 1990). Dans ces propos, Jean Varlet étudie la qualité des relations interrégionales

[56] L'absence de concession de longue durée engendrait un manque de visibilité quant aux capacités de rentabilité des lignes et le nombre important de compagnies aboutissait à un manque de cohérence et à une segmentation des coûts. La concentration des compagnies a alors permis de considérables économies d'échelles, alors que le recours croissant à la garantie d'intérêt a rassuré les investisseurs

en 1980 : cette citation trouve un écho particulier ici, tant on voit que la stratégie infrastructurelle des compagnies ne fait pas des relations interrégionales des couloirs capables de drainer des flux importants.

Ainsi, des axes transversaux décrits dans le paragraphe précédent, tous sont pénalisés par la vitesse, surtout dans la couronne qui entoure largement l'Île de France. L'utopique « Grand Transversal » entre Lyon et Bordeaux projeté par Freycinet ne semble en tout cas pas la priorité des investissements, tant il est réfuté à la fois par les ingénieurs techniquement et économiquement (Ribeill, 1985a). Cet état de fait semble même créer des discontinuités là où la seule structure du réseau ne permet pas d'en identifier : c'est le cas de la vallée de la Durance, en Ardèche, en Haute-Saône ou au cœur de la Bretagne. Ici clairement, les obstacles topographiques semblent avoir influencé le choix stratégique des aménageurs de l'époque.

De même, les axes qui se trouvent au voisinage de la frontière entre les territoires des compagnies connaissent des destins contrariés. C'est le cas de la remontée vers le nord par le Massif Central depuis Montpellier, mais aussi des possibilités de jonctions entre les Compagnies de l'Ouest et du Nord, dont la concurrence fait rage quant aux transports de houilles en provenance du Royaume-Uni (Caron, 1997). La situation la plus floue est entre la Compagnie des chemins de fer de l'Ouest et celle du Paris-Orléans. La seconde arbitre la concurrence entre les axes Limoges-Orléans-Paris et Bordeaux-Poitiers-Orléans-Paris. L'avantage est laissé au second, aux dépens de l'axe vers Toulouse, qui se dirige vers la Compagnie des Chemins de fer du Midi. A l'est, la stratégie de la compagnie donne l'avantage à l'axe vers Strasbourg en passant par Reims et Nancy. Pour autant, certains voient dans la multiplication des chemins de fer d'intérêt local avec le plan Freycinet le passage d'un *« réseau étoilé »* en *« un réseau en échiquier »* d'après Proudhon, où *« la pensée princière, gouvernementale, du réseau étoilé aura disparu dans la multitude des lignes de jonction »* selon Léon Lalanne. Dans l'application de la théorie des réseaux, ce supposé passage traduirait un primat pour la diffusion de l'innovation, au détriment des performances circulatoires. Si la réalité nous apprend que finalement, la structure radiale demeure un élément fort pour les aménageurs de l'époque, nous verrons comment le modèle traduit également cette vision cinétique qui privilégie la branche commune.

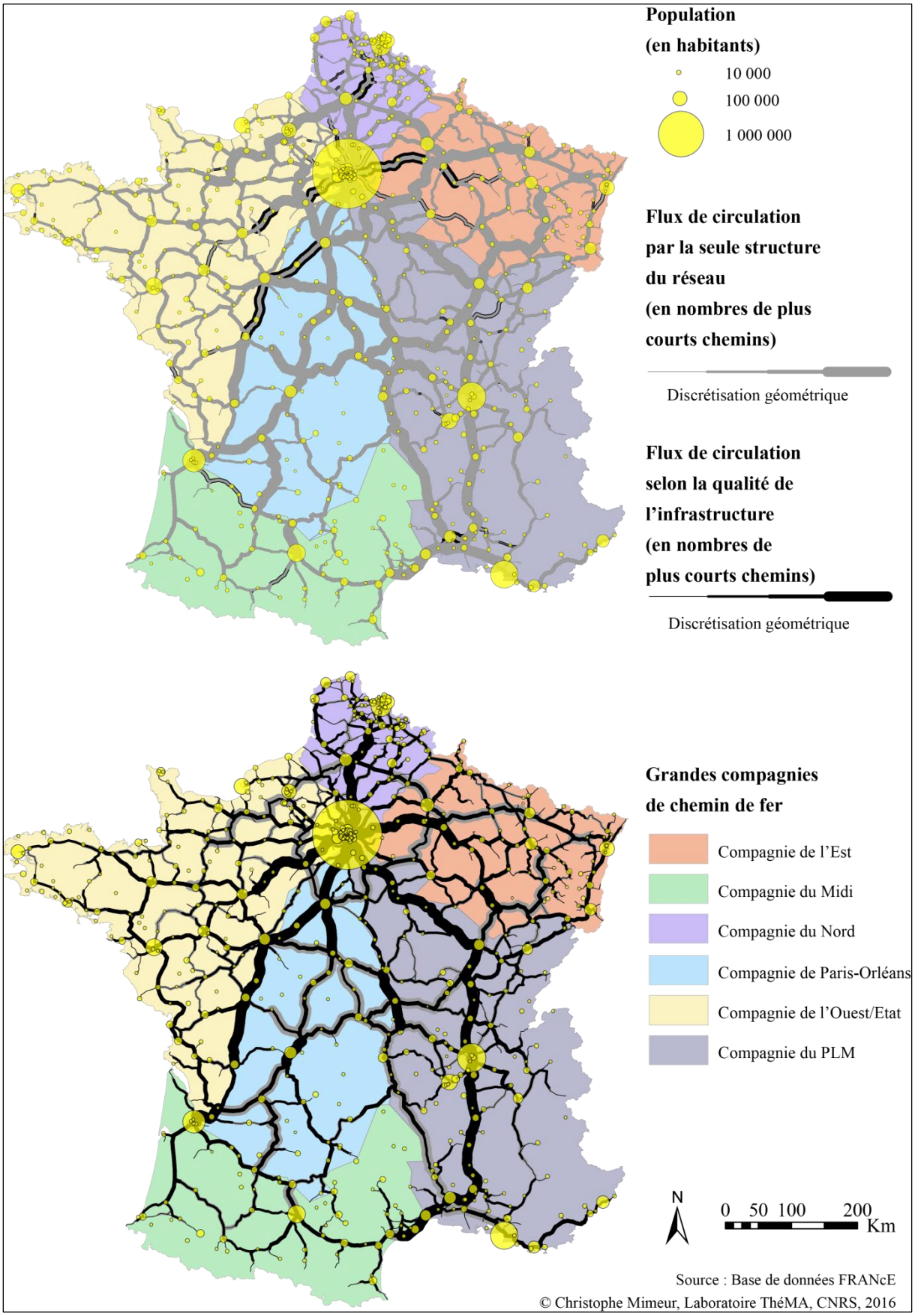


Figure 8. 4. Les capacités circulatoires du réseau ferroviaire français en 1880

Quand on étudie la configuration des capacités circulatoires en 1900, plus de la moitié du plan Freycinet est abouti *(Figure 8. 5)*. Le réseau en échiquier n'est pas à l'œuvre mais on retrouve le caractère en arbre où des embranchements successifs viennent rejoindre les épines dorsales du réseau *(Figure 8. 5. a.)*. Cette répartition spatiale vient par ailleurs confirmer le caractère hiérarchique que l'on a observé dans la permanence de la forme de la distribution des flux *(cf. 8.1.1, p. 313)*. Il est en revanche plus exact de dire que cette structure ne donne plus une place prééminente aux radiales décidées au milieu du XIXème siècle, parce que certains axes transversaux présentent des capacités tout aussi fortes qui permettent de contourner Paris. En revanche, la surimposition de la situation discriminée par les vitesses montre que l'ascendant est pris encore une fois par les radiales. Etonnamment, celle en provenance de Marseille ne semble souffrir d'aucune concurrence qui nécessite de la mettre en avant. Alors, le privilège de la vitesse permet d'arbitrer des situations toujours concurrentielles entre Paris et l'Ouest de la France, entre Paris-Metz-Strasbourg, Paris-Nancy-Strasbourg et Paris-Bâle à l'Est, entre le passage par Amiens ou Saint-Quentin au Nord de la France. Il est alors intéressant de souligner que la concentration permise par la vitesse permet une forte discrimination entre nord et sud de la France, démontrant toujours aussi fort la vigueur de l'attraction parisienne dans les desseins stratégiques des compagnies. Les impacts dans le sud sont plus ponctuels, où on note particulièrement la prime donnée à la circulation vers Toulon et vers Montpellier depuis l'axe Paris-Lyon-Marseille.

Ainsi, en 1900, les principaux axes qui doivent subir l'effet de la vitesse dans leurs capacités de circulation sont ceux qui se trouvaient dans les situations de concurrence *(Figure 8. 5. b.)*. Le passage par Nancy depuis Strasbourg pour rejoindre Paris semble défavorisé, de même que le cheminement vers Dijon, où l'on préfère l'axe par Troyes et Reims. Pourtant, Nancy semble un foyer de population important, attestant donc ici plutôt de choix politiques. A l'ouest, l'axe serré entre les segments Le Mans-Chartres et Tours-Orléans-Paris est lui aussi délaissé. Peu à peu, les axes de prolongement depuis Limoges vers Bordeaux et Toulouse montre que l'axe Bordeaux-Orléans-Paris prend le pas sur les autres. En revanche, on voit que certaines transversales arrivent à conserver des capacités de circulation : le choix de Reims dans l'axe Est-Ouest renforce aussi l'axe Nord-Sud entre Vitry et Laon, permettant des interconnexions progressives entre les compagnies.

Dès lors, on voit bien que la stratégie radiale demeure en dépit de la large diffusion du chemin de fer, et les principales « victimes » se trouvent dans le voisinage immédiat de ces axes dont il a été décidé qu'ils étaient structurants.

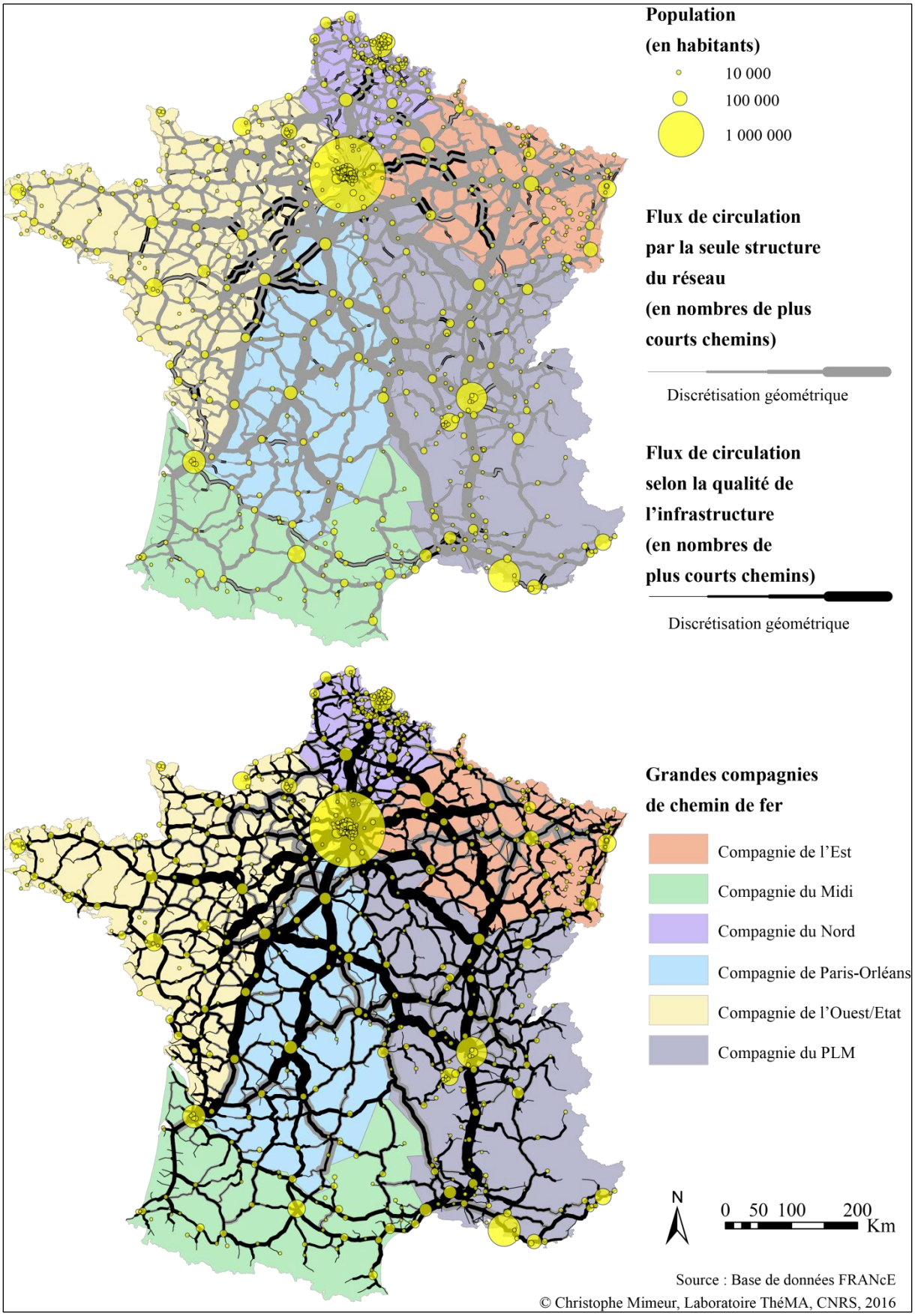


Figure 8. 5. Les capacités circulatoires du réseau ferroviaire français en 1900

En 1930, le réseau est à son expansion maximale et de nouvelles améliorations sur le réseau rapide ont été effectuées *(Figure 8. 6)*. Ainsi, à son apogée, la structure du réseau ferroviaire, malgré son maillage important, conserve une forme en arbre, caractéristique de nombreux autres réseaux ferrés (Bavoux et al., 2005). La structure est ainsi stabilisée depuis 1900 malgré l'ouverture de nouvelles lignes locales, dont l'impact n'est donc que marginal sur la structure globale du réseau *(Figure 8. 6. a.)*. En revanche, la distribution par les vitesses vient confirmer certains choix déjà à l'œuvre à l'aube du XXème siècle. A l'est, l'axe par Nancy et Reims s'impose désormais face à l'axe qui passe par Metz, tandis que l'axe Paris-Bâle gagne en capacité circulatoire. La compagnie du PLM, peu concurrencée vers le sud de la France, voit le renforcement de l'axe depuis Nevers, qui entre alors dans la course à la concurrence avec la compagnie du Paris-Orléans où l'axe depuis Tours et Orléans semble lui aussi s'imposer, alors que la compagnie de l'Ouest renforce les capacités de l'axe qui permet de joindre Bordeaux au Mans par Niort. Plus au Sud, la transversale issue de la loi de 1842 semble trouver une nouvelle vigueur, tant il a été souligné la bataille de la compagnie de Chemin de Fer du Midi face au délaissement de l'Etat (Bouneau, 1990). Enfin, au nord, l'arbitrage demeure difficile et les deux axes conservent de très grandes capacités.

Ces ultimes renforcements se font au détriment d'axes transversaux, surtout ceux qui sont à proximité de Paris *(Figure 8. 6. b.)*. Seules la vallée de la Loire, la liaison entre Lyon et Clermont, entre Dijon et Reims, entre Reims et Saint-Quentin conservent des capacités circulatoires intéressantes. Tous les axes potentiellement concurrents sont pénalisés par la discrimination par la vitesse, alors qu'un nombre important et croissant par rapport à 1900 d'axes ne desservant pas de villes de plus de 5 000 habitants voient leur capacité circulatoire faible et un écart important avec les liens qui concentrent davantage de flux.

Cette analyse a permis de mettre en avant d'un point de vue spatial les capacités circulatoires du réseau ferroviaire français sur le temps long. Elle atteste d'une permanence de la structure du réseau en arbre, malgré un maillage de plus en plus fin que certains jugeront même excessif. La stratégie des aménageurs d'une part, des grandes compagnies d'autre part, répond alors bien à la théorie de l'arête de poisson, où l'épine dorsale – la radiale vers Paris – voit l'apparition de nombreux embranchements, eux-mêmes à l'origine d'une nouvelle hiérarchisation. Ainsi, on retrouve ce même double phénomène transcalaire, qui touchait déjà l'indicateur d'accessibilité moyen, pondéré ou non par la population d'ailleurs. De plus, si la structure voit l'apparition progressive de transversales sur le territoire, la discrimination par la vitesse vient atténuer leur importance, perpétrant la suprématie des radiales vers Paris. On y lit la concurrence entre les compagnies, et les liens de passage sont alors peu nombreux. Les premiers perdants d'une telle hiérarchisation sont les voisins de ces radiales, qui concentrent et aspirent l'ensemble des flux environnants, alors qu'en fin de période, les lignes desservant les campagnes semblent connaître un écart grandissant avec le reste de la hiérarchie. Il semble bien que le processus de concentration et de sélection soit à l'œuvre selon des stratégies économiques de rentabilité mais aussi politiques quand l'arbitrage n'est pas décisif. Alors il convient d'étudier plus précisément comment évolue la distribution des capacités circulatoires et comment elle est influencée par la structure démographique, à l'image des aspirations de l'époque (Lalanne, 1863) et des conclusions contemporaines (Bretagnolle, 2009 ; Pumain, 1982).

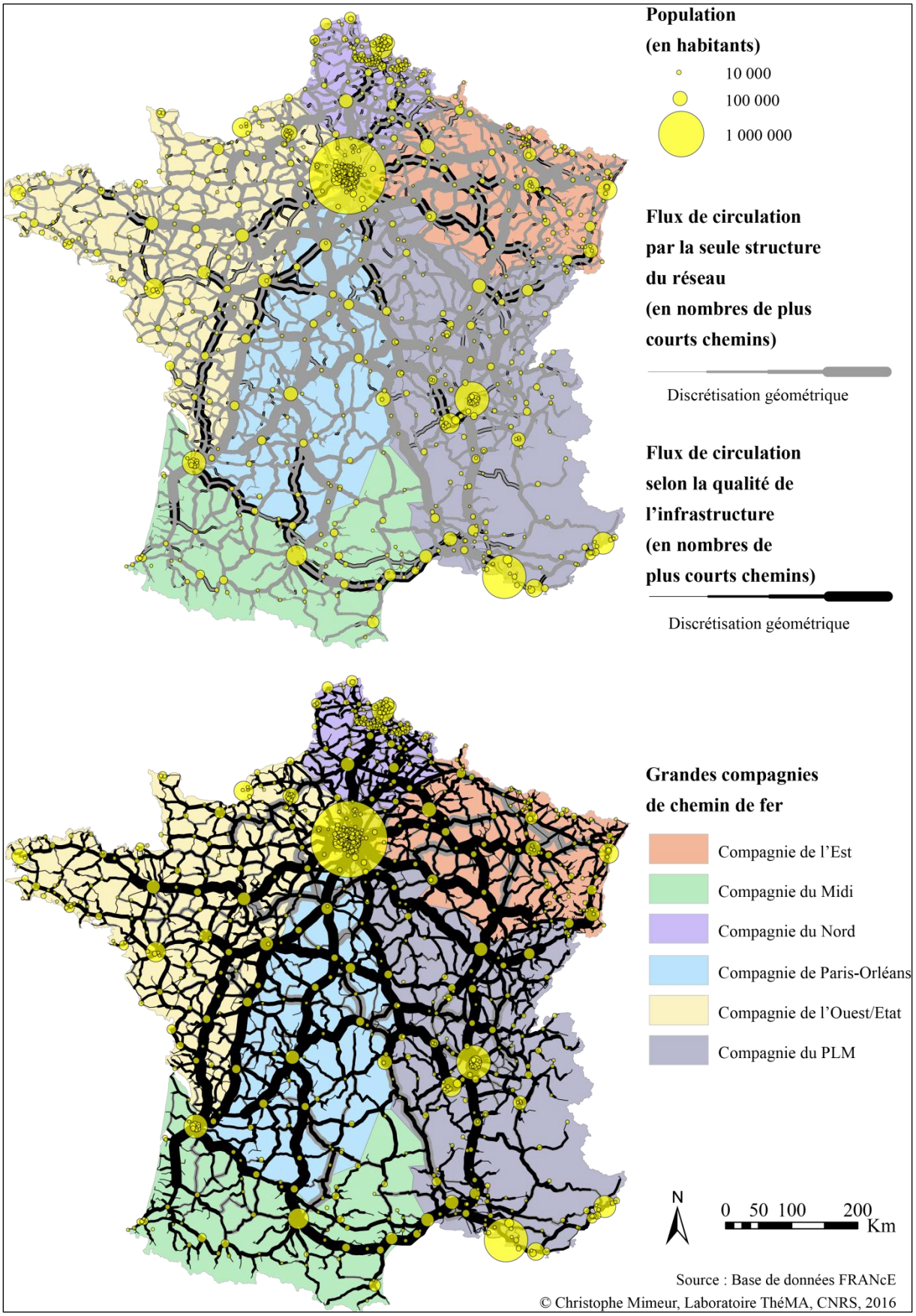


Figure 8. 6. Les capacités circulatoires du réseau ferroviaire français en 1930

8.1.3. **Corridors géohistoriques et niveaux de population**

En créant les grandes compagnies en 1857, les objectifs sont multiscalaires pour l'administration centrale. Le premier est de créer une concurrence modérée pour ne pas que les tarifs de circulation s'envolent et pour permettre à un maximum de voyageurs d'en profiter. Le second est plus caché : l'établissement de ces zones monopolistiques cache en vérité la volonté de renforcer les points de passages obligés, c'est-à-dire les radiales issues du plan qui dessine l'Etoile de Legrand en 1842 (Ribeill, 1985b). A l'intérieur de chaque zone, l'objectif est de venir concurrencer les canaux. A leurs frontières, la bagarre des tracés et la politique de qualité de l'infrastructure se font souvent pour garantir les intérêts de la branche centrale. Mais au-delà de ces stratégies politiques et techniques, l'objectif de l'administration est aussi de rejoindre les villes de premier ordre (Lalanne, 1863). En effet, malgré les volontés vantées de l'homogénéisation, la création de tels monopoles semblaient l'assurance d'une priorité aux chemins de fer économiques à partir de 1860 (Caron, 1997). Il s'agit donc désormais de voir comment les stratégies circulatoires s'appuient sur une structure démographique existante pour des constructions viables, à moindre coût ou en tout cas rentabilisables, en confrontant simplement ces deux variables.

Si on peut rapprocher cette question des centralités déjà étudiées *(cf. 7.2.2, p. 285)*, l'objectif ici est de questionner l'effet dans un seul sens, où les capacités circulatoires du réseau s'adaptent à une structure démographique existante. Est-ce que le tracé du réseau ferroviaire cherche à rejoindre les points extrêmes en ligne droite, ou s'accommode-t-il des noyaux de population rencontrés ? La question est encore une fois abordée selon le double processus à l'œuvre dans la phase de croissance du réseau ferroviaire français : son expansion qui fait évoluer la structure globale du réseau et sa valorisation qui fait émerger une hiérarchie des voies. Ces deux éléments sont perceptibles dans l'étude des capacités circulatoires face à la structure démographique. De fait, nous ne nous intéressons qu'aux communes desservies par le chemin de fer entre 1860 et 1930.

La *Figure 8. 7* confronte la distribution des liens selon leur capacité circulatoire en fonction de la hiérarchie démographique. Pour chaque décennie de l'étude, on regarde comment le réseau s'adapte à la situation démographique qui précède. Par ailleurs, nous ôtons de l'analyse le réseau de 1860 dans la mesure où seule la structure du réseau influence la géographie des circulations au moment de la mise en connexité du réseau. La première partie rend compte de l'adaptation de la structure du réseau, la seconde partie rend compte de l'impact des choix infrastructurels de l'Etat et des grandes compagnies. On voit ainsi comment la structure du réseau ferroviaire français s'adapte sur le temps long à la structure démographique. L'augmentation continue des valeurs illustre l'expansion du réseau et donc le besoin de canaliser davantage de flux. Les liens qui concentrent davantage de capacité circulatoire sont prioritairement dans les villes de plus grande importance, et leur hiérarchie suit la structure existante des populations, dans tous les cas. Au début de la période, seules quelques branches permettent de rejoindre des centres de moindre importance et la structure du réseau ne permet pas de discriminer fortement les capacités circulatoires. En 1870, même l'impact de la différenciation des vitesses n'a toujours pas un rôle prééminent dans la distribution des circulations, où aucun processus hiérarchique n'est à l'œuvre. On pose alors l'hypothèse que cette phase d'installation durable du réseau d'intérêt national est le fruit d'une politique

d'aménagement du territoire décidée il y a déjà fort longtemps et au profit des très grandes villes.

Pourtant, dès 1880, de plus grandes ruptures s'opèrent à l'image de la répartition spatiale observée dans la *Figure 8. 4*. Le renforcement des radiales se traduit par des écarts plus grands dans les capacités circulatoires introduites par les vitesses, davantage que par la seule structure du réseau. Cette stratégie ne suit alors pas forcément la structure démographique, en tout cas pour les plus grands noyaux de population. La décennie suivante, les premiers travaux du plan Freycinet montrent que la structure du réseau a tendance à s'homogénéiser encore davantage avec moins d'écart dans les capacités circulatoires. Si les moyennes n'évoluent pas de manière significative, on constate en revanche que l'introduction des vitesses discrimine davantage certains liens dont la concentration devient de plus en plus importante.

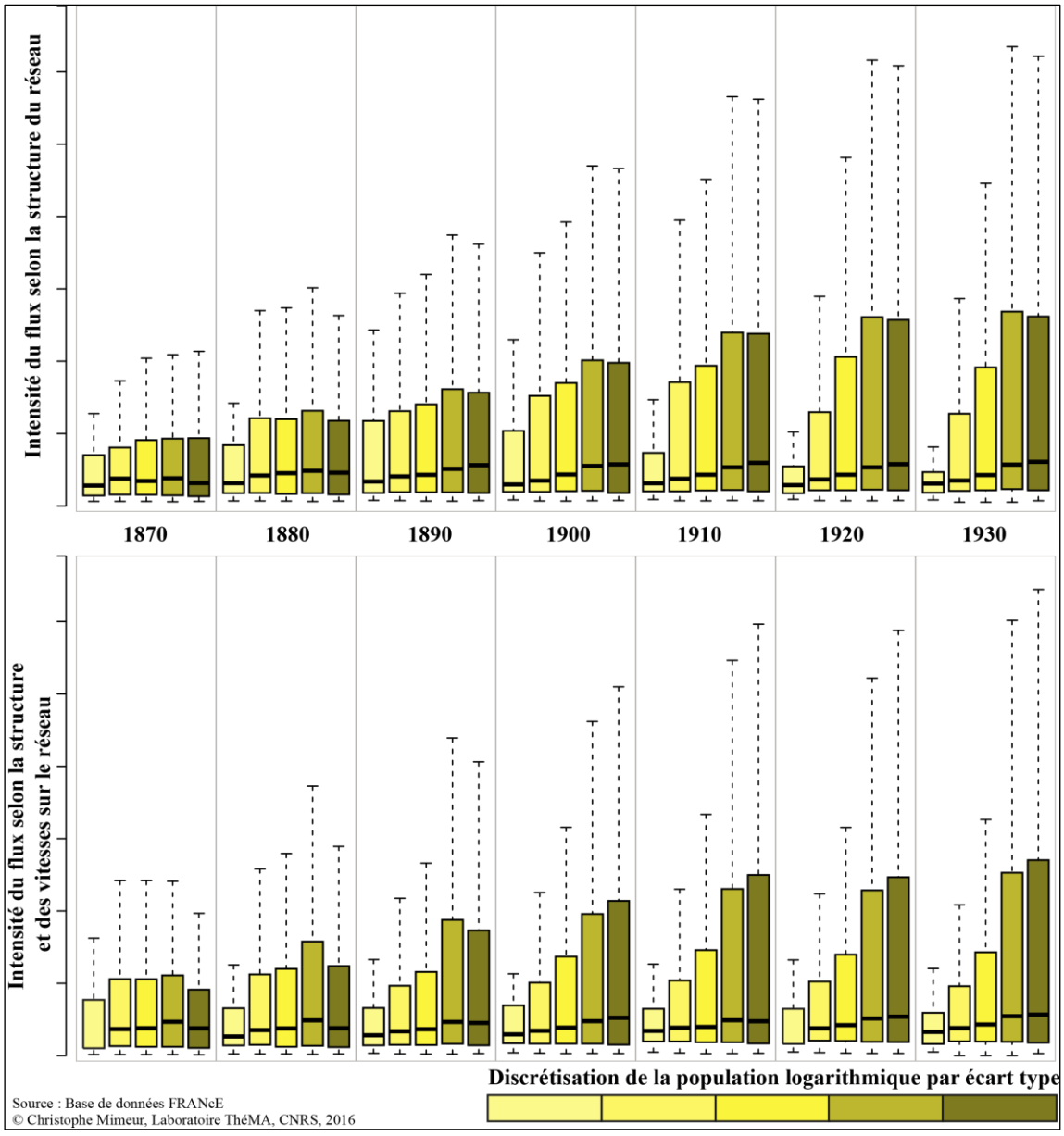


Figure 8. 7. Potentiel de circulation et population entre 1870 et 1930

A partir de 1900, on constate que le processus de hiérarchisation du réseau devient permanent jusqu'à la fin de notre période d'étude et qu'il est le fait du double phénomène lié à la

structure en arbre du réseau et aux choix stratégiques de vitesses le long du réseau. Ainsi, quand le réseau voit des transversales rencontrer des foyers de population importants, l'écart entre les deux dernières classes est réduit voire inversé. En revanche, quand un choix d'implantation des vitesses est effectué, alors la hiérarchie semble rétablie dans tous les cas de notre étude. C'est par exemple le cas de la transversale entre Orléans et Rouen. En 1864, le préfet du Loiret écrivait déjà que *« malgré son utilité relative, la ligne […] ne peut être classée qu'en deuxième ou troisième position »* (Caron, 1997). C'est très perceptible sur les cartes précédentes, où l'on voit le potentiel circulatoire par la structure de plus en plus amoindri par les choix de vitesses. On voit de plus que malgré l'expansion du réseau, les liens à capacité intermédiaire n'évoluent guère dans la hiérarchie alors que la concentration des flux s'opère davantage au détriment des plus petits liens et donc des plus petites villes. Si nous pouvons conclure à une adaptation des capacités circulatoires du réseau à la structure démographique existante, nous ne pouvons pas affirmer que cette adaptation se fait progressivement, au coup par coup, en suivant les fluctuations conjoncturelles. On pose la question d'un phénomène plus structurelle où le réseau s'adapte à une structure préexistante, ancienne, dont la dynamique de fond est très lente et dotée d'une forte inertie (Guérin-Pace, 1993).

Pour cela, nous étudions l'adaptation des capacités circulatoires du réseau entre 1870 et 1930 à partir de la structure démographique qui préexistait en 1866. De cette sorte, on peut juger si le processus de hiérarchie est immédiat ou s'il est le fait d'une dynamique plus longue. La *Figure 8. 8* reprend à la fois la situation contrefactuelle influencée par la seule structure du réseau et la situation qui prend en compte les vitesses. La situation de 1870 est forcément la même puisqu'elle prend en compte la situation démographique de 1866. Au lendemain de la signature du plan Freycinet, le réseau d'intérêt national est constitué. Pourtant, en 1880, la structure du réseau n'obéit pas à une structure hiérarchisée, dans la mesure où l'extension commence à mettre à mal l'organisation en étoile du réseau. Là encore, la discrimination par les vitesses vient corriger le tir et introduire davantage de hiérarchie entre les différents liens, même si les deux dernières classes de population n'obéissent pas à la hiérarchie que l'on pourrait attendre. En revanche, avec les premières ouvertures liées au plan Freycinet, on remarque que la hiérarchie des circulations par la structure du réseau a toujours du mal à s'adapter à la structure démographique préexistante. Là par contre, les choix infrastructurels sont décisifs dans la hiérarchisation des liens dès 1890, où la situation était plus chaotique quand on étudiait les adaptations conjoncturelles. Ainsi à partir de 1900, le double processus de hiérarchisation montre une évolution différenciée, capable d'illustrer la forte inertie liée à la stabilisation d'un réseau sur le temps et où les dimensions cinétiques et adaptatives sont à l'œuvre durant toute la période de croissance du réseau (Dupuy, 1991). Ainsi, alors que la structure du réseau montre une hiérarchie linéaire entre les différentes capacités circulatoires adaptées à la structure démographique, alors le réseau pris dans toute sa dimension – qui allie structure et dynamique – montre que la différenciation des capacités circulatoires en fonction de la structure démographique est plus de forme exponentielle. Par-là même, ce double processus aboutit à une hiérarchisation qui n'est pas le simple fait d'une distribution linéaire, mais qui crée des opportunités de circulation tout en marginalisant d'autres liens. C'est tout le sens que prennent les ruptures plus prononcées entre les différentes classes de circulation quand elles sont confrontées à la population.

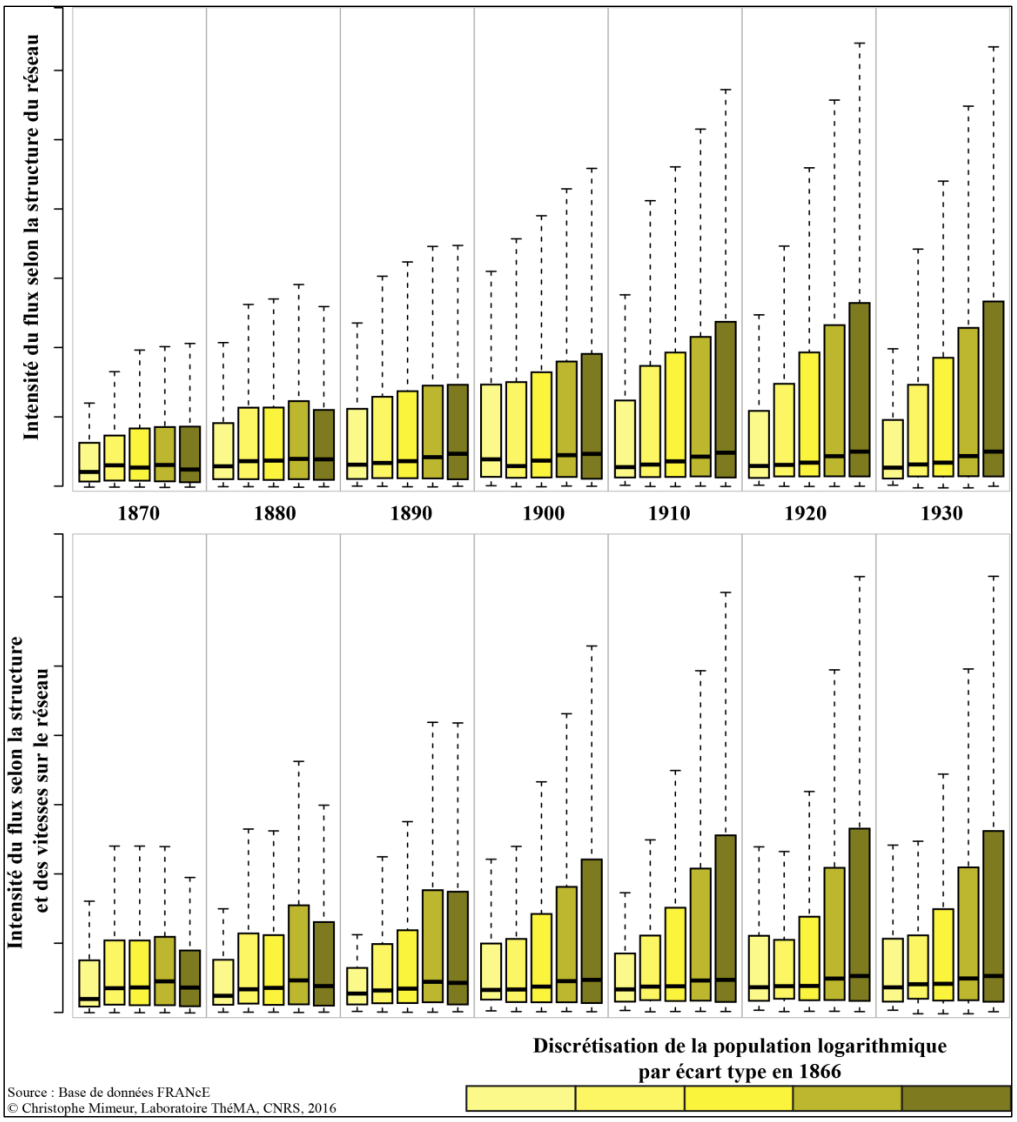


Figure 8. 8. Potentiel de circulation et structure démographique préexistante

Cette même conclusion peut être tirée de l'analyse des communes dont le nombre d'habitants est inférieur à 2 500 habitants *(Figure 8. 9)*. La phase de croissance définie par la théorie des réseaux prend alors ici tout son sens. Les visées d'aménagement sont traduites par l'impact de la structure du réseau dans les capacités circulatoires. Les chemins de fer électoraux prennent leur rôle dans la difficile hiérarchisation des circulations par la structure du réseau, particulièrement dans la période du déploiement du plan Freycinet entre 1880 et 1910. Tantôt les voies présentant les flux les plus faibles traversent des communes peu peuplées, tantôt des voies avec des flux plus forts traversent des campagnes clairsemées en 1860. En revanche, quand on regarde l'impact des choix de vitesses, directement issus des choix d'investissements, on voit comment le système global du réseau s'adapte à la structure démographique de 1860. Cette analyse démontre enfin comment le plan Freycinet a largement été impacté par les disponibilités financières et ne répond donc pas forcément à des logiques réticulaires hiérarchisées[57].

[57] L'étude a également été menée en comparant flux de circulation et situation démographique conjoncturelle dans

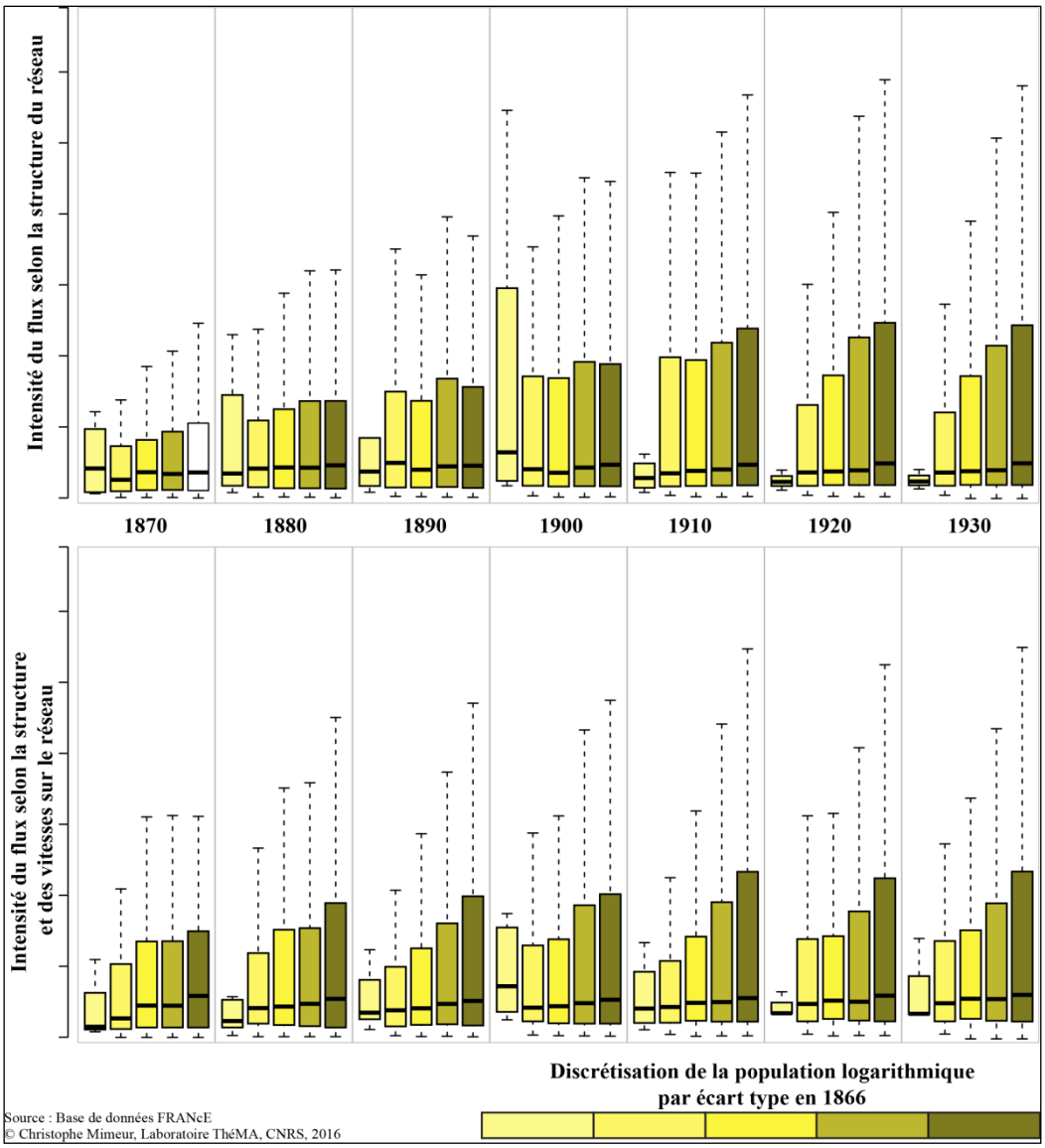


Figure 8. 9. Potentiel de circulation et structure démographique préexistante pour les communes de moins de 2 500 habitants

Ainsi, l'ensemble de cette section a permis de clarifier le vocable *« le mieux possible »* quant à la définition des tracés du réseau ferroviaire français. La littérature pointe le compromis entre trois stratégies (Ribeill, 1985b) : une stratégie des grandes compagnies, une stratégie centralisée et une adaptation à la structure préexistante. L'étude des flux circulatoires a d'abord montré l'évolution des distributions vers davantage de sélection et de concentration des capacités de circulation sur le temps long.

les campagnes. Les résultats ne sont pas montrés ici dans la mesure où l'on constate le même phénomène d'une exécution du plan Freycinet qui n'est pas hiérarchisé mais dont la visée aboutit après 1910 en un réseau hiérarchisé.

L'étude spatialisée des territoires mis en valeur ou marginalisés a permis de montrer l'impact de la stratégie radiale aux dépens des transversales et des lignes d'intérêt local. Par ailleurs, elle a permis de mettre en avant les jeux de concurrence entre des axes à l'intérieur d'une compagnie et entre les territoires de grandes compagnies. Enfin, nous avons vu comment le double processus de hiérarchisation par la structure et les choix d'investissement s'adaptent finalement plutôt bien à une structure démographique préexistante. Cette analyse a permis de rendre mieux compte de la structure, de la dynamique et de la logique (Dupuy, 1991) du réseau ferroviaire français à partir d'une information inédite portant sur les potentialités de circulation en fonction de la qualité de l'infrastructure. Pour évaluer la contribution de chacune de ces stratégies dans les tracés, on se propose de reproduire la stratégie qui s'adapte à la structure préexistante à l'aide de notre modèle d'évolution du réseau de transport. Les écarts au modèle traduiraient alors la force des autres stratégies.

8.2. De l'efficacité radiale à la suggestion de transversales

Si la section précédente a permis de comprendre les stratégies infrastructurelles sur le territoire français pendant la phase de croissance du réseau ferroviaire, on pose l'hypothèse qu'un des facteurs déterminants de l'installation du réseau est son adaptation à une structure démographique préexistante. Nous replaçons cette hypothèse dans les différentes phases de la théorie des réseaux *(cf. Figure 1. 2, p. 33).* Ici, l'application empirique du modèle d'évolution du réseau de transport a pour objectif de reproduire le passage du réseau virtuel au réseau réel (Dupuy, 1991) : en s'appuyant sur le graphe théorique P. Ce dernier traduit hétérogénéité et anisotropie de l'espace, le modèle d'évolution évalue l'ensemble des relations possibles pour ensuite opérer un processus de concentration et de hiérarchie. L'analyse du comportement du modèle sur des formes théoriques montre une adéquation avec le phénomène effectivement observé dans l'évolution des dynamiques circulatoires analysées de manière empirique ci-dessus. Dès lors, nous misons sur la capacité du modèle à opérer les dimensions topologique et cinétique pour apporter de nouveaux éléments d'explication à des affirmations postulant qu'une *« rapide analyse du réseau urbain français aujourd'hui montre une adéquation assez bonne entre la taille des villes et la qualité des infrastructures de transport qui les desservent »* (Baptiste, 2003). Par ailleurs, nous misons sur la capacité du modèle à apporter de nouvelles explications géohistoriques quant à l'agencement des tracés : dans ce cadre, nous interrogeons dans cette section les modalités de l'adaptation du réseau d'intérêt national à une structure démographique préexistante *(cf. Figure 6. 19, p. 255)* : jusqu'à quel niveau / à partir de quel niveau le réseau est-il adapté à la diversité spatio-temporelle des foyers de population ?

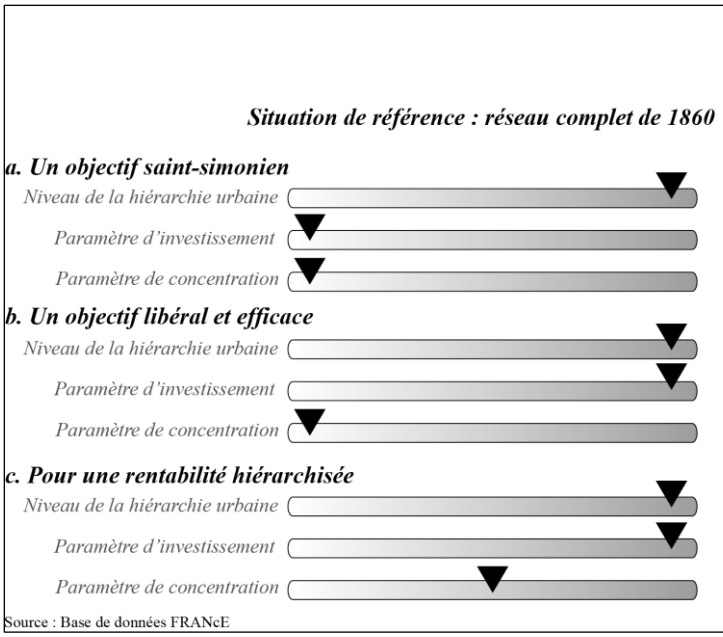
La lecture des simulations s'opère en deux temps. La première est morphologique : quelle est la forme du réseau simulé ? Quelles caractéristiques principales du réseau permettent-elles de reproduire ? La seconde est contextualisée : alors que le précédent chapitre met en valeur la robustesse des estimations, cette analyse vise à analyser les écarts entre le réseau simulé et le réseau réel, comme révélateurs de stratégies qui ont conduit au dessin du réseau ferroviaire.

Nous testons l'évolution du réseau de transport à l'échelle nationale. On le confronte au réseau d'intérêt national stabilisé en 1880, ce qui sera le réseau de référence pour comparer les tracés. Pour tester dans quelle mesure l'expansion du réseau ferroviaire français prend en compte une situation démographique préexistante, nous définissons plusieurs seuils de population, à partir desquels le modèle évalue les projets transactionnels avant d'opérer les mécanismes de concentration et de hiérarchisation pour la croissance endogène du réseau *(Annexe 2)*. Le propos de cette troisième partie est l'évaluation du modèle, c'est-à-dire sa capacité à reproduire le réseau réel : une comparaison systématique des situations simulées et réelles permet de comprendre l'ampleur du critère d'efficience économique et démographique dans le dessin du réseau. Les écarts sont quant à eux à interpréter comme la prédominance d'autres critères, même si l'explication de tous les tracés n'est pas le but ici. La confrontation est chiffrée mais aussi visuelle et complétée par les connaissances géohistoriques existantes. La distribution des flux est dissymétrique, identique aux formes de l'exploration théorique et de l'analyse empirique du réseau ferroviaire, et suggère ainsi une discrétisation géométrique.

8.2.1. **La structure globale du réseau : une prise en compte directe mais partielle de la structure urbaine**

Cette première sous-section vise à étudier les modalités de l'installation des radiales autour de Paris. Répondent-elles à des exigences d'investissement ou de concentration des flux ? Dans un premier temps, nous définissons les villes de premier ordre par les villes de plus de 40 000 habitants en 1860. De cette sorte, le module statique consiste à répartir les flux potentiels par le modèle d'interaction spatiale afin d'alimenter le module d'investissement et mettre à jour les valeurs du réseau créé. Nous testons différents paramètres d'investissement, différents paramètres de concentration[58]. Nous présentons ici les résultats de simulations qui se rapprochent le plus de la structure réelle et une autre situation plus contre-intuitive, pour mieux comprendre l'application des paramètres dans le cadre empirique, à partir de trois situations, dont on résume les paramètres dans la *Figure 8. 10*. Les résultats sont cartographiés dans la *Figure 8. 11*. Nous appelons les différents cas au fil du texte.

[58] Pour davantage de détails sur l'influence des paramètres du modèle, l'exploration théorique du modèle a permis de la qualifier sur des espaces stylisés *PARTIE 1 -6.3.2, p. 224)*. Pour la suite, on qualifiera α comme un paramètre d'investissement et les itérations quand un paramètre de concentration et les figures sont construites selon la même architecture que la *Figure 6. 19 (p. 251)*



Branches divergentes	Kilomètres construits	Population desservie
7	*9 229 km* *Dont 5 300 km sur les axes de Legrand*	*26 %*
14	7 786 km	37 %
7	6 326 km	34 %
4	4 558 km	25 %

Figure 8. 10. Les paramètres du modèle pour caractériser le réseau de premier ordre

La lecture morphologique impose dans tous les cas une structure en étoile, dans laquelle Paris a un poids croissant. Au-delà de la présence du pouvoir central, son poids démographique est important dans la mise en œuvre de la simulation du modèle. Cette structure en étoile ne laisse que très peu de places à des ramifications intermédiaires dans le réseau simulé : quand on prend en compte une structure partielle de la hiérarchie urbaine, ce sont les tracés directs pour rejoindre les principaux centres qui sont préférés. Ce n'est pas une structure en arbre qui émerge.

La lecture du tableau des résultats confronte le nombre de branches divergentes au départ de Paris, le nombre de kilomètres construits à l'heure de la mise en connexité en 1860 et celui créé par la simulation, et le pourcentage de la population desservie. Au départ, avec des paramètres faibles *(Cas a.)* qui favorisent la diffusion de l'innovation ferroviaire sans privilégier une quelconque rentabilité, on observe la présence de 14 branches divergentes au départ de Paris. Il s'agit de maximiser les relations possibles, proches d'une vision saint-simonienne. Les populations desservies promettent d'être nombreuses, plus de 37 % de la population initiale, mais la vraisemblance économique et financière d'un tel réseau peut être facilement remise en cause par une confrontation visuelle. A mesure que l'on augmente le paramètre d'investissement et sans imposer de hiérarchisation *(Cas b.)*, on observe une diminution des axes qui se dessinent, jusqu'à 7 branches divergentes : on répond alors à des exigences économiques, où la rentabilité prend le dessus. Les kilomètres construits chutent pour se rapprocher des 5 300 kilomètres qui constituent l'Etoile de Legrand, alors que 34 % de la population est désormais desservie.

Pour autant, nous ajustons le paramètre de concentration pour encore diminuer les chiffres et se rapprocher de la situation de référence. Le *Cas c* propose de renforcer l'efficacité du réseau en concentrant les flux pour privilégier la canalisation : nous sommes alors dans une situation de rentabilité hiérarchisée. On ne recense plus que 4 branches divergentes, à l'image des points cardinaux. Si le pourcentage de la population desservie est proche de la situation de référence, la concentration a encore une fois fait chuter les kilomètres construits, en dessous de ceux constatés avec l'Etoile de Legrand.

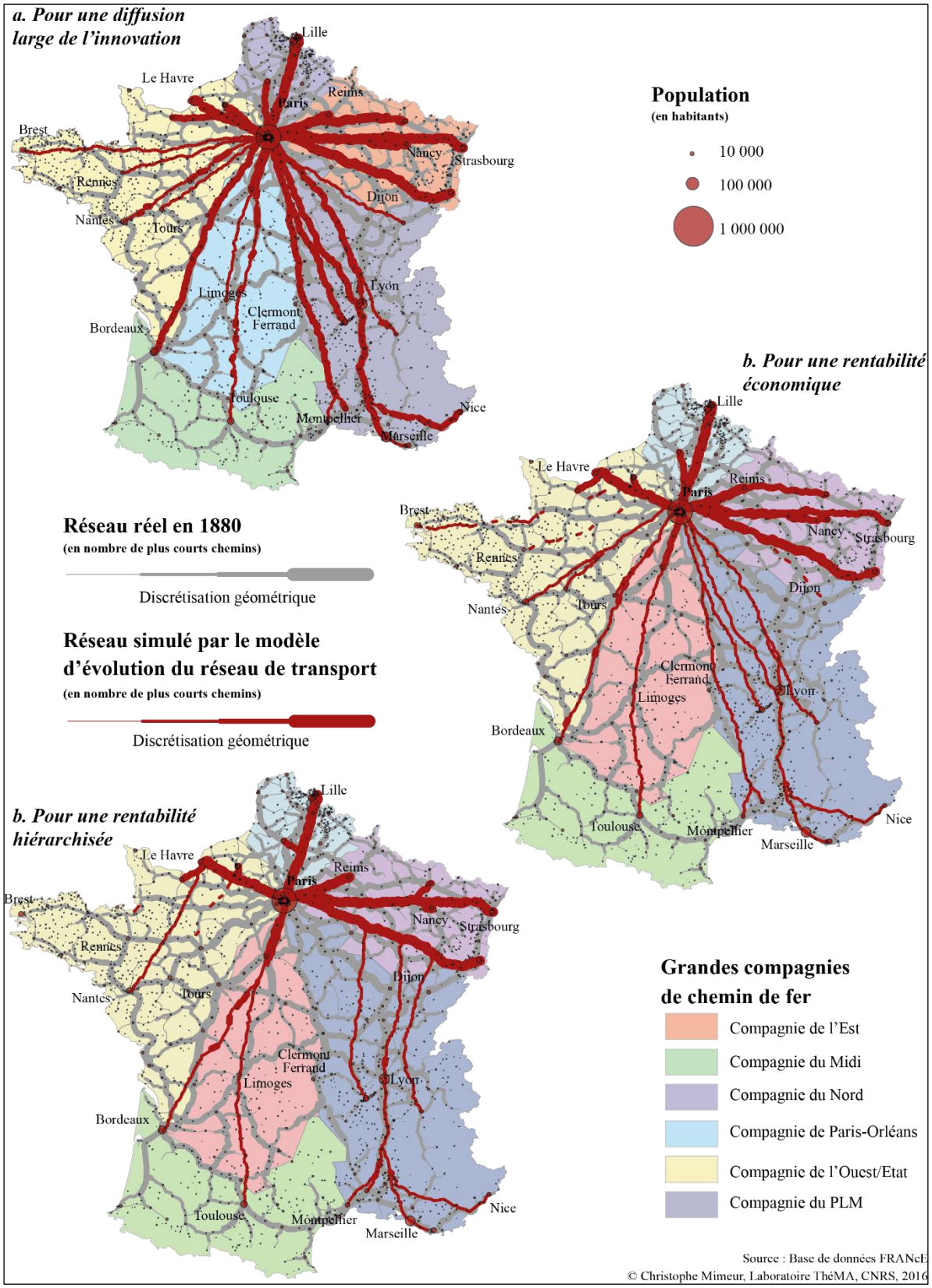


Figure 8. 11. Simulation du réseau pour une structure de 40 000 habitants

Quand les premières études de tracés sortent pour guider l'administration des Ponts et Chaussées, deux doctrines s'opposent. La première est résumée par l'ingénieur en chef Défontaines dès 1835 : *« Les chemins de fer doivent être exclusivement réservés pour les villes de premier ordre, en sorte que leur tracé ne doit nullement se détourner en faveur de localités intermédiaires, dont les transports sont même à éviter comme encombrant la circulation générale du chemin »* (Caron, 1997). Plus nuancé, Minard affirmait dans son *Mémoire sur l'importance du parcours partiel sur les chemins de fer* que le tracé direct ne devait pas *« laisser de côté les grands centres de population agglomérés situés à peu de distance de ce tracé »*. Dès lors, le modèle d'évolution du réseau de transport permet d'évaluer l'étendue de l'application de ces thèses au moment des grandes décisions.

Le *Cas a* propose une multiplicité de lignes qui se dirigent vers les principaux centres, où la hiérarchie mise en place suit un gradient Est/Ouest, bien que cette situation ne soit pas efficiente. Dans le *Cas b*, on identifie les branches principales issues du réseau de Legrand : les branches les plus rentables se dirigent vers Lille, Le Havre et Strasbourg. Les deux suivantes se dirigent vers Bordeaux et Toulouse. Les arbitrages sont en revanche beaucoup plus compliqués pour dessiner des axes vers le sud-est du pays. Le tracé vers Marseille est très direct en passant par Lyon, même si d'autres axes concurrents émergent pour se rendre directement à Montpellier et Grenoble. La densité des villes ne permet pas de privilégier la branche commune même si le paramètre d'investissement opère des sélections, nécessaires à une meilleure efficience économique.

Le *Cas c* propose alors de renforcer l'efficacité du réseau en concentrant les flux pour privilégier la canalisation. Il montre un déplacement de la gravité du réseau vers l'est du pays. Strasbourg et Mulhouse deviennent des centres incontournables, au détriment de la diffusion du réseau jusque dans le sud-est, où le décrochage vers le sud ne se fait qu'à la hauteur de Chaumont pour rejoindre les vallées de la Saône et du Rhône, puis Marseille et Montpellier. Ce réseau simulé propose de nombreux vides, alors que les branches divergentes du réseau de Legrand ont pour objectif de répartir de manière plus équitable les radiales depuis Paris.

Finalement, l'étude parallèle de ces trois situations montre que l'installation des principales branches du réseau de Legrand obéit à des impératifs d'investissement, où l'on favorise la rentabilité économique de l'infrastructure, par un processus de sélection des principaux centres dans une structure déjà élevée de la hiérarchie urbaine. En revanche, l'évolution du paramètre de concentration suggère que les prémices du réseau obéissent à des propriétés hiérarchiques modérées voire même faibles, qui sont une condition pour voir l'apparition la plus proche des branches divergentes effectivement constatées par le réseau de Legrand. L'étude comparée de ces trois réseaux simulés permet en outre de mieux voir où est-ce que les conclusions qui viennent d'être tirées s'appliquent, et dans quels cas elles entrent en contradiction avec le réseau effectivement créé. C'est pour cela que nous nous appuyons sur le second cas *(Cas b)* pour l'évaluation.

La comparaison est valide dans la moitié ouest du pays, en rejoignant Bordeaux par Tours ou encore Toulouse par Limoges. Un autre axe vers Nantes par Le Mans et à proximité d'Angers est aussi identifiable. L'espacement des villes dans le haut de la structure urbaine favorise l'émergence de ces axes structurants.

Il existe en revanche plusieurs niveaux de disparités à l'est, que nous imputons à la plus forte densité de noyaux de population dans cette moitié du pays. Les plus prégnants résident dans la concurrence d'axes, résultant de la difficulté du modèle à privilégier la branche commune de manière vraisemblable. Plusieurs axes se dessinent pour rejoindre la Méditerranée : l'arbitrage peine à être effectif entre Marseille, Montpellier, ou encore entre Lyon et Saint-Etienne. On voit par ailleurs que c'est bien le processus de concentration qui tend à arbitrer ces hésitations dans une certaine mesure, canalisant les flux dans la vallée du Rhône vers Marseille puis Montpellier *(Cas c.)*, même si le modèle ne reproduit pas un axe à part entière se dirigeant vers le sud-est. Enfin, cette dernière situation démontre comment le processus de concentration des flux est celui qui favorise la création de branches divergentes dans d'autres centres que Paris, comme c'est le cas pour Orléans, Nancy ou Montélimar.

Avec un niveau de la structure urbaine supérieur à 40 000 habitants, l'objectif était d'étudier dans quelle mesure le réseau construit participe aux phénomènes cinétiques et adaptatifs de la théorie des réseaux. Cette prise en compte partielle de la structure urbaine privilégie la structure en étoile autour de Paris. La dimension cinétique de ce premier réseau est surtout le fait de la volonté de rejoindre le plus directement possible les villes de premier ordre, parfois au détriment d'un processus de hiérarchisation. Le modèle éprouve toutefois des difficultés à arbitrer quel axe privilégier en présence d'une plus forte densité de villes. Ainsi, nous posons l'hypothèse que ces arbitrages ont surtout été le fait d'une dimension adaptative dans laquelle les décisions d'aménagement paraissent être le facteur le plus probable. C'est notamment le cas du choix à l'est de Paris, dont on a vu que les hésitations ont longtemps perduré entre Nancy, Metz, Reims et Strasbourg. C'est aussi le cas du tracé entre Paris, Lyon et Marseille, dont nous aurons à revenir plus tard dans nos propos. Alors, malgré la présence de grands centres de population à l'est, c'est là que les aménageurs semblent avoir le plus à intervenir pour faire des choix, dont nous posons l'hypothèse qu'ils s'appuient sur un niveau inférieur de la structure urbaine.

Par ailleurs, nous avons vu que le processus de concentration semble décisif dans l'apparition de ramifications dans le réseau, dont nous avons constaté l'importance dans la section précédente. Pour investiguer leur croissance dans le cadre de la simulation, de même que pour affiner les arbitrages entre dimensions cinétique et adaptative, nous poursuivons nos investigations vers des niveaux inférieurs en appliquant les mêmes paramètres.

8.2.2. **L'émergence de carrefours intermédiaires pour desservir la population**

L'objectif de cette sous-section est d'étudier le phénomène d'émergence d'une hiérarchie dans le réseau, à partir d'axes structurants et de ramifications secondaires. Cela requiert la prise en compte d'une structure urbaine préexistante plus large. Les deux premiers cas s'appuient sur les villes de plus de 20 000 habitants et le troisième sur les villes de plus de 15 000 habitants. La *Figure 8. 12* montre l'évolution des paramètres pour appréhender ce phénomène, à partir du réseau principal stabilisé en 1880 comme situation de référence. Cette fois, le critère morphologique que nous retenons est le nombre de carrefours intermédiaires, dont le nombre de branches divergentes sur un point est supérieur à 3. On reporte ensuite le nombre de kilomètres construits et la part de la population desservie. La *Figure 8. 13* confronte les réseaux réels et simulés.

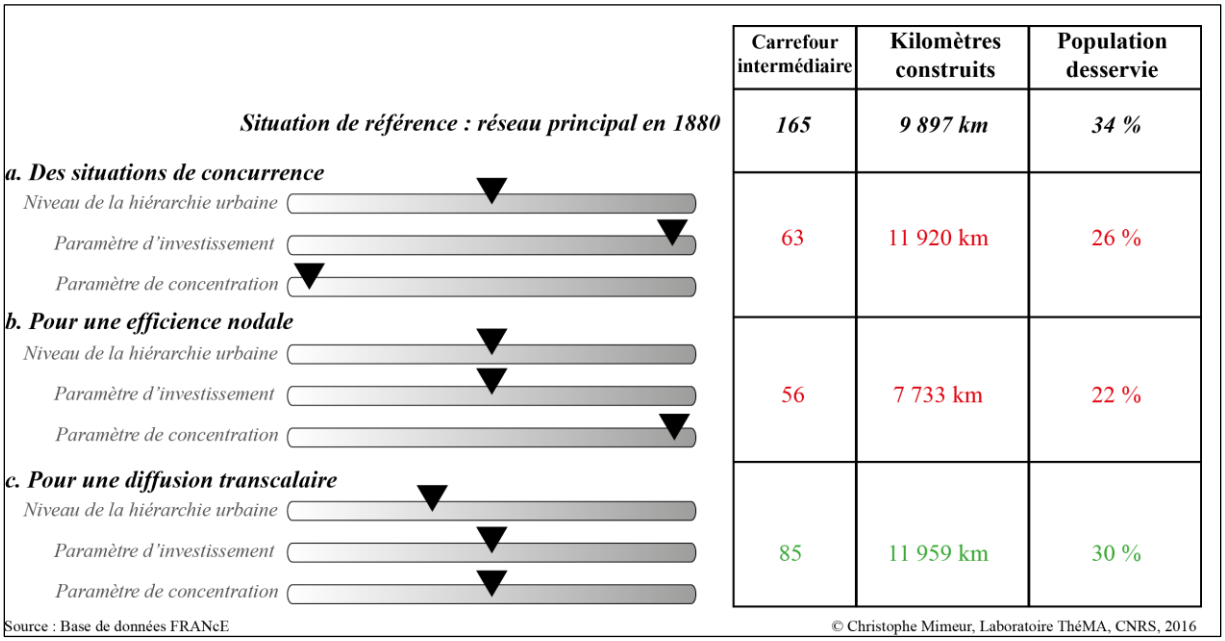


Figure 8. 12. Les paramètres du modèle pour caractériser la nodalité

Dans un premier temps, nous reprenons les paramètres qui ont conduit à l'apparition des 7 branches divergentes depuis Paris. Avec une forte sélection des points à rejoindre de manière prioritaire, nous obtenons un réseau dont la longueur est supérieure à la situation de référence, mais dont la part de la population desservie est peu efficiente *(Cas a)*. Par ailleurs, cette simulation privilégie encore les cheminements directs. La structure en étoile autour de 7 branches est remise en cause : la rentabilité économique semble privilégier le passage par Paris, même si le modèle parvient à créer de premiers nœuds intermédiaires. Il semble bien que le processus de concentration pourrait participer à l'arbitrage de situations concurrentes et à l'émergence de ramifications dans le réseau, pour favoriser sa diffusion auprès d'une population plus large.

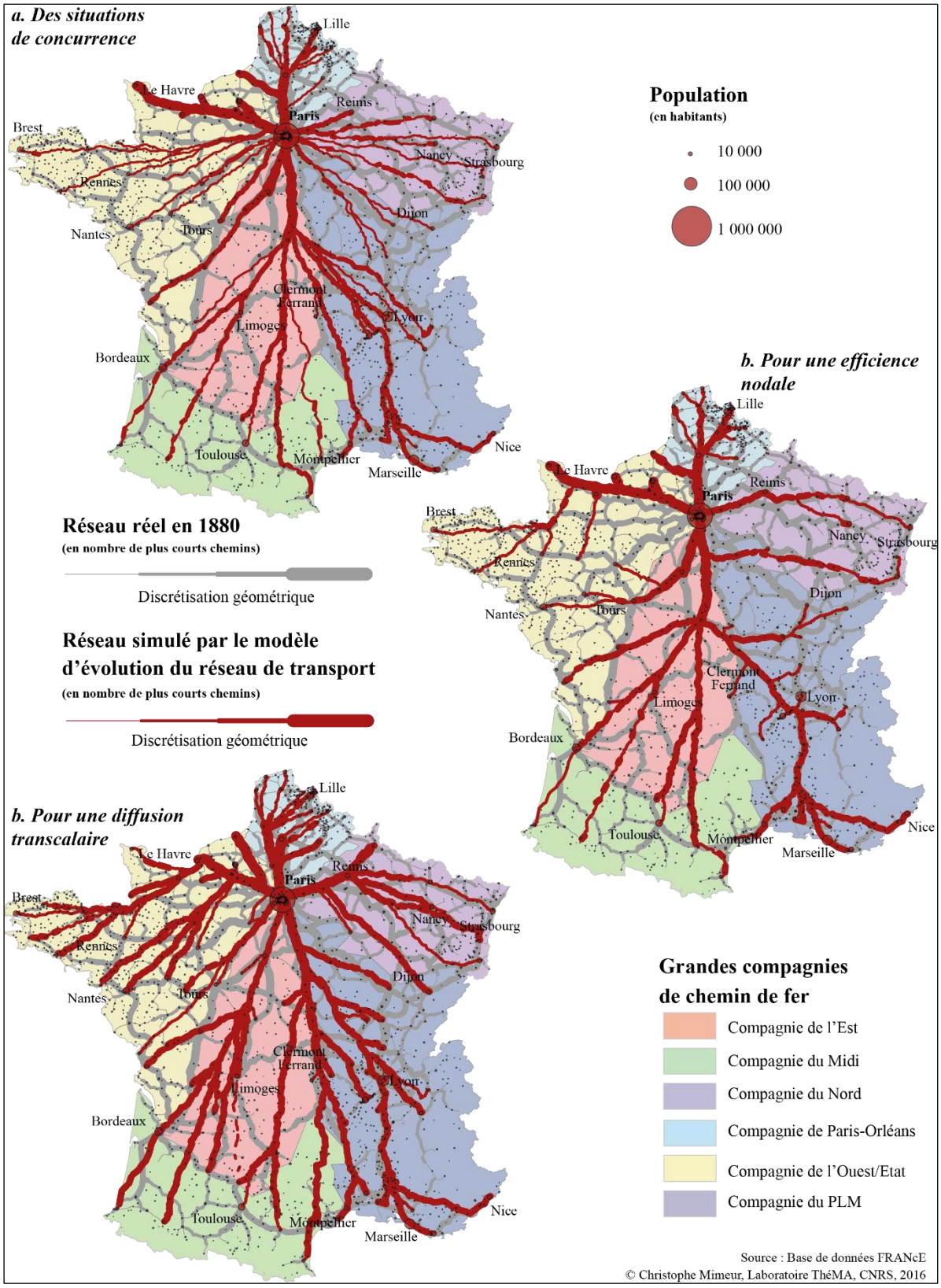


Figure 8. 13. Simulation du réseau pour une nodalité dans le réseau

Le *Cas b* propose un double phénomène : il accentue le processus de concentration pour privilégier la branche commune, tout en abaissant la force du paramètre d'investissement pour permettre une plus large diffusion du réseau. La simulation parvient alors à reproduire la structure stellaire autour de Paris : elle est le fruit d'une forte sélection quand la situation est partielle. Elle est davantage le fruit d'une concentration des flux quand la situation initiale est plus large, qui aboutit toutefois ici à seulement 4 branches. De cette sorte, nous posons l'hypothèse que la simulation arrive à opérer un arbitrage dans les nombreux axes concurrents. Donc, le réseau construit est peu dense, se rapprochant de la situation de référence. Plus étonnement, il ne dessert plus que 22 % de la population initiale. Si le nombre de carrefours intermédiaires est faible, 20 % d'entre eux sont des carrefours complexes à quatre branches ou plus. A ce stade, la dimension cinétique pour un seuil de 20 000 habitants est trop importante pour reproduire le réseau d'intérêt national.

Pour contrebalancer ces faibles performances, nous envisageons une troisième simulation *(Cas c)*. Elle rétablit 6 branches divergentes, plus proche de la situation de référence : définitivement, la répartition des branches autour de Paris est le fruit d'une concentration modérée, afin de favoriser un espacement le plus égal possible entre les branches. Par ailleurs, pour favoriser la diffusion tout en privilégiant l'apparition de carrefours intermédiaires, nous combinons la baisse de la structure urbaine préexistante à 15 000 habitants tout en diminuant le paramètre de concentration. Alors, on augmente les trois indicateurs qui nous rapprochent du réseau de référence.

A ce stade, l'évolution des paramètres montre que la mise en place progressive d'un réseau hiérarchisé résulte bien cette fois de la combinaison des deux paramètres du modèle. La diffusion du réseau secondaire nécessite des différentiels d'investissement modérés pour suggérer la construction de voies moins rentables. Par ailleurs, la dimension cinétique prend corps dans le paramètre de concentration pour privilégier la branche commune, sans l'exagérer toutefois pour favoriser la desserte large des populations. Il s'agit désormais d'étudier dans quelle mesure ces paramètres influencent la géographie du réseau en le rapprochant du réseau effectivement construit. La confrontation s'appuie sur les deux dernières situations décrites dans la *Figure 8. 13*, en s'appuyant également sur des éléments de contexte pour mieux comprendre la position des carrefours intermédiaires.

A la différence de la sous-section précédente, on constate un différentiel entre nord et sud du pays. Dans les deux cas, le modèle arrive à reproduire les principaux axes vers Lille, Strasbourg et Cherbourg. Il parvient même à faire émerger des ramifications vers Compiègne, Boulogne et Arras pour Lille, Nancy pour Strasbourg et Le Havre pour Cherbourg.

Plus on s'éloigne de Paris pour rejoindre le Sud, plus la situation peine à être stabilisée. En effet, on remarque qu'à l'est, la branche directe vers Lyon et Marseille est toujours privilégiée, aux dépens d'un passage par Dijon, tandis qu'une féroce concurrence l'oppose à un axe qui rejoint Montpellier quand la concentration reste modérée : la structure du réseau réel peine aussi à opérer un choix. Dès lors, il semble que le dessin du réseau principal, mais aussi de branches plus secondaires, répond à des enjeux d'investissement et de croissance modérée *(Cas c)*. Pour autant, de premiers carrefours intermédiaires émergent pour couvrir davantage le territoire. Certaines suggèrent de rejoindre la Vallée de la Saône tandis que d'autres permettent davantage de dessertes

dans les régions Lorraine et Alsace. Aussi, la desserte de l'arc méditerranéen passe par Montélimar. Il faut alors souligner un écart entre le réseau simulé et le réseau réel : Avignon est choisi par l'influence de l'ingénieur Paulin Talabot, en tant qu'ingénieur et voisin du Gard. Il s'était fait remarqué dans sa volonté de détruire les remparts de la ville. Le projet fut abandonné au profit d'ouvrages d'art financés localement. Le choix d'Avignon se justifie également dans sa capacité à lier les réseaux venant de Lyon, de Marseille-Nice et de Montpellier : ce n'est pas son dynamisme local qui justifie ici le choix de ce carrefour. En revanche, la possibilité de transporter les houilles par une compagnie à la tête de laquelle se trouve le même Paulin Talabot semble un élément explicatif fort du choix d'Avignon et d'Arles dans le prolongement vers Marseille. A l'est, le réseau simulé semble finalement plus proche de la situation de référence.

A l'inverse, la situation à l'ouest est plus confuse quand on ne privilégie pas une concentration exacerbée. Elle reste valide quand on cherche à rejoindre le sud-ouest avec une branche commune vers Bordeaux, qui fait d'Orléans, Châteauroux et Limoges des passages obligés. Ainsi, on pose l'hypothèse que le besoin d'une branche commune se fait davantage ressentir à mesure que l'on s'éloigne du centre de gravité du réseau *(Cas b)*. Pour autant, on voit que le modèle opère un choix entre l'axe Paris-Tours-Poitiers-Bordeaux et l'axe Paris-Orléans-Limoges-Toulouse en proposant un tracé intermédiaire dont les ramifications permettent de rejoindre ces centres secondaires *(Cas c)*. Nous avons vu que la Compagnie du Paris-Orléans a fait le choix du premier aux dépens du second, alors que la seconde simulation semble donner l'avantage au premier *(Cas b)* : on pose l'hypothèse de facteurs exogènes qui participent à ce choix, parmi lesquels les influences de la Compagnie du Paris-Orléans et de la Compagnie de l'Etat pour rejoindre le littoral semblent importantes. Finalement, l'axe vers Toulouse sera l'une des dernières radiales à être achevée.

La situation paraît encore plus confuse quand on compare le réseau réel et le réseau simulé en Bretagne. Les deux structures urbaines prises comme situation initiale ne permettent pas d'isoler un axe Paris-Rennes, dont on sait qu'il a été ajouté plusieurs années après les premières décisions de l'Etoile de Legrand. Alors, le cheminement de Brest se fait par la Normandie, tandis que la desserte de Nantes est opérée depuis Tours.

Finalement, nous retenons de ces trois dernières simulations la capacité du modèle à hiérarchiser le réseau : c'est par le processus de concentration que le modèle parvient à créer des centralités dans le réseau en séparant branches principales et ramifications secondaires. La structure globale du réseau répond bien à une structure intermédiaire à l'Est du pays, qui parvient à identifier des centralités et des possibilités de circuler en province. La prise en compte d'une structure plus précise semble requise pour étudier le phénomène à l'œuvre dans l'ouest du pays.

On peut par ailleurs identifier des possibilités de rapprochement entre certains terminus définis par la simulation : des potentialités de circuité dans le réseau commencent ainsi à être envisageables, sans que le modèle ne soit pas capable de les reproduire entièrement. On pose l'hypothèse que la proximité de ces points peut être un facteur exogène qui participe à la décision de la construction d'une transversale, nous l'envisagerons dans la prochaine sous-section.

8.2.3. **De potentielles transversales dans un arbre hiérarchisé**

Désormais, nous avons deux critères morphologiques pour évaluer la performance d'une simulation : le nombre de branches divergentes et le nombre de carrefours intermédiaires. Pour envisager la circuité dans le réseau, il s'agit d'évaluer dans quelle mesure on peut envisager des potentielles transversales dans le réseau. Pour cela, on s'appuie sur une structure fine de la situation initiale, de 5 000 à 2 500 habitants. Comme précédemment, nous repartons des derniers paramètres les plus performants dans les simulations précédentes *(Figure 8. 14)*. Les deux simulations sont reproduites dans la *Figure 8. 15*.

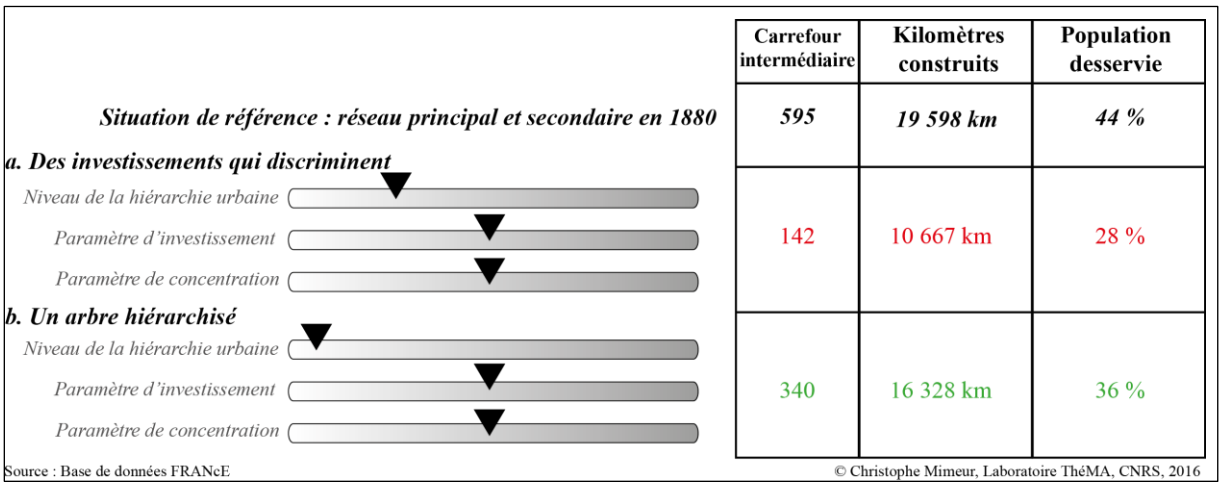


Figure 8. 14. Les paramètres du modèle pour caractériser le réseau secondaire

Cette fois, nous prenons les réseaux principal et secondaire comme situation de référence en 1880. L'application des mêmes paramètres ne permet de couvrir que 28 % de la population *(Cas a)*. Même si la structure urbaine initiale est plus fine, les kilomètres construits demeurent stables, alors que les ramifications ont été favorisées. Pourtant, on est encore loin de la situation de référence.

Alors que les paramètres de construction nous sont parus efficients pour une part importante du territoire, nous misons plutôt sur le niveau de l'initialisation pour faire converger branches communes, ramifications secondaires et possibilités de circuler entre les provinces. Avec une situation initiale qui prend en compte toutes les communes de plus de 2 500 habitants, le nombre de kilomètres est considérablement augmenté, et se rapproche de la situation de référence. Il en de même pour le critère des carrefours intermédiaires. Le pourcentage de la population desservie est lui aussi renforcé pour atteindre 36 %, à 8 points de la situation de 1880. Dès lors, nous concluons que la mise en place du réseau d'intérêt national est le fruit d'une modération en termes d'investissement et de concentration : cette modération est la condition de la diffusion du rail, parfois au détriment de la rentabilité, qui est toutefois optimisée en privilégiant la branche commune. De cette sorte, la simulation aboutit à une structure en arbre hiérarchisé *(Cas b)*. Il s'agit de confronter les différentes simulations à la morphologie du réseau réel.

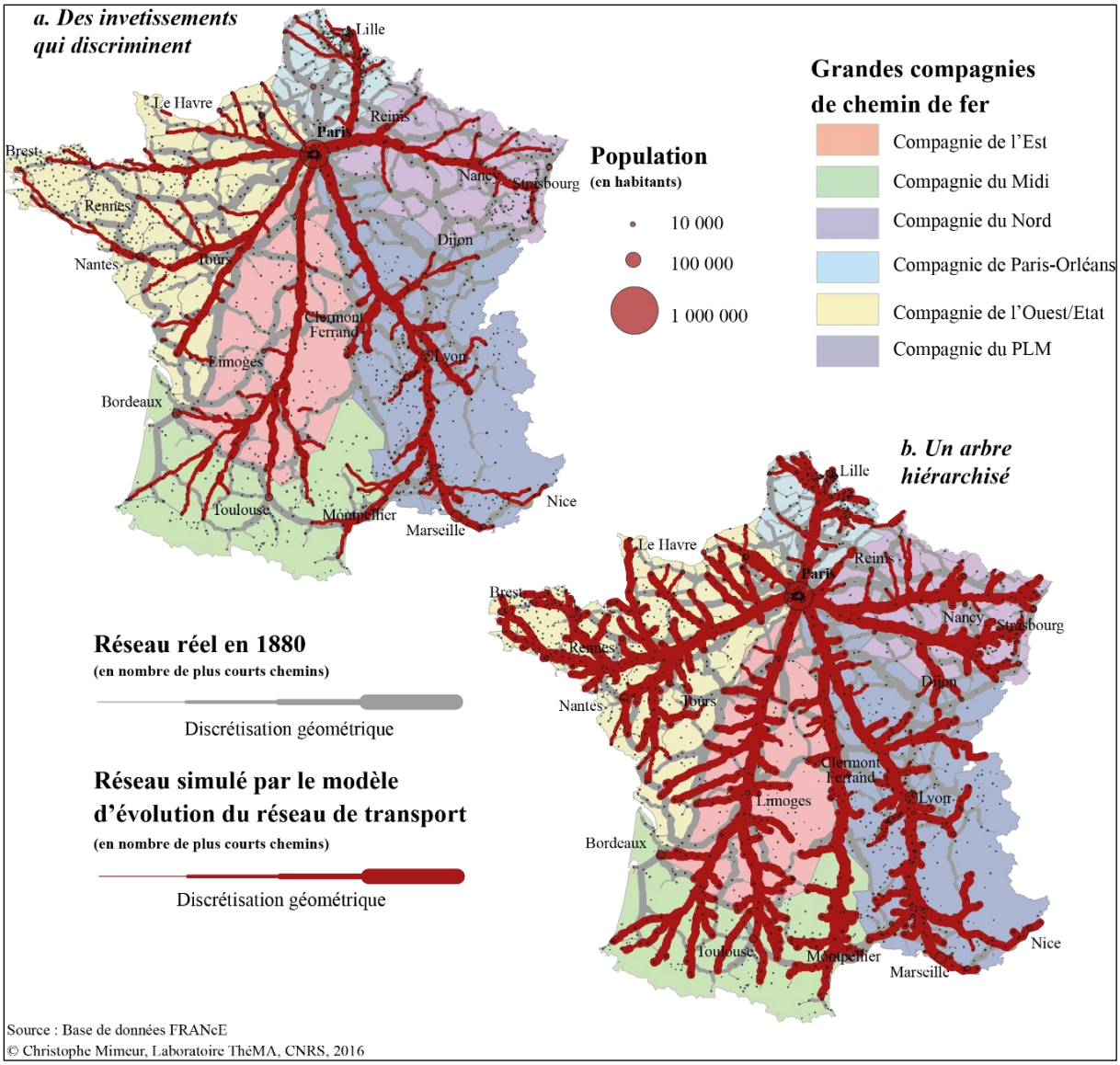


Figure 8. 15. Simulation du réseau pour une structure plus complète

Malgré la déficience des principales caractéristiques du réseau simulé, le *Cas a* nous permet de confirmer la stabilisation des principales radiales autour de Paris, au prix des paramètres choisis et quel que soit le niveau initial pour simuler la croissance du réseau. A l'est, la situation paraît elle aussi semblable avec un renforcement des performances nodales que l'on peut davantage identifier. La seconde simulation propose des critères plus proches de la situation de référence. Une structure plus fine fait émerger à l'est des axes secondaires *(Cas b)*, mettant à profit les performances du modèle à créer des carrefours intermédiaires, afin que l'infrastructure se propage dans la Champagne, la Lorraine, le nord de la Bourgogne et de la Franche-Comté.

Dans l'autre moitié du pays, la structure prise en compte par le modèle *(Cas a)* permet d'intégrer la densité importante de communes peuplées à l'Ouest du pays. Les littoraux nord et sud sont traversés par le réseau simulé et leur proximité laisse imaginer les possibilités de se joindre pour mieux circuler. La structure initiale la plus fine *(Cas b)* permet en outre de prendre davantage en compte les foyers de population entre Paris et la pointe bretonne, qui suggèrent un cheminement par Le Mans et Rennes.

Au sud-ouest, le modèle transcrit la concurrence qu'il existe effectivement pour rejoindre la frontière espagnole, par Tours et Poitiers dans un cas, par Orléans et Limoges dans le second. La situation est éclaircie à un niveau encore plus fin *(Cas b)*, qui donne décidément l'avantage au réseau qui se rapproche davantage du littoral, venant renforcer notre hypothèse sur l'existence de facteurs exogènes à cette décision.

La structure du réseau simulé ne propose toutefois pas de capacités circulatoires. Pourtant, la structure en arbre hiérarchisé qui ressort de la dernière simulation *(Cas b)* vient appuyer l'hypothèse énoncée dans la sous-section précédente, selon laquelle les extrémités de cette structure peuvent participer à la réflexion sur les possibilités de tracer des infrastructures transversales. Pour cela, nous proposons une mesure de la distance qui sépare chacun de ces points. Cette démarche exploratoire nous conduit à définir un critère de distance, par laquelle nous soumettons l'hypothèse qu'un investissement permettrait de faire émerger des performances circulatoires dans le réseau. Par une démarche progressive pour définir ce critère, nous présentons dans la *Figure 8. 16* les résultats avec un critère de distance de 50 kilomètres.

Parmi tous les segments possibles, nous avons choisi de supprimer les arêtes qui se croisent pour plus de clarté dans l'interprétation. Il est intéressant de souligner que parmi les potentielles transversales tracées par des traitements géomatiques, certains d'entre eux correspondent avec les liaisons transversales constatées dans le réseau réel : il s'agit de faire le lien entre les différentes radiales du réseau, de même qu'entre les différentes compagnies.

C'est notamment le cas entre les compagnies du Nord et de l'Est où Laon était identifié comme un centre à desservir et se retrouve au cœur d'un axe Amiens-Reims quand on lui joint de potentielles transversales. Cet exemple a vocation à se répéter sur l'ensemble du territoire. Il est aussi étonnant de voir comment ces possibles liaisons participent à la mise en cohérence de réseaux que l'on peut qualifier de plus régionaux. C'est le cas de l'axe Nord/Sud en Alsace que le modèle n'arrive pas à reproduire. Cet axe tend à se prolonger pour lier Saône et Rhin.

Quand on descend plus au sud, on remarque aussi la présence de sous-systèmes qui font sens avec le réseau réel de 1880, avec l'identification de la vallée de la Saône d'une part, le cheminement par la Bresse d'autre part. La forte densité de foyers de population favorise aussi la mise en connexité de ces centres qui ne paraissent pas assez rentables à l'échelle nationale : des liaisons entre Auvergne et Rhône-Alpes se dessinent de manière multiple pour desservir le bassin de Saint-Etienne. Dans le même ordre d'idée, les potentielles transversales identifient une ceinture qui relie la vallée de la Loire au littoral Atlantique, en suggérant des lignes effectivement existantes, par Nevers, Vierzon, puis Tours. De semblables suggestions sont présentes pour rejoindre Angers au Mans par exemple.

Les performances circulatoires sont améliorées quand on prend en compte un critère de proximité entre les centres à desservir d'après la dynamique du modèle. En Bretagne, ces avancées nous permettent d'envisager un collier littoral qui se propage jusque sur les côtes normandes jusqu'au Havre, qui rapproche alors la simulation de la réalité. Ces conclusions ont tendance à se vérifier également sur le littoral charentais et vendéen, de même que sur l'arc méditerranéen. Enfin, cette analyse permet d'identifier les rapprochements très étroits pour dessiner la transversale entre Atlantique et Méditerranée.

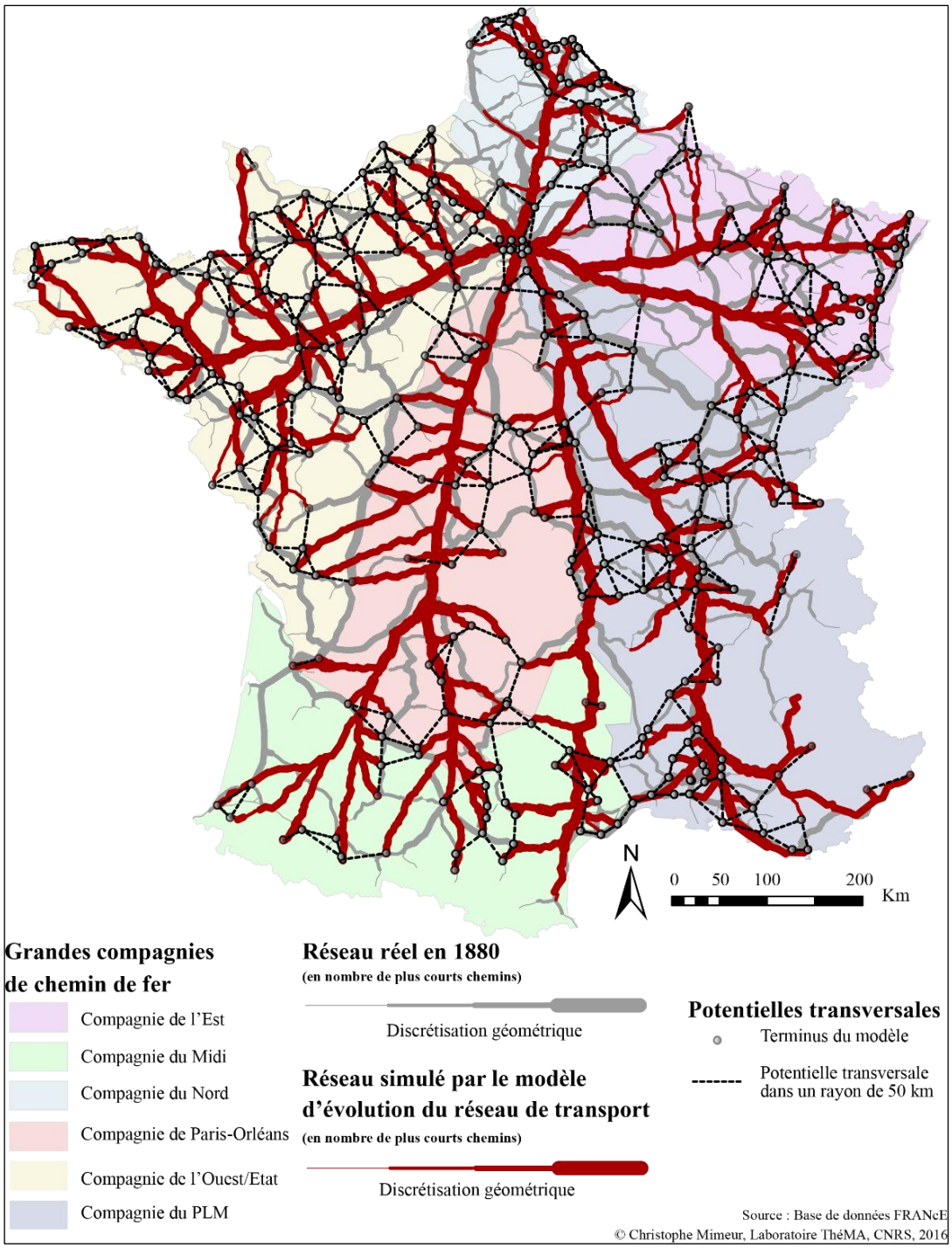


Figure 8. 16. De potentielles transversales dans le réseau national

Finalement, cette exploration du réseau d'intérêt national par le modèle d'évolution d'un réseau de transport nous a permis de montrer comment la structure démographique préexistante a été décisive dans le tracé des principaux axes de la première phase de croissance du réseau ferroviaire français. Ainsi, la structure stellaire autour de Paris répond à une structure initiale des villes de plus de 40 000 habitants, ainsi qu'à des objectifs de rentabilité économique, dont la vision hiérarchique n'était pas requise à ce niveau des branches directes.

Par ailleurs, la simulation du réseau à d'autres seuils de population montre comment la structure stellaire doit s'accommoder de choix d'investissement dont la discrimination est modérée, afin de conserver un écartement homogène des branches dessinant l'étoile autour de Paris. Ce choix modéré participe aussi à la diffusion du réseau, pour desservir davantage de population. Pour optimiser ce processus, le paramètre de concentration favorise la branche commune et une structure en arbre. Le modèle est alors capable de créer des ramifications qui illustrent le phénomène de hiérarchisation par la création de carrefours intermédiaires. A l'est, la stabilisation du réseau national est le fruit d'une prise en compte des villes de plus de 15 000 habitants. A l'ouest, les fortes densités de communes peuplées obligent à une prise en compte plus complète de la hiérarchie urbaine, sans que les simulations ne parviennent à dessiner des transversales pour rejoindre ces centres intermédiaires. Pour cela, la création de potentielles transversales à partir de distances entre les arêtes du réseau simulé participe à l'appréhension de la circuité dans le réseau, qui semble obéir à d'autres paramètres que ceux de ce modèle. Pour autant, nous souhaitons tester ces paramètres à des échelles plus locales : parviennent-ils alors à reproduire un réseau selon une structure encore plus fine ? Les décalages observés permettent-ils d'envisager des circulations dans le réseau ? Dans quelle mesure peut-on intégrer des facteurs exogènes dans le modèle ?

8.3. Enrichir les comparaisons par des explorations locales

Parmi les trois types de stratégies qui influencent le tracé des voies ferroviaires, nous avons vu comment l'efficience économique à une macro-échelle a présidé le dessin des principaux rayons partant de Paris puis comment les niveaux inférieurs de la hiérarchie ont influencé certains tracés. D'autres ont été détournés de cheminements optimaux par des volontés politiques ou une marque intense d'aménagement du territoire. Cette section vise à explorer la combinaison de ces trois aspects à des échelons inférieurs : au-delà des axes structurants, l'objectif est d'évaluer comment le modèle d'évolution d'un réseau de transport est capable d'envisager la seconde phase de la croissance du réseau ferroviaire français et la circuité dans le réseau.

Nous explorons ici plus particulièrement deux zones d'études : l'aire d'influence de la Compagnie des Chemins de fer du Nord, dont la stratégie est multiscalaire ; l'aire d'influence du Paris-Lyon-Méditerranée, qui voit la desserte du massif alpin et des actes d'aménagements décisifs dans le dessin du réseau. Ces deux cas permettent d'explorer comment le modèle d'évolution du réseau de transport peut être sophistiqué.

8.3.1. **Le réseau du Nord : entre frontière, littoral et bassin industriel**

Cette première sous-section a pour objectif de transposer les paramètres qui ont conduit à la simulation du réseau national à une échelle plus locale. Dans quelle mesure la situation initiale participe-t-elle à l'établissement d'une structure globale du réseau ? Le choix de la Compagnie des Chemins de Fer du Nord n'est pas un hasard : ce sous-réseau n'est pas un réseau en étoile, même si on suspecte Paris de rester le centre incontournable pour le dessin général du réseau. Il s'agit alors

de tester son rôle dans l'évolution des paramètres quand le centre principal n'est pas au cœur du territoire. Par ailleurs, il semble que l'architecture globale du réseau réel répond à plusieurs enjeux, résumés en deux phrases dans l'ouvrage de François Caron : *« les lignes du Nord formaient dorénavant un réseau dont l'extension répondait aux logiques d'ensemble du nouveau système : intensifier la desserte dans la zone déjà contrôlée, lutter contre les concurrences dans les zones frontières. Mais un troisième facteur était intervenu : il s'agissait de fournir au trafic de la ligne principale, déjà au bord de la saturation, des cheminements alternatifs »* (Caron, 1997).

Pour tester l'adaptation de l'infrastructure à ces trois types de priorité, nous simulons la croissance du réseau sur le territoire de la Compagnie du Nord. Nous adoptons ici la même grille de lecture qu'auparavant, en faisant varier les paramètres à trois niveaux différents de la hiérarchie urbaine, qui prennent en compte de plus en plus de communes à l'initialisation *(Figure 8. 17)*. Comme la capacité du logiciel permet de s'appuyer sur une large partie de la hiérarchie urbaine, nous choisissons le réseau complet de 1930 comme situation de référence. La confrontation visuelle entre réseau réel et réseau simulé est permise par la *Figure 8. 18*.

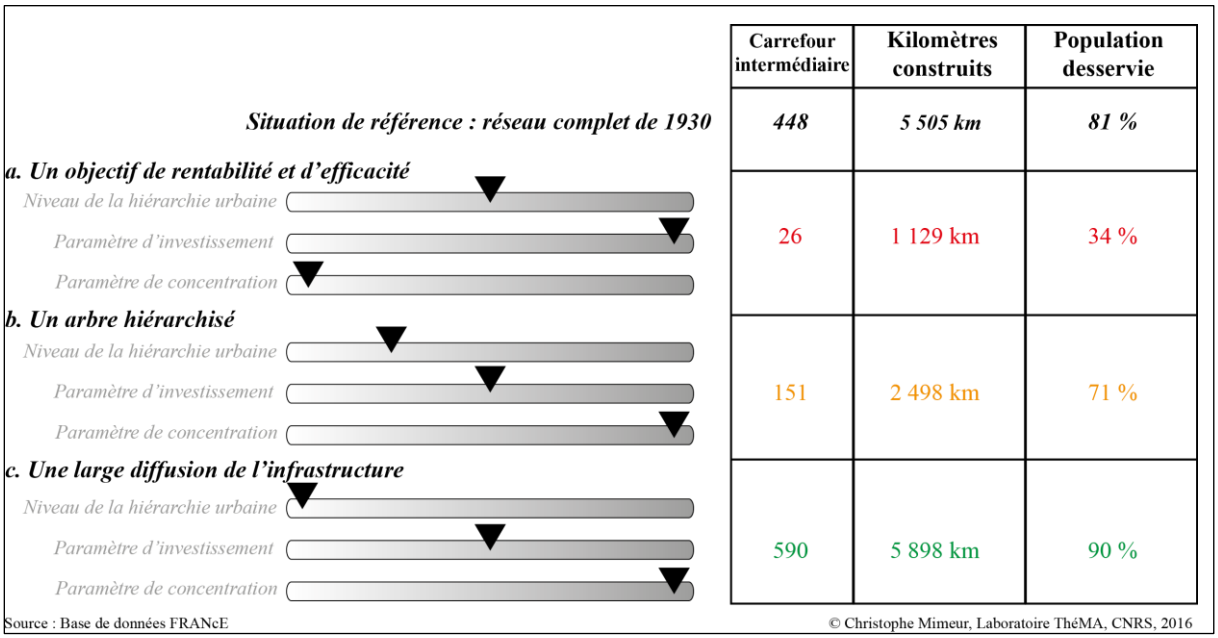


Figure 8. 17. Les paramètres du modèle pour simuler l'infrastructure sur le territoire de la Compagnie des Chemins de Fer du Nord

La première simulation a pour objectif d'étudier quels principaux centres le modèle choisit de desservir quand on choisit des paramètres qui favorisent exclusivement la rentabilité et l'investissement, au détriment de la diffusion du réseau *(Cas a)*. De manière attendue, le réseau ainsi simulé ne compte que 1 129 kilomètres et ne couvrent que 34 % de la population. Les performances nodales de ce réseau sont faibles : à l'image du réseau national, ces paramètres favorisent la branche directe et ne se détourne pas de son principal objectif. La confrontation visuelle permet en revanche de mieux saisir les priorités du modèle dans le développement du réseau. Il propose trois branches principales au départ de Paris qui correspondent aux trois axes établis en 1860, attestant d'une naissance du réseau qui ne répond qu'à des objectifs de rentabilité sans privilégier la diffusion du réseau. Le premier axe cherche à rejoindre le littoral, sans qu'un

arbitrage réussisse à opérer au-delà d'Abbeville, tant les tailles de Boulogne, Calais et Dunkerque sont comparables. Dans la lignée des hypothèses sur les potentielles transversales, on peut alors envisager une transversale qui permet les dessertes successives de Abbeville, Boulogne puis Calais et Dunkerque.

Le second axe est celui qui permet de se diriger vers Lille par Cambrai, qui évite toutefois Amiens, qui est pourtant l'axe historique entre Paris et Lille et qui permet la desserte d'Arras et Valenciennes. On pose l'hypothèse que la forte densité de communes très peuplées – plus de 10 000 habitants dans le *Cas a* – favorise cet axe plus à l'Ouest que le réseau a effectivement créé. Alors, un troisième axe se dessine vers Laon. L'adoption de ces paramètres ne permet pas de privilégier la branche commune. De cette sorte, nous misons sur l'adoption de paramètres modérés, à l'image de ceux utilisés dans la situation nationale, en abaissant la situation initiale de la hiérarchie urbaine.

Le *Cas b* favorise la diffusion de l'infrastructure avec un paramètre d'investissement modéré et la hiérarchisation du réseau pour privilégier la branche commune. Le nombre de kilomètres augmente alors quand on prend en compte les communes de plus de 2 500 habitants. Près de 71 % de la population est desservie et la stratégie de concentration permet l'apparition de 151 carrefours. Les performances de ce réseau restent inférieures au réseau réel. Pourtant, on voit que le réseau arrive à concentrer les flux au départ de Paris pour un premier carrefour important près de Compiègne, alors qu'il se situe à Creil dans la situation de référence. On peut dès lors postuler des effets exogènes dans le choix de cette ville, et on peut soumettre l'hypothèse que la situation de carrefour à Creil peut lui procurer un effet « réseau » dans la mesure où la situation initiale ne privilégie pas ce choix.

A partir de ce premier nœud, l'arbre hiérarchisé identifie les deux priorités de cette branche : rejoindre les fortes densités à l'est de la région, autour de Maubeuge, mais aussi desservir la région lilloise, en assurant un passage par Arras. Depuis cette branche, la simulation se propage jusqu'au le littoral à la frontière belge. Si le modèle met bien en évidence ces deux axes, il montre aussi un axe concurrent qui permet de se diriger vers le littoral en passant par Amiens, devant rejoindre ensuite Boulogne et Calais. A ce stade de l'initialisation, on remarque qu'à un niveau plus local, des concurrences subsistent avec un fort paramètre de concentration. Dans le réseau réel, les arbitrages ont été effectués de la façon suivante : le partage du flux est opéré au sud-est et permet l'émergence de deux axes, vers Amiens et vers Saint-Quentin. A un niveau inférieur, une autre ramification permet de rejoindre directement le littoral. Ainsi, deux hypothèses se posent : soit le chemin par le littoral est un choix délibéré d'aménagement, soit il répond à une structure préexistante plus maillée.

Pour cela, la troisième simulation permet de prendre en compte les communes de plus de 800 habitants, en conservant les mêmes paramètres qui régissent une diffusion de l'infrastructure par un paramètre d'investissement modéré et une forte hiérarchisation avec un paramètre de concentration élevé. Le réseau obtenu est le plus proche de la situation de référence, bien qu'il surestime le nombre de kilomètres construits, de même que les nœuds dans le réseau et la population desservie. Pour autant, il permet d'isoler les branches communes qui sont sélectionnées. Les branches au départ de Paris répondent davantage à la réalité, avec une branche alternative qui se dirige vers Laon à l'extrême est du territoire d'étude. En revanche, il montre bien l'existence d'une

branche commune pour ensuite se séparer vers Saint-Quentin d'une part et Amiens d'autre part. Cette concentration renforcée participe à la plus large diffusion du réseau jusque sur le littoral de la Manche vers Abbeville. Toutefois, il ne parvient pas à rejoindre les principaux ports au nord d'Abbeville.

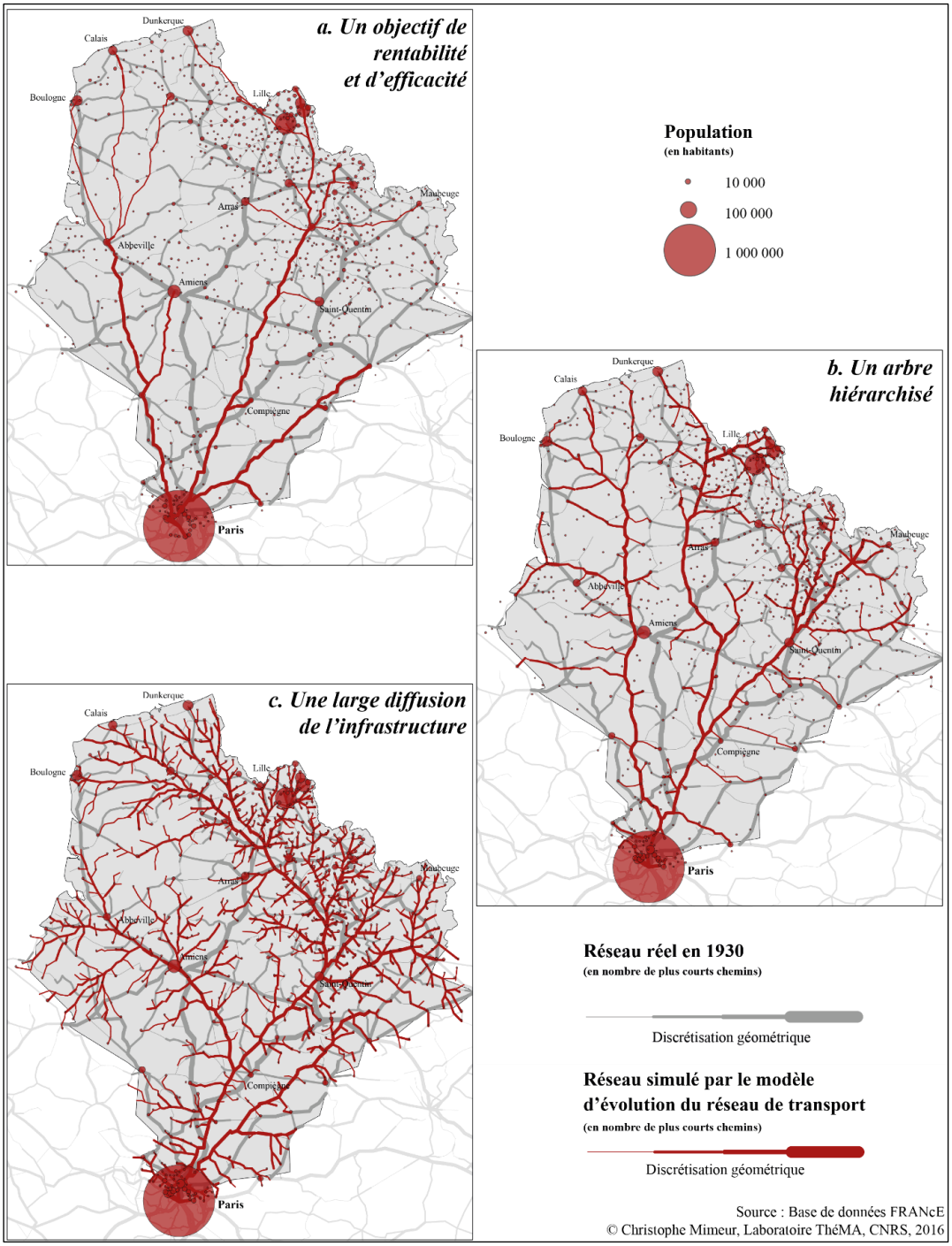


Figure 8. 18. Simulation de l'infrastructure sur le territoire de la Compagnie des Chemins de Fer du Nord

Pour autant, la proximité des points terminaux des arêtes dessinées par le réseau suggère encore davantage la jonction de ces points pour construire une infrastructure côtière. Cela vient confirmer que c'est à partir d'une structure plus complète que la circuité peut être envisagée, sans qu'elle soit dessinée par la simulation, mais seulement suggérée. Il en est de même pour l'axe historique entre Amiens et Lille : il répond à un tracé direct et nécessite la prise en compte d'une situation partielle de la hiérarchie urbaine Au nord et à l'est du territoire, on remarque une densité élevée créée par la simulation, qui propose une liaison le long de la frontière belge, doublé d'un axe vers le littoral. Cette dernière simulation suggère des inégalités dans la densité d'infrastructures.

Nous proposons de la confronter à la densité effectivement constatée avec le réseau réel de 1880, à partir d'un carroyage de 10 kilomètres de côté *(Figure 8. 19)*. Après avoir confronté les deux situations, leur soustraction permet d'évaluer la qualité de l'estimation des densités par la simulation.

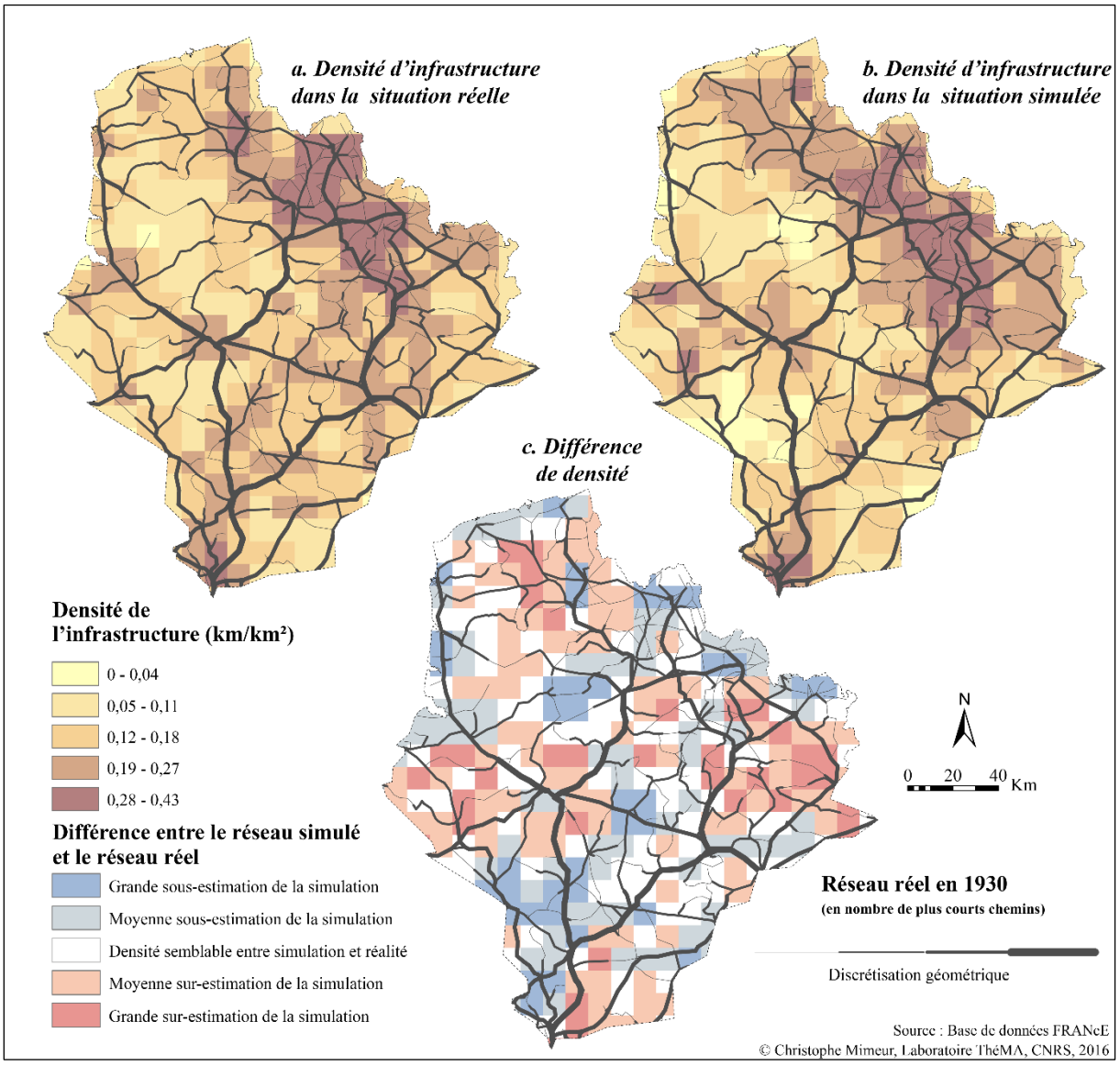


Figure 8. 19. Comparaison des densités d'infrastructure entre réseau réel de 1930 et réseau simulé

L'étude successive des deux situations *(Figure 8. 19. a. et b.)* montre des similitudes que le réseau simulé arrivent à reproduire. La situation réelle montre bien un déséquilibre entre est et ouest dans la dotation en infrastructures, où les fortes densités de communes peuplées entre Lille et Valenciennes semblent influencer la croissance du réseau ferré. Le réseau simulé suggère même de fortes densités dans l'ensemble de la zone frontalière avec la Belgique, de Saint-Quentin jusque sur le littoral. Ailleurs, les plus fortes densités sont constatées au départ de Paris, ainsi que le long des principaux axes, entre Abbeville et Saint-Quentin. La situation est plus diffuse dans la situation simulée, où le paramètre de concentration privilégie la branche commune au départ de Paris et ne diffuse que peu l'infrastructure jusqu'à Amiens ou jusqu'à Compiègne.

La *Figure 8. 19. c.* confronte directement les deux situations. Elle montre la sous-estimation du modèle au nord-ouest de Paris pour rejoindre le littoral, de même que le long de celui-ci. En effet, le réseau simulé privilégie les liaisons vers l'Ouest et on voit comment les décisions d'aménagement et la cohérence avec les autres réseaux participent au rééquilibrage progressif du réseau *(Figure 8. 19. a.)*. En revanche, la simulation surestime le réseau au nord-est du territoire, où on suppose qu'elle est influencée par la forte densité de communes peuplées. La surestimation qu'il subsiste entre Lille et le littoral est à chercher dans les difficultés du modèle à sélectionner des axes avec un paramètre de concentration modéré : il tend à accentuer les disparités.

La simulation du réseau à une échelle plus locale montre comment des facteurs exogènes viennent contrarier la situation d'un réseau simulé. La simulation est capable de reproduire un réseau dense au nord-est du territoire de la Compagnie du Nord tandis qu'il sous-estime la construction à l'Ouest. Nous posons alors l'hypothèse que les dimensions d'aménagement participent au rééquilibrage dans le dessin des tracés, de sorte à rejoindre le littoral de part et d'autre de la région. Enfin, la circuité dans le réseau est envisagée par le modèle d'évolution du réseau de transport sans qu'il arrive à en sélectionner un parmi d'autres. Les règles qui régissent la construction du réseau privilégient les ramifications et la structure en arbre, parce que la concentration ne permet pas la redondance de chemins permise par une meilleure connectivité. A ce stade, on peut donc conclure que notre modèle favorise la connexité plutôt que la connectivité. Pour autant, le rapprochement des extrémités de certaines branches laisse à penser à une jonction possible et rentable, notamment dans les zones les moins denses à l'ouest de la région.

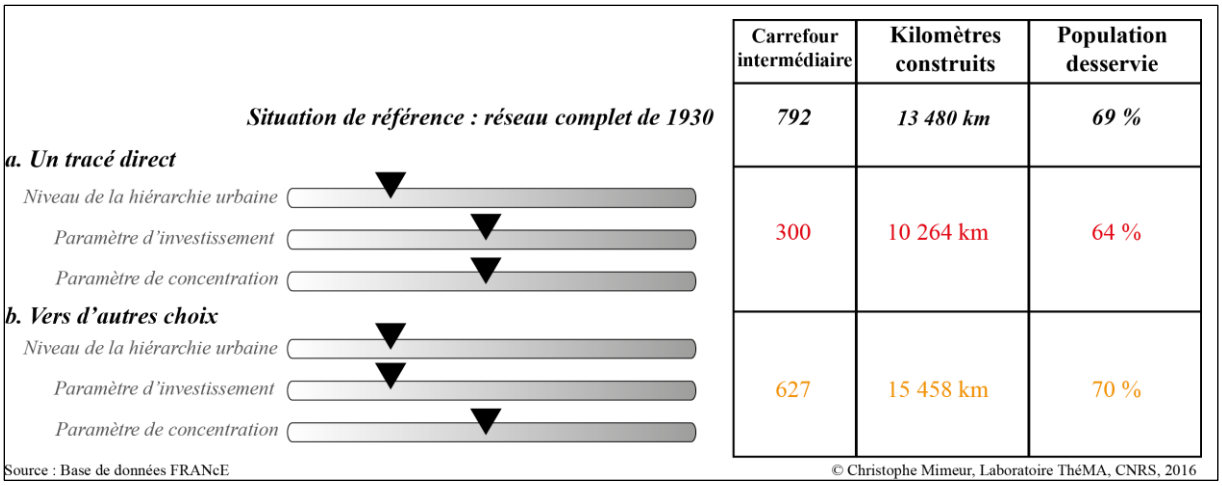
Il s'agit désormais d'éprouver le modèle à la même échelle, sur un autre territoire, où l'on connaît les contraintes d'aménagement qui ont prévalu sur l'efficience économique, et qui nécessite d'adapter le modèle pour en rendre compte.

8.3.2. **Le réseau du PLM : la force du politique pour éviter le Morvan**

L'axe Paris-Lyon-Marseille a été tracé par l'Etat en attribuant les concessions à plusieurs compagnies concentrées dans la Compagnie du Paris-Lyon-Méditerranée en 1857. La simulation du réseau à l'échelle nationale montre qu'il s'agit d'une des seules radiales que le modèle n'arrive pas à reproduire, parce que ce dernier suggère toujours un axe direct entre Paris et Lyon, alors que

la réalité du tracé propose un passage par Dijon, avant de cheminer le long de la vallée de la Saône puis du Rhône. Jean-Jacques Bavoux souligne que de nombreux témoignages s'accordent à dire que « *Dijon doit sa fortune à l'ingénieur Darcy* ». Dès lors, l'analyse des tracés de cette région doit se faire dans « *l'histoire mouvementée du flux Paris-Lyon* » (Bavoux, 1994).

L'objectif de cette sous-section est d'explorer comment le modèle d'évolution du réseau de transport se comporte à l'échelle du territoire de la Compagnie du PLM. Trace-t-il toujours un axe direct ? Suggère-t-il un passage par Dijon ? Si non, comment peut-on transcrire un acte fort d'aménagement dans le modèle, afin d'étudier la configuration du réseau qui en résulte. Notre démarche est de montrer dans quelle mesure l'évolution des paramètres peut participer à une meilleure compréhension de la situation réelle, en mobilisant à la fois le modèle, le contexte local et l'arrière-plan historique pour la construction du réseau ferroviaire français. Les paramètres des simulations pour répondre à ces questions sont contenus dans la *Figure 8. 20*, en s'appuyant sur le réseau maximal de 1930 comme situation de référence.



Carrefour intermédiaire	Kilomètres construits	Population desservie
792	*13 480 km*	*69 %*
300	10 264 km	64 %
627	15 458 km	70 %

Figure 8. 20. Les paramètres du modèle pour simuler le réseau du PLM

La première simulation reprend le même niveau de la hiérarchie urbaine et les mêmes paramètres que pour l'installation du réseau national *(cf. Figure 8. 14.c.)*. Les performances sont inférieures à celles du réseau réel, même si elles permettent de desservir 64 % de la population initiale *(Cas a)*. Pour autant, en prenant les communes de plus de 2 500 habitants, cette simulation ne produit que 300 carrefours intermédiaires. La confrontation visuelle des deux réseaux suggère toujours un trajet direct entre Paris et Lyon. Il traverse l'Yonne et le Morvan pour rejoindre Lyon. Cette branche commune permet la desserte simultanée de villes situées sur l'axe Paris-Lyon-Marseille réel et sur l'axe Paris-Nevers réel. Cette configuration est encore renforcée quand les paramètres sont augmentés. On peut alors postuler que le cheminement vers Dijon pour rejoindre Lyon répond à un paramètre d'investissement plus faible qui favorise une plus large diffusion du réseau.

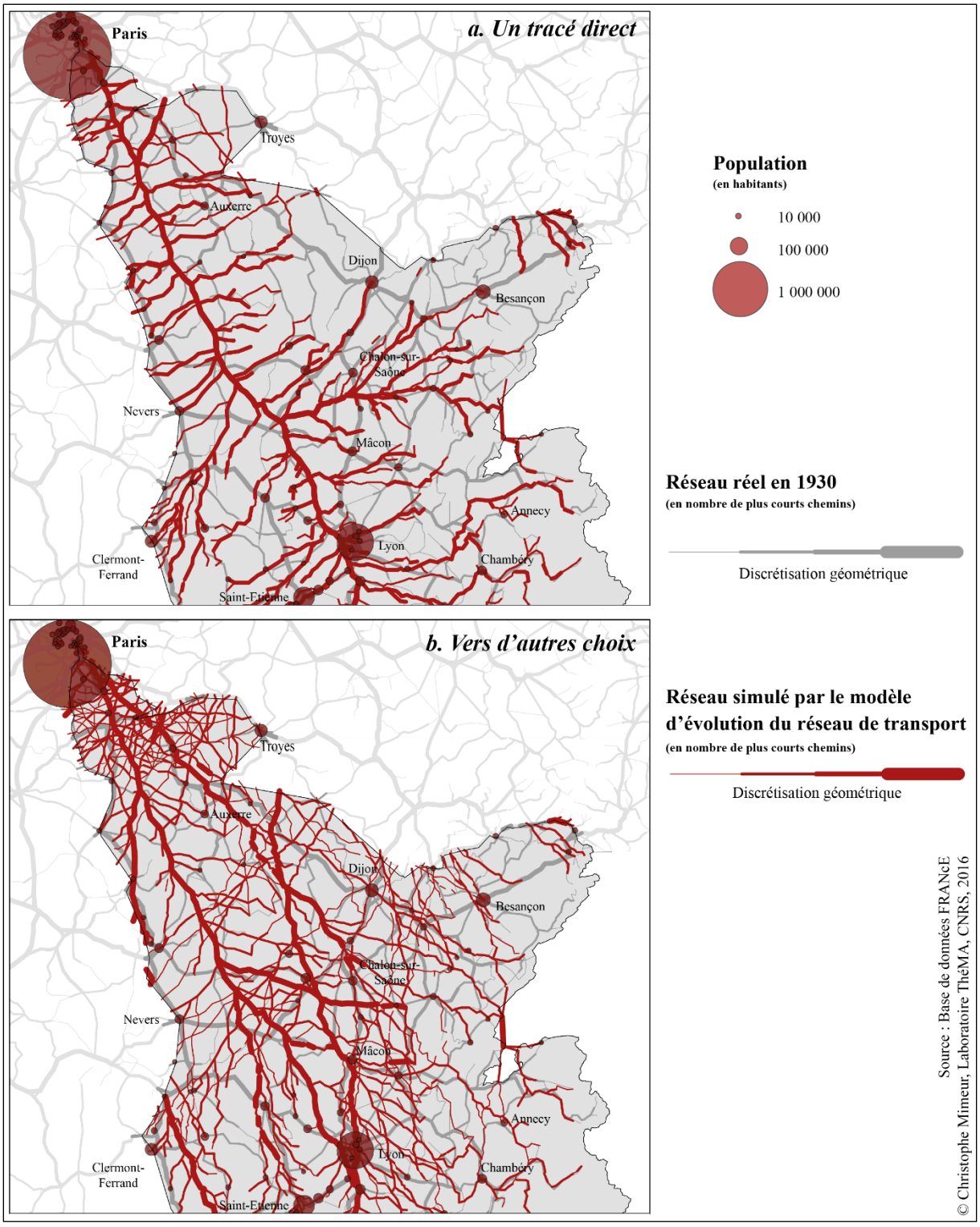


Figure 8. 21. Simulation de l'infrastructure sur le territoire du PLM

La seconde simulation propose un réseau qui paraît peu vraisemblable mais qui permet d'étudier l'étendue de la possibilité de cheminer. Pour la première fois, on peut identifier des capacités circulatoires dans le réseau *(Cas b)*. Pour autant, le nombre de kilomètres construits est supérieur au niveau construit, malgré de bonnes performances pour les deux autres critères. Cette

confrontation a surtout pour objectif de montrer quels cheminements sont envisageables pour rejoindre Dijon et longer les vallées de la Saône et du Rhône. Si le tracé direct est toujours préféré, un tracé se prolonge jusqu'à Dijon, qui se superpose à l'infrastructure réelle mais dont la place dans la hiérarchie est faible. L'alternative qui ressort le plus est un cheminement par Auxerre, puis qui se dirige vers l'Autunois pour rejoindre la vallée de la Saône à Mâcon.

Etonnamment, le modèle nous pousse à réintroduire des éléments de contexte qui ont recours à la très longue durée et dans une concurrence très ancienne entre deux axes de passage en Bourgogne : l'axe du Bourbonnais et l'axe Bourguignon (Bavoux, 1994). Ce débat a peiné à être tranché, tant *« les voyageurs de Paris à Lyon préfèrent cette route (du Bourbonnais) à celle qui se dirige par Mâcon, sans qu'on puisse en donner une cause valable »*[59]. Dès lors, l'administration centrale semblait avoir les coudées franches pour instaurer une nouvelle voie, la plus rapide, qui consiste à traverser la Bourgogne. Pourtant, ce n'est pas celle qui a été retenue à cette époque : les débats se concentraient autour du choix entre Chalon et Dijon. Si *« aucune n'offrait un avantage matériel décisif, le tracé dijonnais a manifestement triomphé sur d'autres terrains »* (Bavoux, 1994). Ce sont ces autres terrains que nous allons traiter en les introduisant de manière simplifiée dans le modèle, à partir de la lecture des décisions politiques et commerciales :

La concurrence avec la Compagnie du Paris-Orléans et les arbitrages de l'Etat poussent à l'abandon progressif des projets nivernais, dont la priorité se trouvait abaissée parce que contenue entre les radiales vers Toulouse d'une part et Lyon d'autre part. Par ailleurs, la volonté de l'Etat-Major était de repousser un maximum l'axe des frontières, mais *« la chance de Dijon fut de profiter des rapports entre ces deux lignes [Paris-Lyon et Paris-Strasbourg] »* (Bavoux, 1994). Alors dans ce cadre, la position de Chalon-sur-Saône est mise à mal : c'est d'ailleurs confirmé par la simulation du réseau de transport à l'échelle nationale où l'axe Paris-Alsace tend à privilégier parfois une branche en direction de Dijon mais qui n'arrive jamais à se prolonger vers la vallée de la Saône. Désormais, tous les ingrédients sont là, sinon pour encourager un tracé, pour en éliminer d'autres.

De cette sorte, nous choisissons de contraindre le modèle d'évolution du réseau de transport en le contraignant à ne pas s'implanter dans une zone. Nous la définissons comme un losange entre Paris, Dijon, Lyon et Nevers : le tracé direct n'est désormais plus possible, le décrochement par Chalon-sur-Saône non plus. En revanche, nous conservons les deux axes préférentiels et historiques du Bourbonnais et de la Bourgogne. Nous exécutons le modèle à partir de cette nouvelle situation initiale *(Figure 8. 22)*. Un avantage indéniable est la possibilité d'étudier l'arbitrage entre les deux axes historiques. En revanche, l'inconvénient principal est la perte d'information sur le réseau local dans la zone où nous forçons le modèle à ne plus implanter d'infrastructures. Dans ce sens, nous retenons toutefois l'importance d'un axe qui traverse le bassin minier de Saône-et-Loire pour rejoindre la Côte vers Dijon. Nous pouvons dès lors interpréter le nouveau réseau simulé au seuil initial de 1 000 habitants. Nous conservons les paramètres de la dernière simulation pour favoriser la diffusion de l'infrastructure. Les principales caractéristiques sont résumées dans la *Figure 8. 23*.

[59] D'après l'Almanach historique et politique de la ville de Lyon cité par Jean-Jacques Bavoux en 1994.

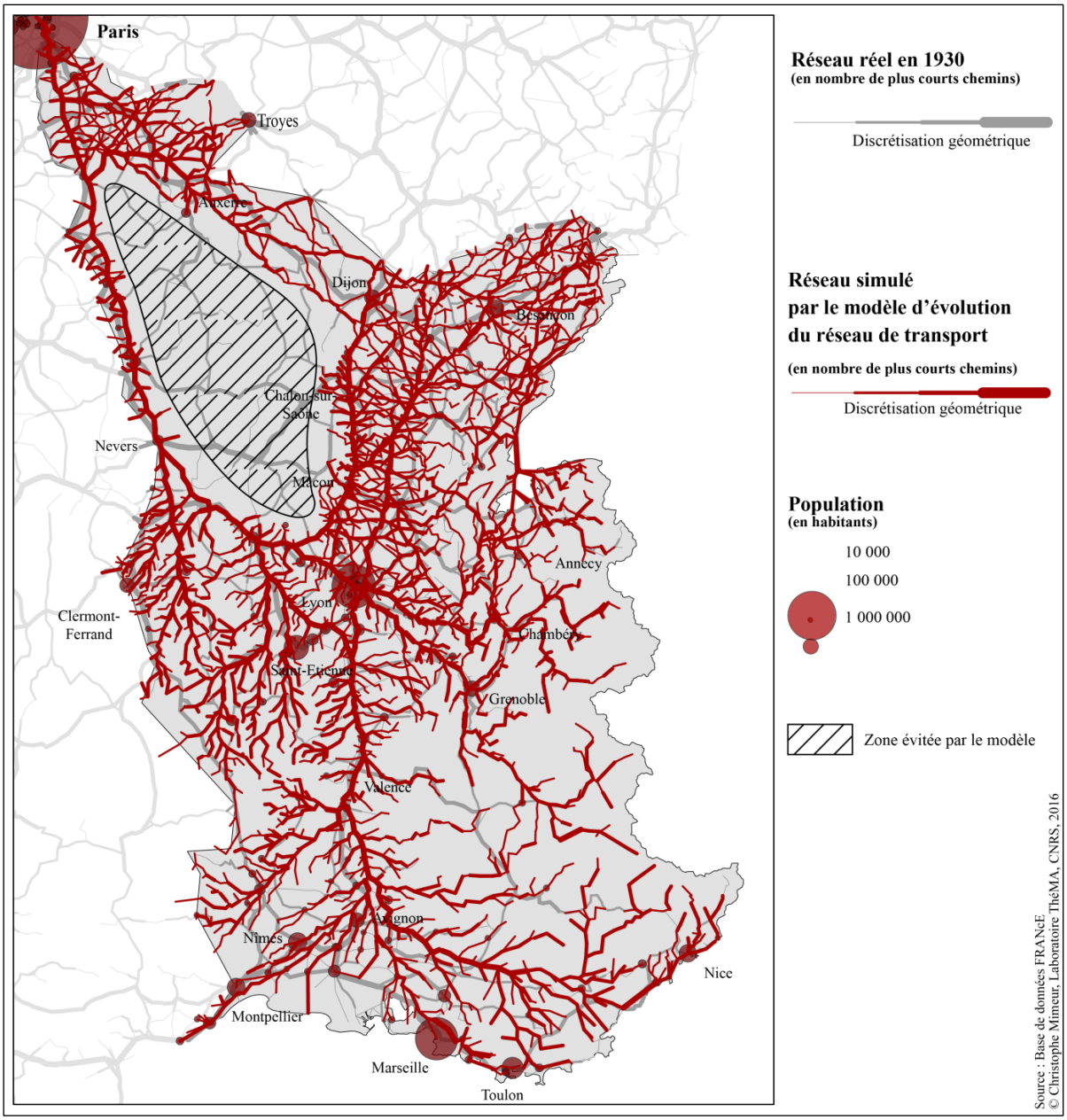
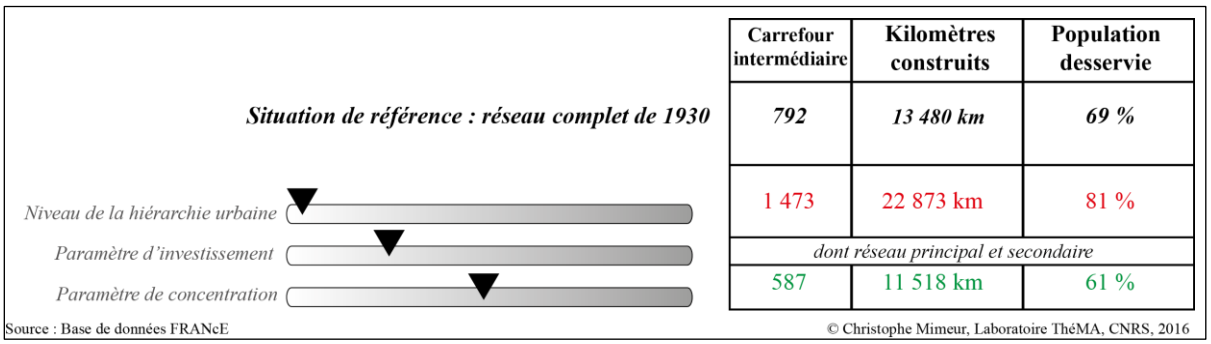


Figure 8. 22. Simulation de l'infrastructure sur le territoire du PLM avec une contrainte dans le modèle



	Carrefour intermédiaire	Kilomètres construits	Population desservie
Situation de référence : réseau complet de 1930	*792*	*13 480 km*	*69 %*
	1 473	22 873 km	81 %
	dont réseau principal et secondaire		
	587	11 518 km	61 %

Figure 8. 23. Les paramètres de simulation avec une contrainte dans le modèle

Les paramètres aboutissent à une diffusion exacerbée du réseau, avec près de 23 000 kilomètres construits qui desservent 81 % de la population. Ainsi, nous sélectionnons les valeurs les plus élevées des capacités circulatoires sur le réseau simulé, à partir des deux premières classes qui obéissent à une distribution géométrique. Dès lors, on se rapproche du réseau effectivement construit en 1930, avec 61 % de la population desservie par près de 12 000 kilomètres.

Le réseau simulé suggère désormais les deux axes historiques qui se dirigent vers le sud-est par le Bourbonnais et vers Dijon. Cette répartition répond à une cohérence globale de la structure stellaire au départ de Paris. Le partage de cette première branche commune s'établit en Fontainebleau et Montereau-Fault-Yonne à l'image du réseau réel. L'axe entre Paris Nevers est bien reproduit par le modèle en privilégiant la ligne droite, même jusqu'à Clermont-Ferrand. Le partage à Vichy pour rejoindre l'agglomération lyonnaise mais aussi Saint-Etienne est fidèlement reproduit par ailleurs, avec un tracé qui privilégie fidèlement la ligne droite.

La fameuse diagonale entre Paris et Dijon est cette fois bien représentée et le décrochement vers Auxerre est lui-même représenté par le réseau simulé. Le tracé entre Auxerre et Dijon est influencé par les obstacles topographiques, et la percée de Blaisy semble ainsi un facteur important qui a permis de dévier le tracé par la vallée de l'Ouche. La contrainte imposée au modèle, qui traduit la volonté des aménageurs du passage par Dijon, parvient par la suite à modéliser le réseau au sud et à l'est de l'agglomération dijonnaise. La côte viticole est traversée par le chemin de fer qui rejoint ensuite la vallée de la Saône jusqu'à Lyon.

La prise en compte d'un paramètre d'investissement faible permet en outre l'identification de capacités circulatoires qui sont à un niveau élevé de la hiérarchie du réseau simulé. L'axe bressan qui remonte vers la vallée du Doubs est ainsi modélisé, de même que ces raccordements à l'armature principale à hauteur de Mâcon, Chalon et Dole. Cette forte densité de chemin de fer semble fidèlement reproduite à l'est de Lyon. On peut conclure ici qu'en dépit d'un acte fort d'aménagement qui privilégie le passage de l'axe Paris-Lyon par Dijon, le déploiement du réseau répond par ailleurs fidèlement à une structure démographique préexistante.

Une deuxième valeur ajoutée du modèle peut être retranscrite grâce à la simulation du réseau sur le territoire du PLM. L'initialisation du modèle permet d'établir des tracés qui tiennent compte des obstacles topographiques : l'exemple du relief alpin est criant. Les cheminements vers Albertville et la vallée de la Tarentaise, Chambéry et la vallée de la Maurienne, Briançon et la vallée de la Durance sont fidèles aux obstacles. Pour autant, le modèle ne simule pas les axes transversaux entre Grenoble et Nice, via Gap et Digne-les-Bains, attestant de l'importance d'autres facteurs dans l'implantation de cette ligne.

Finalement, ces deux focus régionaux nous ont permis d'intégrer différents niveaux dans l'interprétation de la simulation d'un réseau de transport. Le premier permet de comparer le déploiement du réseau dans la Compagnie des Chemins de fer du Nord. Il montre que le déséquilibre de la dotation en infrastructures est le fait d'une densité de villes importantes. Ainsi, la volonté

d'aménagement de rééquilibrer le réseau vers le littoral a permis à l'axe Paris-Amiens-Lille de devenir structurant.

Par ailleurs, l'étude du déploiement du réseau dans la Compagnie du Paris-Lyon-Méditerranée vient confirmer l'importance dans la densité de villes dans la capacité du modèle à traduire des circuités dans le réseau. Aussi, des décisions majeures d'aménagement ont pu être directement pris en compte dans le modèle, venant confirmer le statut de la Bourgogne comme celui d'un *« espace de circulation « sous influence » par où transitent des trafics sur lesquels les décideurs n'exercent pas d'emprise à peine une empreinte »* (Bavoux, 1994).

Plus globalement, ces conclusions viennent apporter des éléments de réponse aux affirmations de Jean-Marc Offner sur le développement endogène des réseaux de transport (Offner, 1993b) : s'il s'agit pour lui d'une thèse réfutable, alors la comparaison d'un réseau simulé et d'un réseau réel montre bien la contribution d'une situation initiale au développement endogène du réseau ferroviaire français. Il démontre comment la performance radiale a souvent été pensée et privilégiée par les ingénieurs de l'époque. En revanche, le modèle éprouve de grandes difficultés à faire émerger des capacités circulaires dans le réseau. Nous avons proposé des pistes pour les envisager : en calculant des potentiels de distance ou en forçant le modèle à ne pas traverser certaines zones.

L'étude progressive des paramètres et des configurations amène à l'interprétation des écarts, qui sont à lire comme l'intervention de facteurs exogènes, comme l'intervention et la pression politique d'un notable local. Ces écarts peuvent être aussi lus par l'établissement de potentielles liaisons, qui participent au renforcement d'un système territorial à des échelles intermédiaires. Les premières transversales proposées par notre méthode montrent comment elles peuvent participer à des systèmes régionaux.

Conclusion

Ce huitième chapitre a permis de saisir les nombreuses avancées en termes de modélisation dynamique en géographie pour étudier l'adaptation des tracés du réseau ferroviaire français à une structure démographique préexistante. Il s'inscrit durablement dans notre question de recherche, dans la mesure où son fort apport méthodologique participe à la construction de connaissances géohistoriques, et ainsi à la montée de la préoccupation thématique des effets structurants entre transport et territoire : cela va dans le sens de notre positionnement qui défend la valorisation de la donnée pour la production de résultats originaux *(Figure 2. 6, p. 86)*. Pour cela, le modèle de génération d'un réseau de transport traduit le déploiement endogène d'une structure émergente, comme un préalable indispensable pour *« faire réseau »* (Ribeill, 1985b). Ce modèle traduit bien la dimension cinétique (Dupuy, 1987b) traduite par les théoriciens du réseau et par les paramètres d'investissement et de concentration des capacités circulatoires. Alors la structuration de l'information géohistorique nous permet d'éprouver le graphe modélisant l'expansion du réseau ferroviaire pour étudier les capacités circulatoires issues des choix d'investissement dans la vitesse ferroviaire. Le double phénomène de concentration et de hiérarchisation traduit bien cette dimension cinétique, qui apparaît indispensable à la diffusion de la fièvre ferroviaire, mais qui apparaît tout aussi contradictoire avec le mythe de l'association universelle.

Ces similitudes de formes nous autorisent à simuler le déploiement du réseau ferroviaire à l'aide du modèle d'évolution du réseau de transport. Au-delà de la dimension cinétique pris en compte dans les modules statiques et dynamiques du modèle, l'interprétation des écarts entre le réseau et le réseau réel participe à l'explication de la dimension adaptative (Claval, 1981 ; Raffestin, 1980), elle aussi évoquée par les théoriciens du réseau. Le modèle montre que la structure stellaire est le fait d'une prise en compte partielle et par le haut de la structure démographique tandis que le choix de transversales est une décision forte d'aménagement, dont les impacts ne sont perceptibles qu'à des structures initiales plus larges. Le modèle démontre aussi la nécessaire hiérarchisation afin de rendre possible le déploiement du réseau dans des régions plus reculées. Il laisse entrevoir quelques opportunités pour établir des transversales, dont l'efficience économique peine toutefois à être démontrée.

Enfin, des analyses plus locales de la simulation d'un réseau de transport dans des régions plus compactes démontrent une relation entre la densité de villes et la densité de la dotation en infrastructures. Les écarts entre les réseaux réels et simulés montrent comment la dimension adaptative cherche à rééquilibrer les inégalités dans un souci d'équité territoriale. Enfin, la sophistication du modèle par la contrainte du tracé permet de traduire la force des décisions politiques dans le dessin du réseau réel, qui laisse par ailleurs entrevoir de nouvelles perspectives dans le développement du modèle de génération d'un réseau de transport dont les préoccupations peuvent d'ailleurs dépasser le cadre de ce travail de recherche, pour envisager les mécanismes qui ont conduit à la décroissance des réseaux de transport, à l'heure où l'installation des grandes régions réinterrogent la réduction du réseau ferroviaire.

Chapitre 9. LA BASE DE DONNEES FRANCE, UN PROJET ENTRE ELARGISSEMENT ET APPROFONDISSEMENT

Introduction

La première partie de la thèse s'est attachée à la définition des concepts et des cadres épistémologiques de notre question de recherche. La seconde partie a permis de définir les contours méthodologiques de l'analyse à de telles échelles. La troisième partie s'est attachée à démontrer comment ils peuvent contribuer à la connaissance géohistorique du lien entre réseau et territoire. Ce chapitre vient remettre en perspective la question de recherche à partir des questions posées et des résultats produits.

Les concepts d'élargissement et d'approfondissement sont empruntés à la construction européenne (Dezert et Vandermotten, 2008) : pendant longtemps, leur combinaison est parue incompatible, comme si l'élargissement avait été une source de blocage pour l'approfondissement du projet européen[60]. Alors que nous avons replacé tout au long de cette thèse la base de données FRANcE au cœur d'un plus vaste projet de SIG-H, nous transposons ces concepts d'élargissement et d'approfondissement dans ce chapitre qui ouvre de nouvelles perspectives quant à son utilisation, sa conceptualisation et sa capacité à produire de nouvelles connaissances géohistoriques. En questionnant le dimensionnement qu'il convient à ce projet, nous remettons en perspective la *Figure 2. 6 (p. 86),* qui porte sur les challenges d'un projet géohistorique, en démontrant les nécessaires allers et retours entre préoccupations thématiques et méthodologiques. Plus largement, il s'agit de repenser le projet de la base de données FRANcE dans l'évolution des humanités numériques, dont le champ de recherche évolue d'une approche d'abord quantitative vers de nouvelles potentialités à couvrir de multiples aspects de la géographie du passé (Gregory et Geddes, 2014). Dès lors, ce chapitre vient confirmer le cap et le positionnement à adopter pour développer ce projet ou à en initier d'autres, en faisant appel à l'interdisciplinarité et l'exigence de la définition des objets de la recherche (Grossetti, 2011 ; Pumain, 2005, 2014), afin d'asseoir la scène géohistorique dans le champ de la géographie (Bol et Ge, 2005 ; Cresswell et al., 2015 ; Grataloup, 2015) et dans le paysage de la géographie française en particulier.

La métaphore de l'élargissement va ainsi nous permettre d'analyser les perspectives scalaires d'un tel projet, tant la notion d'échelles a été au centre de notre question de recherche. On pose ici les questions de l'ouverture à d'autres échelles spatiales, à d'autres échelles de temporalités ainsi qu'à d'autres échelles d'innovations ou de réseaux (9.1.). Ces nouvelles possibilités n'entravent en revanche aucunement les nombreuses opportunités d'approfondissement, qui sont même complémentaires. Il s'agira alors d'étudier le sens que l'on souhaite donner à ce projet, à travers le nécessaire dialogue avec d'autres disciplines et l'appel à de nouvelles méthodes, en posant également la question de la diversité des publics auxquels se destine l'approfondissement du projet (9.2.).

[60] L'idée d'une complémentarité est alors au centre d'une « Question d'Europe » défendue par la Fondation Robert Schuman, Centre de recherches et d'études sur l'Europe, en 2007, à l'occasion de l'adhésion de la Bulgarie et de la Roumanie. http://www.robert-schuman.eu/fr/questions-d-europe/0049-apres-l-adhesion-de-la-bulgarie-et-la-roumanie-en-finir-avec-l-opposition-entre-elargissement

9.1. L'élargissement : une théorie des réseaux *« à la française »* ?

La construction de la base de données FRANcE a permis de mettre en œuvre un modèle conceptuel de données que nous avons voulu évolutif *(cf. 3.1.3, p. 111)*. Cette adaptabilité a été éprouvée quand il a fallu diversifier l'information géohistorique, en introduisant des indicateurs de développement économique à des échelles agrégées ; et quand il a fallu transposer l'information géohistorique aux attributs du graphe, à travers ses nœuds et ses liens. Dès lors, au-delà des résultats produits dans le cadre de la thèse, nous estimons que le modèle conceptuel de données peut concourir aux questionnements sur un *« modèle générique »* du développement des réseaux techniques (Offner, 1993b). Si l'analyse empirique du développement du réseau et la mise en œuvre du modèle d'évolution du réseau de transport ont montré l'adéquation entre la croissance du réseau ferroviaire français et les mécanismes d'une théorie des réseaux, largement construite dans la littérature française, ces mêmes composantes régissent-elles le développement du réseau ferroviaire sur d'autres terrains ou à d'autres échelles ? Conformément à nos réflexions sur les échelles *(cf. Chapitre 2, p. 64)*, nous voyons dans nos perspectives de recherche des opportunités d'élargissement des échelles spatiales et temporelles, en questionnant la longue durée et l'échelle de l'innovation. Ici, l'expérience acquise sur la collecte de l'information, la capitalisation du corpus et les analyses démontrent la complémentarité des trois aspects, et comment une structure conceptuelle générique peut faciliter la très consommatrice phase de collecte.

9.1.1. **Elargir le cadre spatial**

L'élargissement du cadre spatial dans le cadre de l'adaptation du modèle conceptuel de données PONT passe alors par la considération des trois volets du MCD : le réseau, la population et le territoire. Les principales perspectives retrouvent l'essence des projets de SIG-H, en posant les questions de l'adaptabilité nécessaire de l'outil, problématique bien connue de la discipline (Knowles, 2005).

Il s'agit d'abord d'évoquer la composante *« Network »* du modèle conceptuel. Initié en 2007, le programme européen dans lequel s'est inscrite la numérisation du réseau ferroviaire français *(cf. 3.1.2, p. 106)*, a fait l'objet de multiples publications quant aux conditions de sa conception et de son inventaire (Martí-Henneberg, 2013 ; Morillas-Torné, 2012). Pour autant, si des analyses ont été largement citées dans le cadre de ce document à propos d'études de cas nationales et régionales, aucune étude intégrée de l'impact du réseau ferroviaire à l'échelle européenne n'a été menée à ce jour sur la longue durée, à notre connaissance (Kasraian et al., 2016). On connaît en revanche l'étude des évolutions des accessibilités, sur des périodes plus courtes, portant sur le réseau autoroutier européen (Dupuy et Stransky, 1996) et le réseau à grande vitesse européen (Gutiérrez et Urbano, 1996). Dès lors, il semble naturel de réinscrire la base de données FRANcE dans des perspectives plus larges, en prenant appui sur le corpus constitué par l'équipe espagnole de Jordi Marti-Henneberg.

Cette base de données a déjà montré comment la France est le pays qui concentre le plus de kilomètres de voies ferrées à l'apogée du système ferroviaire en Europe *(Table 9. 1)*. Elle permet de réinterroger à de plus larges échelles et de manière comparée la phase d'évolution de l'innovation ferroviaire *(Figure 1. 3, p. 34)*. Se caractérise-t-elle dans tous les pays par un phasage en trois actes, de la mise en connexité à la maturité en passant par une phase de développement très accélérée et équitable sur l'ensemble du territoire concerné ? La temporalité de ces phases obéit-elle aux règles de l'innovation technique ? Cette approche permet de remettre en perspective les proximités entre l'innovation technique et sa diffusion spatiale (Dupuy, 1993 ; Garrison, 1990).

Rk	Pays	1870	1900	1930	2000
1	France	15,344	32,905	37,137	29,648
2	Germany	15,796	37,011	33,106	28,961
3	Italy	5,384	12,820	15,802	16,061
4	United Kingdom	20,115	28,280	30,939	15,505
5	Poland	-	-	4,471	14,736
6	Spain	4,992	9,725	10,621	11,583
7	Sweden	1,081	3,238	4,162	8,712
8	Czech Republic	-	-	3,510	7,247
9	Hungary	2,633	14,222	5,767	6,542
10	Finland	0	2,255	3,827	5,179
11	Austria	4,368	9,478	4,719	4,894
12	Norway	325	1,700	3,193	3,424
13	Switzerland	1,312	2,862	3,286	3,255
14	Slovakia	-	-	2,332	3,109
15	Belgium	2,575	3,799	4,080	2,689
16	The Netherlands	1,217	2,525	3,144	2,480
17	Portugal	632	2,145	3,079	2,453
18	Denmark	681	2,231	3,206	2,254
19	Greece	6	6	0	1,596
20	Ireland	2,866	3,825	3,111	1,521
21	Luxembourg	150	281	281	195
22	Liechtenstein	0	9	9	9

© Thomas Thevenin

Table 9. 1. Nombre de kilomètres de voies ferrées par pays dans le corpus européen entre 1870 et 2000, d'après Thévenin, 2010

Une telle initiative requiert de nouvelles expertises quant aux sources qui ont permis cette digitalisation, à partir de laquelle nous pouvons reconstruire l'histoire des réseaux ferroviaires en Europe. En effet, les initiatives se sont jusqu'à présent focalisées sur l'évolution de la structure du réseau, à travers ses tracés et son inscription territoriale par la présence d'une gare dans les unités spatiales étudiées. Désormais, il s'agit de s'interroger sur la concordance, la convergence des

informations à propos des différents réseaux afin de produire des analyses qui couvrent l'ensemble du territoire européen et portant sur l'évolution des accessibilités. Nous retrouvons alors un écueil inhérent à chaque projet de SIG-H (Gregory et Healey, 2007b), qui se doit de croiser des sources très hétérogènes et souvent partielles. Nul doute que ces difficultés ont contribué jusqu'à maintenant à la rareté des résultats. Dans ce cadre, il s'agit d'éprouver la stratégie de généralisation des vitesses, sur la base du travail effectué à l'échelle française, en croisant les sources littéraires et historiques, capables de nous informer quant aux vitesses exercées sur le réseau européen. Le réseau britannique a déjà fait l'objet d'une numérisation détaillée de lignes de chemins de fer à partir de l'atlas de Michael H. Cobb. En revanche, les vitesses de circulation n'y sont précisées que pour certaines lignes. Alors, dans la lignée de cette thèse, l'objectif est de tester la reproductibilité de cette approche sur d'autres terrains.

Avec l'incorporation des vitesses de circulation à l'échelle européenne, la construction de nouveaux graphes fonctionnels du réseau ferroviaire permettra de produire de nouveaux indicateurs d'accessibilité sur le temps long et d'appréhender l'évolution des vitesses. Ainsi, alors que de nombreuses études se sont attachées aux axes les plus structurants dans la construction européenne (Rietveld, 1989 ; Vickerman, 1995) – ceux qui permettent de rejoindre les centres de décision, les pôles économiques –, cette généralisation a aussi pour objectif de mettre l'accent sur les réseaux secondaires, comme potentiel facteur de convergence européenne, au regard des développements économiques et démographiques souvent inégaux depuis le milieu du XIXème siècle. Ainsi, *« si le problème des réseaux secondaires est aujourd'hui posé, c'est en regard du développement des grands réseaux européens dont on a l'impression qu'ils contribuent à un espace plus concentré, et surtout, plus dualisé : d'une part, les zones à forte accessibilité autoroutière, ferroviaire et aérienne, d'autre part, des zones éloignées de ces réseaux rapides et dont on peut craindre qu'elles restent évincées des grands axes de développement. Le problème des réseaux secondaires et de l'intégration des chaînes de transport devient alors un problème d'arrimage aux portes d'entrée des grands réseaux. Il ne peut être traité sans que soient prises en compte les grandes tendances de nos systèmes de transport »* (Bonnafous, 1993). De cette sorte, il s'agit d'étudier comment, sur la longue durée, un réseau européen a été pensé autour de quelques axes structurants, influençant ainsi la stratégie d'investissement infrastructurel, tant par la présence même du rail que par l'amélioration de la vitesse sur le temps long. Dès lors, une approche par la hiérarchisation du réseau semble reproductible à partir des indicateurs de graphe utilisés dans le cadre de la thèse. Il s'agit aussi d'étudier dans quelle mesure les décisions nationales ont anticipé la construction européenne en termes d'infrastructures de transport : ces axes structurants sont-ils permanents sur le temps long ? La stratégie *a priori* isolée de chaque Etat a-t-elle contribué à la construction d'une cohérence des infrastructures de transport à l'échelle de l'Europe ? La stratégie des institutions européennes s'est-elle appuyée sur les corridors construits auparavant ou a-t-elle rebattu les cartes dans une visée d'aménagement du territoire ?

Pour autant, il ne faut pas lire dans ce projet la stricte application de larges échelles spatiales. Ainsi, nous avons vu comment des focus régionaux permettent de soulever des problématiques plus locales quant aux règles qui ont régi la construction du réseau ferroviaire français *(cf. 8.3.2, p. 351)*. A l'échelle européenne, l'étude d'espaces transfrontaliers semble d'après nous une opportunité décisive dans l'étude des relations entre transport et territoire, et permet de questionner le rôle de l'infrastructure dans le dépassement des contraintes politiques, géopolitiques ou administratives.

Ainsi, on sait comment le contexte politique a eu des impacts dans les stratégies de tracé, tant en Europe de l'Est qu'en Europe de l'Ouest. Là encore, il s'agit d'étudier dans quelle mesure la modélisation des données géohistoriques peut participer au débat sur les effets structurants et s'insérer dans un champ de recherche plus large, afin de donner des clés de réflexion ou des éléments participant à poser de nouvelles questions quant au rôle des frontières dans la structure des réseaux, sur le temps long, mais aussi dans une période contemporaine. Cette perspective pose également l'hypothèse de la prise en compte des frontières à l'intérieur même du modèle d'évolution du réseau de transport. On peut ainsi imaginer contraindre le modèle à investir ces liens transfrontaliers avec des coûts pondérés, et donc bien plus élevés. L'une des questions posées est alors la mesure dans laquelle le potentiel de circulation et d'échange est assez fort pour tout de même permettre l'investissement qui est alors très onéreux. L'adaptation du module d'investissement du modèle d'évolution d'un réseau de transport est alors nécessaire à la prise en compte de ce facteur exogène : jusqu'à quel seuil l'investissement est-il rentable ? Quelle signification géohistorique donner à ce seuil ?

Plus généralement, l'extension de nos analyses à l'échelle européenne semble avoir toute sa place dans la directive INSPIRE, qui prévoit une infrastructure unique d'information géographique dans l'Union Européenne (Noucher et Gautreau, 2013). Dans le cadre de cette directive, on retient plusieurs obligations qu'il est nécessaire de transposer à notre projet : *« la constitution de catalogues de données »*, *« l'application de règles d'interopérabilité »*[61]. La question de l'infrastructure à imaginer dans le cadre de ces ambitions fera l'objet de la section suivante. Dès lors, si la question de l'harmonisation des données est primordiale pour appréhender la modélisation des données qui concernent le réseau, l'étude de leurs relations et interactions avec d'autres éléments du territoire requiert la même exigence pour les composantes *« Population »* et *« Territory »* du modèle conceptuel.

A l'échelle européenne, la question du découpage administratif est certainement celle qui affecte le plus le volet *« Territory »* du modèle PONT. Ainsi si le *Modifiable Area Unit Problem* est traité depuis plusieurs décennies dans la littérature traitant de l'analyse spatiale (Wilson et Fotheringham, 2008), il trouve une résonance particulière sur le temps long et à l'échelle de l'Europe. Nous interrogeons dans ce cadre les découpages territoriaux et leur hétérogénéité dans le cadre de l'étude de plusieurs Etats. Ainsi, la diversité des découpages administratifs pose la question de l'homogénéisation des analyses et de la pertinence des agrégations choisies pour les analyses. Elles sont toutefois parfois nécessaires et c'est dans ce sens qu'a été constituée la Nomenclature des unités territoriales statistiques (NUTS). Ces questionnements s'inscrivent alors dans la démarche constructiviste adoptée dans la thèse, sur la contextualisation de l'information nécessaire à l'interprétation de sa mesure (Desrosières, 2010). L'une des réponses proposées à ces enjeux est la prise en compte de l'échelon administratif le plus fin (Gregory et Ell, 2007 ; Knowles, 2005). Cette initiative a déjà été suivie dans de nombreuses analyses géohistoriques à l'échelle nationale ou binationale (Alvarez, Franch et Martí-Henneberg, 2013 ; Franch, Morillas-Torné et Martí-Henneberg, 2013 ; Schwartz, Gregory et Marti-Henneberg, 2011 ; Silveira et al., 2013) : il s'agit de

[61] http://inspire.ec.europa.eu/

la paroisse au Royaume-Uni *(parish)* et au Portugal *(freguesia)*, la commune en Espagne *(municipio)* et en France. Ainsi, si les analyses françaises à l'échelle de la commune ont montré la persistance du découpage communal sur le temps long, le croisement de données de population provenant d'autres pays requièrent le recours à des opérations spatiales, parmi lesquelles l'interpolation permet de contourner le problème des mutations des frontières administratives (Gregory et Ell, 2006).

Si le problème des découpages administratifs trouve un écho dans les problématiques géomaticiennes, alors il pose également la question de la diversification de l'information géohistorique à de telles échelles, en élargissant les possibilités dévolues au volet *« Population »* du modèle conceptuel de données. L'objectif est de spécifier des modèles de relations plus fins, afin de limiter le fameux biais de variables omises qui tend à donner une importance surestimée aux effets du réseau. C'est encore une fois par le prisme des agrégations imposées par la nature des données que nous devons faire face à l'hétérogénéité de la donnée. Par ailleurs, sa relative rareté met souvent à mal la large couverture et nous contraint à des analyses à des échelles plus réduites. En revanche, dans notre positionnement épistémologique, leur multiplication participe à la véracité des analyses (Pumain, 2014). A notre connaissance, il existe des données très intéressantes portant sur les mouvements migratoires à l'échelle des paroisses britanniques : leur analyse promet de donner davantage de détails quant aux mutations démographiques, dues à l'accroissement naturel et aux mouvements migratoires, et à leurs liens avec la dynamique des réseaux de transport.

Par ailleurs, si nous avons déjà saisi l'opportunité d'ajouter des variables de spécialisation et de diversité économiques afin de contrôler l'effet du réseau sur les dynamiques démographiques, il existe de nombreuses autres données historiques entre 1800 et 1925[62]. Il s'agit de données qui ont été saisies dans les années 1980 à partir des ouvrages de la Statistique Générale de la France par des chercheurs franco-américains basés à l'*Inter-University Consortium for Political and Social Resedarch (ICPSR)*, à Ann Arbor, dans le Michigan. Pour autant, la nature des données est très hétérogène dans la mesure où la disponibilité n'est que souvent partielle, à des échelons souvent agrégés aux arrondissements ou aux départements. Mais dans le cadre d'analyses régionales, il semble indéniable que ces données participent à la prise en compte de davantage d'hétérogénéité dans la modélisation des relations entre transport et territoire. Comme enjeu inhérent à des études de long terme, il est difficile de mener des études longitudinales, dans la mesure où certaines variables ne sont présentes que pour un seul point dans le temps. Pour autant, il peut s'agir d'une étude isolée pour l'appréhension d'un phénomène plus particulier, comme l'accroissement des populations rurales et urbaines à l'échelle des départements. Il s'agit d'une difficulté qu'il faut chercher dans la conception de la statistique et de la quantification au XIXème siècle. Il faut interroger cette conception pour comprendre et anticiper les possibles analyses du lien entre réseau et territoire : dans les méthodes qui ont cours entre 1820 et 1900, *« le partage est ainsi net, entre, d'une part, une statistique couvrant le territoire par des recensements exhaustifs, publiés notamment selon le découpage départemental et, d'autre part, des monographies […] portent sur des espaces restreints et fluctuants, correspondant aux aires d'action ou d'influence d'acteurs*

[62] http://www.insee.fr/fr/bases-de-donnees/default.asp?page=tableaux_sgf.htm

variés, désignées par le vocable flou » (Desrosières, 1994). L'intérêt de l'utilisation de telles informations réside dans la mise à disposition aisée de ces données, dans la mesure où la phase de collecte ne requiert pas un recours à l'archive.

Finalement, cette sous-section nous permet de voir que l'extension des échelles spatiales pose directement la question des sources historiques, de leur diversité, leur hétérogénéité mais aussi leurs nouvelles opportunités. Ainsi, la pluralité des contextes oblige à éprouver le potentiel adaptatif du modèle PONT :

- L'élément *« Territory »* doit ainsi s'accommoder de la diversité historique des découpages administratifs inhérents à la collecte d'autres informations et à leur mutations sur le temps long dans d'autres pays européens ;
- L'élément *« Population »* peut aussi être diversifié vers des informations socio-économiques, dont la couverture nationale à une échelle fine est en revanche une limite ;
- L'élément *« Network »* semble revêtir des opportunités très riches, avec la prise en compte de caractéristiques temporelles et techniques, nécessaire à l'évaluation des accessibilités sur le temps long

D'ailleurs, il semble que ce dernier élément ne peut plus se limiter au seul réseau ferroviaire. Cela a déjà été envisagé dans cette thèse à partir d'une exploration de l'instrumentation des réseaux, mais la prise en compte d'autres réseaux paraît désormais possible, afin d'élargir le cadre temporel des relations entre réseau et territoire.

9.1.2. **Elargir le cadre temporel**

En questionnant à nouveau les échelles de temps, nous replaçons la croissance des réseaux dans l'analyse des innovations, de leur succession et de leur évolution. Dans son analyse du développement des réseaux techniques, Jean-Marc Offner affirme que *« c'est à l'échelle séculaire que perdurent les réseaux techniques, avec la même dénomination : métro parisien, réseau ferroviaire, système d'assainissement haussmannien… »* (Offner, 1993b). Il s'agit alors de s'imprégner de vocables qui n'étaient pas abordés dans le cadre de ce travail pour l'étude des effets entre transport et territoire, tels que la permanence, la reproduction, la superposition, la complémentarité, la concurrence, la substitution des réseaux de transport.

Comme nous l'avons déjà vu, *« l'arrivée d'un nouveau venu n'est que rarement synonyme d'éviction pour les réseaux antérieurs. Il y a plutôt adaptation, reconfiguration des positions »* (Offner, 1993b). Il existe à ce propos des monographies techniques, autour de l'histoire des technologies et de la succession des innovations. Pour autant, l'extension de l'échelle temporelle dans notre projet SIG-H permet de questionner ces éléments dans la dimension épistémologique que nous avons adoptée jusqu'à présent. Elle permet donc de ne plus se concentrer que sur le réseau

ferroviaire mais d'étudier les phases de croissance des autres réseaux. Dès lors, nous questionnons encore une fois la longue durée, à l'échelle des innovations qui sont mises en valeur, mais également cette fois la *« très longue durée »* (Grataloup, 2015). L'objectif est d'étudier en quoi la permanence, le parallèle des tracés entre les axes structurants des transports : il s'agit ici de mettre encore une fois à l'épreuve la composante *« Network »* du modèle conceptuel de données. Cet aspect a été déjà abordé dans la mobilisation des réseaux postaux et des réseaux de voies romaines dans l'exploration de l'instrumentation des temps de parcours moyens *(cf. 5.3.2, p. 209)*, et est l'occasion de questionner l'accessibilité géohistorique en fonction des terrains, mais aussi en fonction des périodes d'étude. Dans ce cadre, l'adaptation des modèles comportementaux est un préalable indispensable à la modélisation des réseaux. Dans le cas du réseau de voies romaines, nous avons fait l'hypothèse, qu'il faut considérer avec une très grande prudence, d'une quasi-intégralité des déplacements à pied, qui sont contraints par la présence d'axes pensés sous l'Empire Romain, avec différentes qualités de revêtements, ainsi que la structuration théorique du lacis de chemins de terres (Verdier et Bretagnolle, 2007) qu'il pouvait exister par ailleurs mais dont le franchissement est beaucoup plus difficile.

Dans le cadre de la modélisation des déplacements sur les voies postales, la littérature historique est beaucoup plus prolixe. L'impressionnant travail entrepris depuis le début des années 2000 par Anne Bretagnolle et Nicolas Verdier a permis de reconstituer le linéaire de voies postales tout au long du XVIIIème siècle jusqu'à la pose des premières rails (Bretagnolle, 2009 ; Bretagnolle, Giraud et Verdier, 2010 ; Bretagnolle et Verdier, 2005). La prise en compte des voies postales permet également de modéliser un système de transport qui obéit à certaines règles en termes d'infrastructures et de services qui y sont délivrés. A ce titre, il serait particulièrement opportun d'étudier l'évolution de ce système de transport au cours des deux siècles qui l'ont vu prospérer : des analyses similaires à celles entreprises dans cette thèse pourraient être conduites pour analyser la stratégie des tracés, la place de la rentabilité des échanges dans le dessin du réseau. L'approche par les vitesses pourrait permettre d'établir des parallèles entre la croissance du réseau postal et du réseau ferroviaire et ainsi de pouvoir affiner les estimations des effets de réseau à l'aide de la méthode des variables instrumentales, variantes dans le temps long.

Dans le cadre de l'extension de la dimension temporelle, il semble naturel de vouloir s'attacher aux possibilités de déplacement induites par le développement des voies navigables. *« La navigation intérieure représente une des formes d'utilisation des rivières et des fleuves dans la conquête de l'espace »* (Le Sueur, 1989), et sa large utilisation participe au développement des relations et des échanges dans la période préindustrielle. La prise en compte d'un tel système de transport participerait selon nous aux discours portant sur les rapports entre le transport et le territoire, tant les voies navigables semblent engendrer *« des espaces spécifiques, construits et aménagés par les hommes avec ténacité, en fonction d'impératifs économiques et techniques et en tenant compte de facteurs géographiques »* (Le Sueur, 1989). Le modèle comportemental inhérent à ce réseau de voies navigables questionne en revanche l'adaptabilité du modèle PONT, tant les discontinuités sont naturelles et nombreuses. C'est sans doute dans l'exploitation d'autres types de multigraphes qu'il faut chercher des solutions, capables de prendre en compte plusieurs modes de déplacement.

Sans doute, la prise en compte sur le temps long des réseaux de transport ne peut se

dédouaner de la fièvre routière, qui a opérer au début du XIXème siècle (Arbellot, 1990). Défini comme le siècle des routes (Reverdy, 1995), le XIXème siècle est ainsi l'occasion de se poser la question de la mise en réseau des routes sur le territoire français, que Georges Ribeill qualifie avant cela d'*« impensable »* (Lepetit, 1986), parce qu'à l'époque, tous les chemins de terre sont *« à peu près équivalents et […] ne sont dénommés routes que parce que le trafic à longue distance les emprunte »*. La littérature historique ne s'accorde pas sur la période à laquelle les routes et leur hiérarchisation font qu'il existe un système de transport routier, avec les propriétés que cela engendre. Pour Georges Reverdy, les cinquante premières années du XIXème siècle ont permis d'*« achever le réseau des grandes routes […] puis on a créé, souvent de toutes pièces, un réseau de routes départementales ; […] et jusqu'à la fin du siècle, ce sera la construction de l'immense réseau des chemins vicinaux »* (Reverdy, 1995). Ainsi, un décret de 1811 fait appliquer une numérotation systématique des routes à l'échelle nationale (Beyer, 2004). Ce développement fait l'objet de deux documents statistiques chargés de rendre compte de l'état des routes en 1824 et 1837. Ces sources, numérisées par la Bibliothèque Nationale de France, montrent à quel point le développement du réseau routier est d'abord pensé en termes de financement, et non pas dans une logique d'aménagement à l'échelle nationale (Lepetit, 1986). Pour ce dernier, l'une des principales difficultés qui fait que l'on a du mal à penser que le réseau ait été réfléchi est l'absence de hiérarchisation visible dans les premières cartographies du réseau routier. En effet, ces éléments existent mais n'ont pas été repris, attestant du déficit des ingénieurs de l'époque à envisager le maillage routier comme un réseau hiérarchisé (Lepetit, 1986). Parmi ces informations existantes, le décret de 1811 instaure trois classes de routes, de la radiale depuis Paris jusqu'aux routes d'intérêt régional, en passant par les liaisons vers des villes de province de moindre importance (Arbellot, 1990 ; Studeny, 1995).

L'étude d'un tel réseau mérite un dialogue avec le laboratoire COGIT de l'IGN, dont un des axes consiste à caractériser les évolutions territoriales à partir de la construction de bases de données historiques, en lien avec l'EHESS, à travers la vectorisation de cartes anciennes. A ce propos, des premiers contacts ont été initiés pour obtenir les minutes cartographiques de la carte d'Etat-Major, qui ont été par ailleurs mises en ligne par l'IGN dans son portail public[63]. L'objectif d'une telle entreprise est d'adapter la méthode éprouvée pour la saisie du réseau ferroviaire français. A partir d'une information vectorielle existante, il s'agit de remonter le temps selon la même méthode composite : l'information de base envisagée est celle contenue dans les données Route 500 de l'IGN. A partir de là, en fonction des minutes dont nous disposons, il s'agit de renseigner la présence ou l'absence des tronçons pour une date donnée. Il est également envisagé de qualifier la hiérarchie du réseau à partir de cette même source, dont la légende est construite de manière cohérente pour en rendre compte (Robert et Costa, 2009), ainsi qu'en s'appuyant sur des sources statistiques secondaires, capables de nous renseigner sur la qualité des infrastructures, surtout dans la première moitié du XIXème siècle.

Dès lors, l'étude combinée de plusieurs réseaux de transport pose la question de la permanence de la structure par leur superposition et leur complémentarité : dans quelle mesure la dimension stellaire du réseau ferroviaire français est-elle transposable à d'autres réseaux ? Est-il possible de dater et de contextualiser la mise en place d'une telle structure ? Dans ce cadre,

[63] http://professionnels.ign.fr/scanhisto

réinvestir la structure des voies romaines semble d'ailleurs une piste intéressante, malgré un déficit d'informations sur les réseaux de communication à l'époque médiévale *(cf. 5.3.3, p. 211)*.

Pour intégrer la dimension routière, une conséquente nouvelle phase de collecte de l'information géohistorique semble indispensable. Dans la mesure où nous avons mis à l'épreuve notre modèle conceptuel de données et démontré comment des analyses sont possibles à mener dans divers champs, le recours à cette nouvelle phase semble se justifier, parce qu'elle permet aussi d'exploiter le réseau ferroviaire sur un temps plus long, jusqu'à nos jours. Ainsi, la méthode développée pour la construction du multigraphe a déjà été mise à l'épreuve pour d'autres réseaux de transport.

Jusqu'à maintenant, l'étude du réseau ferroviaire jusqu'à nos jours ne pouvait être menée dans la mesure où le mode de transport dominant ne pouvait pas être pris en compte dans la modélisation des effets du réseau (Flonneau et Guigueno, 2009). L'objectif est de couvrir une période très large, depuis la fièvre routière qui semble avoir prise sous le Premier Empire, jusqu'à nos jours. A l'intérieur de cette croissance du réseau routier, l'histoire montre comment la dimension d'aménagement a régi le tracé des certains axes structurants, en fonction de visées militaires pour franchir les Alpes, ou dans l'esprit d'association sous la Restauration pour améliorer les nombreux tronçons lacunaires et pour permettre la construction d'ouvrages d'arts pour traverser les cours d'eau. On retrouve le même Legrand dans les grands projets routiers de l'époque, pour garantir l'entretien des routes existantes (Reverdy, 1993). La phase de décollage du réseau ferroviaire doit aussi être une période d'intérêt pour l'étude du réseau routier, qui aboutit à la déchéance de nombreuses routes, en favorisant la décentralisation de la compétence routière.

Rapidement, la situation change dans la phase de croissance du réseau routier. Le début du XXème siècle est celui des bouleversements dans la hiérarchie routière qui est nécessaire à *« l'adaptation du réseau à la voiture »* (Beyer, 2004). Pour autant, la littérature pointe un décalage entre l'apparition de cette innovation technique et une nouvelle étape du développement du réseau routier après la seconde guerre mondiale, tant le réseau déjà dense pouvait suffire jusqu'alors (Reverdy, 1995) : cette première partie de siècle s'attache alors à lutter contre la poussière et à améliorer l'infrastructure existante, sans toutefois prendre en compte *« aucune hiérarchisation des besoins, ni donc des itinéraires »*. Dans ce cadre, l'utilisation du graphe fonctionnel et de l'étude des capacités circulatoires peut apporter des clés d'analyse et d'explication dans la construction d'un véritable réseau, en évaluant les décalages entre la structure du graphe et les choix stratégiques qui sont appliqués sur le réseau en termes d'investissement et d'amélioration de conditions de circulation.

Enfin, l'étude du réseau sur le temps long permettra de mettre à profit les études éprouvées dans le cadre de la thèse, appliquées par exemple aux différentes tranches du premier plan quinquennal entre 1952 et 1956, à partir des chapitres portant sur les grands itinéraires, les autoroutes, ainsi que sur les équipements routiers des grands centres urbains et industriels. Pour les périodes ultérieures, de premières pistes ont aussi déjà été lancées dans le cadre d'une consultation de recherche dirigée par Mohamed HILAL dans le cadre du programme PUCA, à laquelle nous avons collaboré. Cette opération a permis la saisie de 72 700 kilomètres de routes entre 1970 et 2000, attestant de la présence, du type et de la largeur des routes. Cette période couvre la seconde étape du développement autoroutier et les contournements routiers croissants des pôles urbanisés (Hilal, 2012).

L'élargissement du cadre spatial nous a permis de dégager des perspectives de recherche quant au déploiement de l'innovation ferroviaire sur le Vieux Continent : des perspectives d'ensemble et comparatives sont envisagées pour les croiser avec des indicateurs de dynamisme territorial afin d'étudier le lien qu'ils entretiennent sur le temps long. Ce dernier participe aussi à l'élargissement progressif des échelles temporelles : ces dernières perspectives montrent comment le cas français revêt encore d'importants enjeux, en questionnant la permanence des réseaux sur la très longue durée et sur la succession des systèmes de transport jusqu'à nos jours. La confrontation de ces doubles dimensions spatiales et temporelles invite alors à questionner la permanence du concept de réseau, des contextes historiques jusqu'à des enjeux contemporains.

9.1.3. **Questionner la permanence du concept de réseau**

L'évolution des réseaux de transport sur le temps très long permet désormais d'évaluer dans quelle mesure *« l'idée d'une aptitude intrinsèque du réseau à croître sans souci de quelconques obstacles »* (Offner, 1993b) n'est pas si juste que cela. Ainsi, nous pouvons désormais poser la question de la permanence d'un cycle tel qu'il est décrit par Mensch *(Figure 1. 3, p. 34)*. Si Gabriel Dupuy a montré dans son schéma tendanciel que l'arrivée de l'automobile n'a pas engendré une décroissance des réseaux de transport en commun, si l'arrivée de la télévision n'a pas engendré la décroissance du téléphone, si l'arrivée de l'électricité n'a pas provoqué le recul du gaz, alors l'étude empirique de quelques réseaux de transport en particulier montre que ce phénomène de croissance n'est pas automatique. L'exemple des Etats-Unis est à ce propos très évocateur *(Figure 9. 1)*.

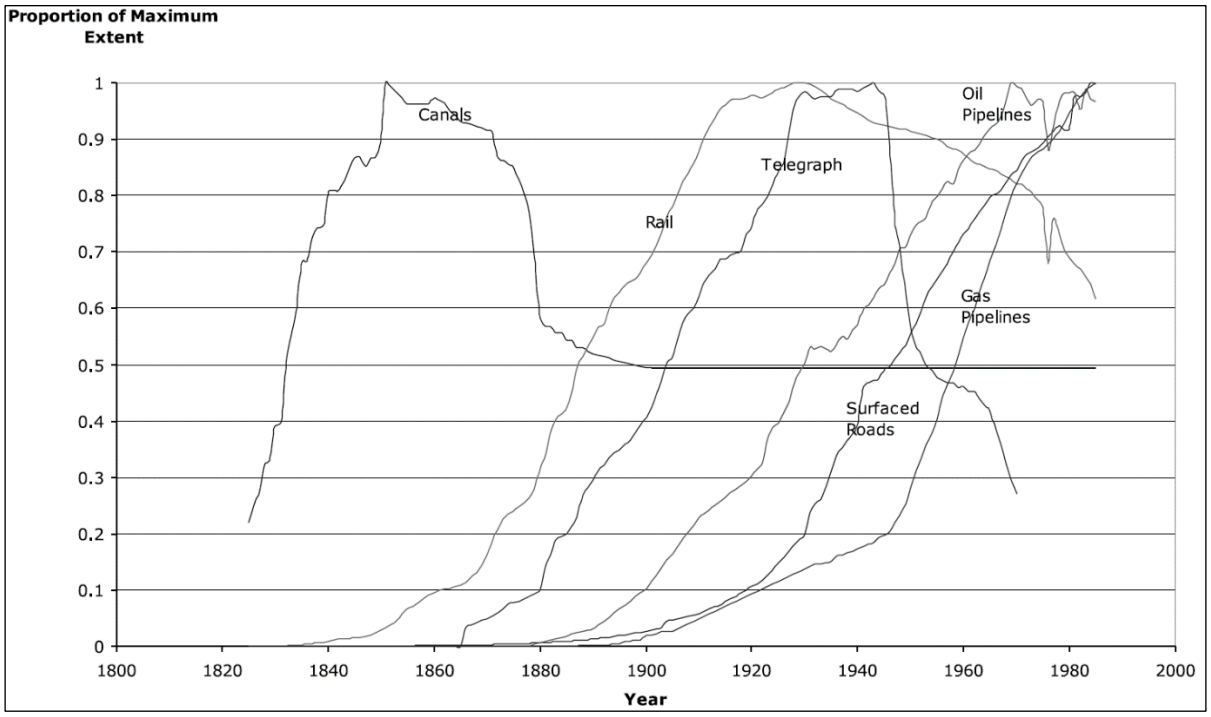


Figure 9. 1. La succession et la diffusion des réseaux aux Etats-Unis, d'après GARRISSON, LEVINSON, 2014

Dans ce cas, tous les réseaux ne sont pas logés à la même enseigne. Il est par exemple tout à fait remarquable de voir comment le réseau des canaux en Amérique a connu un apogée très éphémère, alors même que le développement du chemin de fer n'en était qu'à ses balbutiements. Par la suite, le réseau s'est stabilisé en 1900 jusqu'à nos jours. Est-ce que le réseau le plus développé était adapté à la situation existante, était-il en surcapacité ? En quoi la situation stabilisée répond-elle aux besoins de manière efficiente, équitable ? Voilà le genre de question auquel notre modèle d'évolution d'un réseau de transport doit pouvoir répondre. La rétractation de ce réseau répond-elle à la prise en compte d'un certain niveau de la structure économique ou démographique ? Le questionnement de la permanence d'un réseau permet ici de remettre en perspective différents éléments développés au cours de la thèse et de les articuler. Le modèle dynamique du réseau de transport est alors un moyen pour enrichir les connaissances quant à la diffusion d'un réseau, à sa rétraction : ses paramètres permettent de montrer dans quelle mesure le réseau participe à une concentration des flux ou à une lâcheté de sa structure, en rapport avec une situation initiale. Ce cas isolé des Etats-Unis doit nous permettre de transposer ces questions sur des territoires qui nous sont plus proches.

Dans ce domaine, la littérature sur l'évolution des réseaux s'est également penchée sur la phase de décroissance des réseaux. Alors que le raisonnement conduit au cours de la thèse a questionné l'amélioration des liens, la création de centralités, l'inverse questionne l'abandon de certains liens, la rétraction de certains nœuds. Cet aspect-là peut être selon nous étudié de plusieurs manières, en mettant à profit le plan d'expérimentation initié dans le cadre de la thèse et qui nous permet de davantage nous interroger sur le concept de vitesse différenciée (Ollivro, 2000).

Cela permet d'abord de se réapproprier le multigraphe pour alimenter les indicateurs d'accessibilité. Nous devons imaginer la structuration d'un multigraphe qui combine les différents modes de transport. Dans ce cadre, le modèle comportemental doit être adapté et une attention plus particulière au concept de rupture de charge (Stathopoulos, 1997) doit être portée, ainsi qu'au respect d'une chaîne de déplacement cohérente, dans la mesure où la vitesse autorisée sur autoroute est parfois supérieure à la vitesse moyenne d'un train. Dans cette perspective, les enseignements que nous avons tirés de l'utilisation d'un graphe triangulé nous poussent à poursuivre nos investigations par la comparaison de distances-temps théoriques et de distances-temps sur un réseau routier réel, avec des vitesses différenciées. Pour l'instant, il s'agit toujours de raisonner sur une offre potentielle de transport, estimée à partir de l'infrastructure. Nous reviendrons sur les enjeux de la prise en compte d'une qualité de la desserte dans la section suivante. Pour autant, dans la mesure où la part modale de l'automobile a pris une place croissante aux dépens du rail (Dupuy, 1999), il semble que cet aspect doit être pris en compte dans la modélisation ultérieure des effets de réseaux.

Cette étape est un préalable à l'étude des effets de réseau à partir de spécifications semblables à celles entreprises dans la thèse. Elles nous semblent en effet porter de sérieuses attentes quant à l'explication d'une perte d'intérêt d'un réseau au profit d'un autre. Par ailleurs, l'analyse comparée des gains d'accessibilité sur le temps long peut peser sur le pouvoir explicatif des dynamiques territoriales. A ce propos, de premières pistes ont été lancées dans un projet[64] traitant

[64] Nous avons pris part à ce projet dans la mesure où les temps de parcours ferroviaires sur le réseau du XIXème siècle ont été utilisés comme variables instrumentales pour caractériser les temps de parcours sur le réseau du

des retombées économiques indirectes des grandes infrastructures de transport depuis les années 70. Les chercheurs montrent que l'infrastructure permet de capter une richesse externe et une fuite du revenu local : dans leurs conclusions, ils montrent que les villes les plus dotées en aménités urbaines bénéficient d'une *« circulation invisible des richesses »* (Bahoken et al., 2016). Pour cela, ils ont utilisé la méthode des variables instrumentales : les temps de parcours sur le réseau ferroviaire du XIXème siècle que nous avons construit ont permis d'instrumenter les temps de parcours à partir des années 1970. Ce type de projet vient confirmer le positionnement de notre recherche, qui vise à la fois à diversifier l'information contenue dans l'étude des réseaux de transport, mais aussi celle qui a un pouvoir explicatif local, de sorte à ne pas surestimer l'effet du réseau dans les dynamiques territoriales, et ainsi répondre aux critiques répétées d'un manque de rigueur méthodologique (Offner, 1993a). Dans ce cadre, nous souhaitons pouvoir investiguer plus en détail l'économétrie spatiale qui nous permet de complexifier l'approche de l'interaction par des modèles spatiaux, mais aussi par des modèles locaux, tels que la *Geographically Weighted Regression* (Brunsdon, Fotheringham et Charlton, 1996).

De l'autre côté du plan d'expérimentation, une perspective vise à qualifier cette phase de décroissance selon des critères morphologiques et hiérarchiques. Les mêmes auteurs que ceux cités dans la construction de notre modèle s'y sont déjà intéressés (Xie et Levinson, 2008). Cela nous permet de poser des jalons dans le développement du modèle d'évolution d'un réseau de transport : le constat de départ est simple et montre que malgré la croissance de la littérature sur l'évolution des transports, la plupart assume une structure croissante qui n'autorise pas un lien à être pénalisé ou abandonné.

La situation initiale n'est alors plus la même. La *Figure 9. 2* montre les premières pistes conceptuelles de la construction d'un tel modèle. Alors que notre modèle de croissance suppose un réseau de projets transactionnels exhaustif, la prise en compte d'une décroissance du réseau part de l'apogée de ce dernier. A titre d'exemple, la modélisation empirique de la décroissance du réseau ferroviaire français passe par la prise en compte de son étendue maximale entre 1920 et 1930, en respectant la hiérarchie alors installée dont l'étude a fait l'objet du chapitre précédent, tout en conservant l'ensemble des autres liens qui servent de rabattement vers le réseau ferroviaire[65]. Le principal changement réside dans la définition et les paramètres à donner au module dynamique du modèle d'évolution du réseau de transport. En nous inspirant de la littérature, nous évoquons ici quelques pistes quant aux paramètres à intégrer.

Le modèle doit toujours opérer un choix d'investissement qui doit être à double sens et non systématique : il s'agit bien d'une analyse coût-bénéfice entre le coût attribué à un lien – son impédance – et le revenu qu'il peut apporter – le potentiel circulatoire qu'il renferme. L'arbitrage pour l'investissement ou le désinvestissement est le nouveau critère. Il requiert la définition d'un seuil à partir duquel nous postulons que les revenus potentiels d'un lien ne suffisent plus à justifier l'infrastructure. La littérature suggère la prise en compte de coûts fixes et de coûts variables, en fonction de la qualité et de la circulation le long d'un lien. On peut imaginer que le modèle d'investissement consiste à améliorer l'infrastructure ayant des revenus élevés, ainsi qu'à diminuer

XXIème siècle.

[65] La discussion sur la valeur de ces liens spécifiques revient à discuter la modélisation de la complémentarité des réseaux et à l'évaluation du rabattement vers le réseau ferroviaire à partir d'un réseau routier déjà hiérarchisé.

sa qualité – appréhendée par son impédance – si les revenus ne sont pas suffisants. De cette sorte, nous ne devons autoriser la modification d'un lien seulement dans la mesure où il accueille déjà une infrastructure ferroviaire. De la même manière que dans le cas du modèle développé au cours de la thèse, ce nouveau modèle doit prendre en compte un paramètre de concentration, qui traduit la capacité du réseau à privilégier la branche commune. Finalement, le principal changement réside dans la possibilité de faire décroître la vitesse d'un lien, alors que cette modification n'est pas possible sur l'ensemble des liens de la situation initiale, mais seulement sur ceux qui possèdent une infrastructure.

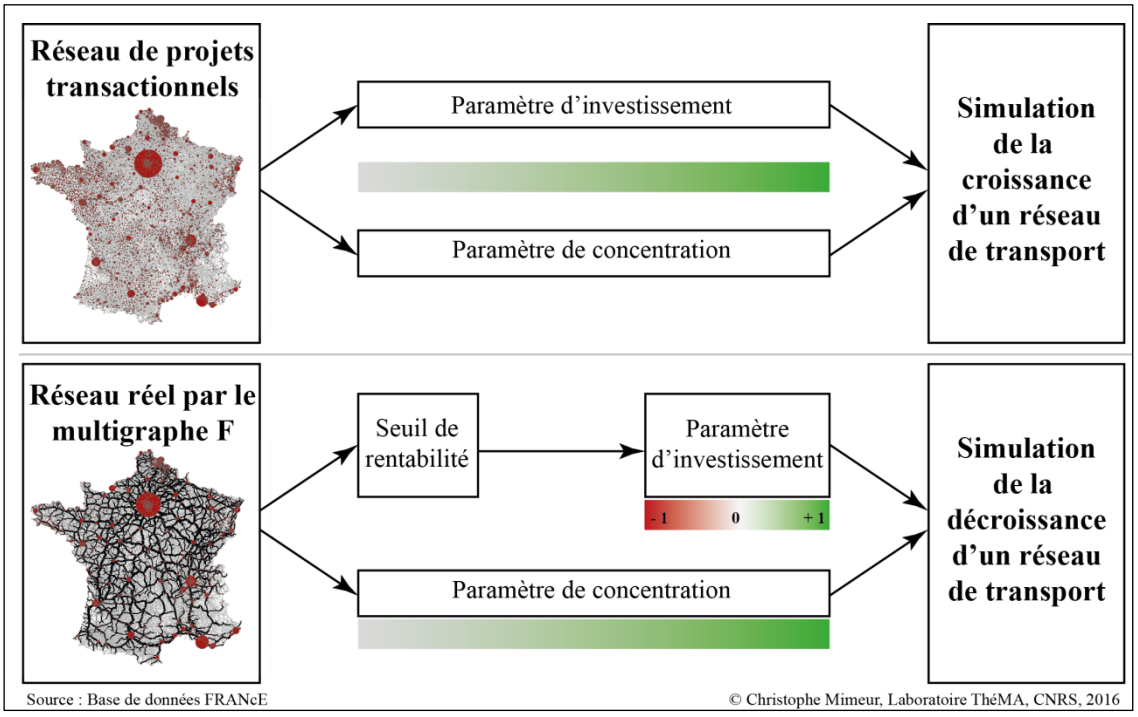


Figure 9. 2. De la croissance à la décroissance des réseaux de transport

Ces dernières réflexions incitent à replacer le rôle des aménageurs dans le dessin des grands réseaux de transport. L'ensemble des modélisations envisagées dans le plan d'expérimentation de la thèse permet de questionner les discours et les actes d'aménagements. En effet, la prise en compte d'indicateurs d'accessibilité fonctionnelle a permis de mettre en exergue des différentiels polynucléaires dans la phase de croissance du réseau ferroviaire français, en questionnant la portée des gains d'accessibilité. La dimension de panel a permis de montrer comment la hiérarchisation du réseau a un pouvoir explicatif sur l'évolution différentiée des trajectoires de populations, alors que la possession d'une infrastructure revêt encore un effet marginal qui lui est propre. Ce questionnement peut alors inciter l'aménageur à réfléchir à la portée d'une infrastructure : est-elle à envisager en termes de temps, de coût ? Notre étude montre qu'en dépit de temps de parcours plus faibles, l'effet direct du réseau n'est plus significatif dans un rayon supérieur à une heure de marche. Par ailleurs, notre modélisation vient encore justifier le recours à la mémoire du réseau, tant

l'ancienneté de la connexion a un impact positif mais non-linéaire dans l'explication de l'évolution des niveaux de population. Dès lors, l'accessibilité ne semble pas pouvoir se suffire à elle-même et démontre encore une fois la force des enjeux de contextes locaux. Des développements ultérieurs sont donc suggérés vers des focus régionaux, qui passent également vers une diversification de l'information géohistorique.

Par ailleurs, notre positionnement au cœur de la controverse sur les effets structurants vient questionner le rôle de structures initiales dans les actes de décisions d'aménagement, à travers la forme des réseaux. Le modèle d'évolution d'un réseau de transport a montré sa force à reproduire les formes radiales du réseau et ses difficultés à dessiner des transversales. A partir des règles simples, auxquelles nous avons donné des significations géohistoriques, on peut questionner les volontés des aménageurs, dans un temps plus ou moins long, face aux discours d'équité territoriale et aux volontés de rééquilibrage. Ainsi, *« les documents de planification des transports insistent depuis les années 1960 sur la nécessité de construire des infrastructures de transport circulaires. […] Le constat est assez net : depuis un demi-siècle on planifie des rocades et on réalise des radiales »* (Commenges, 2013). Dans le cadre de sa thèse, l'auteur propose deux explications possibles : l'une se rapproche de la littérature des sciences expérimentales et met en avant une forme universelle des réseaux dont l'expansion tend à la hiérarchisation et à la concentration, l'autre se rapproche de la littérature de la socio-économie des transports où l'on privilégie la rentabilité par la concentration des flux.

L'une des réponses à apporter réside sans doute dans l'articulation des deux sens de l'interaction entre réseau et population et donc de notre plan d'expérimentation. Ainsi, il nous semble que les études sur l'effet propre du réseau peuvent alimenter une complexification des règles qui établissent l'interaction spatiale de même que celles qui régissent l'investissement et la concentration. Dès lors, nous souhaitons imaginer une autre structure pour l'initialisation du modèle de croissance – puis de décroissance – de l'évolution d'un réseau de transport. Il s'agit finalement de réintroduire progressivement d'autres éléments de contexte dans l'initialisation du modèle pour évaluer leur part dans l'explication des tracés du réseau de transport réel. Ces questionnements se rapprochent d'enjeux très contemporains, dans la mesure où le nouveau découpage régional questionne la structure des réseaux ferroviaires, entre perspectives radiales et transversales, entre rentabilité et équité. L'analyse du découpage des grandes régions et de son impact sur la structure des réseaux est possible en utilisant l'ensemble des méthodes du plan d'expérimentation, sur la portée de l'effet du réseau et sur la force de la priorité radiale sur l'ensemble.

Cette première section nous a permis de voir comment le projet entrepris au cours de la thèse peut encore élargir ses échelles spatiales et temporelles. Les premières perspectives dégagées ainsi que les premiers pas entrepris dans ce sens montrent que le modèle conceptuel de données est adapté à une échelle européenne, ainsi qu'à une échelle temporelle encore plus large, tout en questionnant des enjeux de géomatique et de géohistoire. Cette nouvelle phase d'élargissement du corpus est aussi justifiée par les nombreuses perspectives d'analyse du couple entre réseau et territoire, s'inscrivant dans les grandes familles de modélisation définies dans le plan d'expérimentation. Désormais, il ne s'agit plus de se cantonner à la croissance d'un réseau mais à l'évolution comparée des réseaux, à leur complémentarité et leur concurrence ainsi qu'à leur tendance à la décroissance dans certains cas. Elle est nécessaire à l'étude de l'effet du réseau dans

la socio-économie des transports ainsi qu'à l'étude de l'abandon de certains tronçons des infrastructures, en posant un regard qui permet de questionner la forme universelle des réseaux et les discours des principaux aménageurs. Toutefois, l'ampleur d'un tel projet vient également questionner son dimensionnement et son approfondissement. De cette sorte, nous défendons dans la section suivante une meilleure place pour la géohistoire sur la scène scientifique française, par le prisme des infrastructures de projets géohistoriques et des publics visés.

9.2. L'approfondissement : méthodes, outils et publics

L'approfondissement d'un projet géohistorique s'appuie selon nous sur cinq piliers qui le différentient d'un projet historique ou d'un projet géographique (White, 2010). Ces cinq piliers ont déjà tous été évoqués au long de cette thèse et le projet tel qu'il a évolué jusqu'à aujourd'hui répond à ces cinq enjeux, de manière plus ou moins intense :

- Le projet géohistorique est collaboratif ;
- Le projet géohistorique met en avant les aspects cartographiques ;
- Le projet géohistorique s'appuie sur une utilisation intensive de l'informatique ;
- Le projet géohistorique est un projet évolutif ;
- Le projet géohistorique doit mettre en exergue l'espace.

Pour autant, l'évolution de la discipline, particulièrement dans la sphère anglo-saxonne, a d'abord mis l'accent sur les aspects quantitatifs avant de s'ouvrir progressivement et plus largement aux humanités et à la manière de les exploiter avec le numérique (Gregory et Geddes, 2014). Cette section s'attache ainsi a démontré comment notre projet géohistorique, parmi tous les autres, ne doit pas se cantonner au monde académique qu'il lui semble le plus proche mais doit encore s'ouvrir à d'autres disciplines et approches. Finalement, nous questionnons ici les enjeux d'une plus grande collaboration dans le projet, à partir de son dimensionnement et des publics auxquels il peut s'adresser.

9.2.1. **Un projet pour le partage et la valorisation du patrimoine ferroviaire**

Il a été montré au cours de la thèse comment le débat portant sur les effets structurants des infrastructures de transport dépasse la sphère académique (Bazin et al., 2011). Nous soutenons alors ici l'idée selon laquelle l'approche par les larges échelles temporelles et spatiales doit tout aussi dépasser le cadre de la recherche, qu'elle soit géographique ou historique. Pour autant, la collaboration des disciplines à l'intérieur de réseaux d'experts doit pouvoir contribuer à la préservation et la mise en valeur du patrimoine ferroviaire français. Un outil pour une patrimonialisation du réseau ferroviaire passe par le partage des données dans le monde académique et institutionnel et leur valorisation pour un public plus large.

Si la notion de patrimoine semble tirer sa légitimité du passé, il est beaucoup moins évident de la rapprocher de questionnements spatiaux (Desmichel, 2012). Défini comme l'ensemble des *« héritages matériels ou immatériels reconnus par les sociétés, afin d'être transmis aux générations futures »* (Veschambre, 2007), le patrimoine ne suscite l'intérêt de la discipline géographique qu'à la fin des années 1980. Pour autant, l'archéologie a saisi depuis plus longtemps la force d'un SIG dans les activités de protection, d'investigation et de communication des fouilles vers une société contemporaine de plus en plus ouverte (Kokalj et al., 2006). La géographie culturelle s'est emparée de la notion de patrimoine pour se confronter à la question du sens et de la valeur des lieux, parce qu'elle se retrouvait aussi au cœur du tournant spatial déjà évoqué dans les humanités. Toutefois, dans sa lecture de l'introduction du patrimoine dans la discipline géographique, Vincent Veschambre identifie une géographie historique à la marge du processus de patrimonialisation. Pour autant, la géographie s'est peu à peu emparée d'une demande sociale de connaissance et de valorisation du patrimoine (Hertzog, 2011).

Si les patrimoines culturels, naturels et religieux ont fait irruption il y a déjà longtemps, le *« patrimoine industriel »* est l'un des derniers qui a été reconnu (Edelblutte, 2008). En retard par rapport à la Grande-Bretagne, le patrimoine industriel n'est exploré en France qu'à partir des années 1970, sous l'impulsion des travaux de Maurice Daumas. La reconversion des friches industrielles dans le cadre du renouveau urbain ne place que rarement l'approche géographique au centre de ses questionnements (Edelblutte, 2008). Parmi les éléments sauvegardés au nom du patrimoine industriel, le patrimoine ferroviaire est une composante importante de l'histoire des techniques au même titre que de nombreuses autres infrastructures, comme les voies navigables, les anciennes voies romaines. Selon nous, l'engouement tardif qui lui est destiné est à lire de la même manière que celui du patrimoine industriel en général, où *« la volonté de ne pas conserver le patrimoine industriel provient aussi de la pratique politique de l'usine »* (Edelblutte, 2008). Les nombreuses vagues de fermetures nées de la concurrence routière (Ribeill, 1985a) se sont souvent soldées par le revêtement des anciennes lignes par du bêton. C'est sans doute pour cette raison qu'une partie du patrimoine ferroviaire a longtemps été oublié (Ribeill, 1993). En 2009, un état des lieux de l'inventaire du patrimoine ferroviaire a été fait à partir de deux bases de données – Mérimée et Palissy – du ministère de la Culture et la Communication (Simon, 2009). Il montre que 72 % de l'inventaire concerne le patrimoine ferroviaire bâti, entre gares et usines de matériel ferroviaire. Marqueur des conclusions sur l'abandon temporaire des friches industrielles, seulement 10 % de cet inventaire concerne les voies ferrées à proprement parler. Ainsi, la croissance de l'intérêt pour le patrimoine ferroviaire n'est pas à chercher dans les marqueurs territoriaux des tracés ferroviaires mais plutôt sur le fonctionnement d'un système ferroviaire largement influencé par un contexte historique et politique fluctuant ainsi que par une organisation entrepreneuriale et financière particulière (Caron, 1992, 1997).

Créée en 1987, l'Association pour l'histoire des chemins de fer (AHICF) – devenue Rails & Histoire[66] en 2010, rassemble des professionnels ou anciens professionnels du rail ainsi que des chercheurs, sociologues, géographes, mais majoritairement des historiens. Malgré la présence de quelques géographes dans les instances dirigeantes de l'association, on voit comment la discipline historique est largement représentée. L'histoire de l'entreprise, l'histoire sociale des « cheminots »

[66] www.ahicf.com

sont des thèmes privilégiés de l'association, de même que des comparatifs réguliers avec d'autres initiatives européennes et internationales. En 1998, l'association a consacré pourtant un colloque au patrimoine ferroviaire en donnant une large place à l'histoire du patrimoine roulant de la SNCF, des gares, et dans une moindre mesure la réhabilitation des sites ferroviaires et le rôle patrimonial des anciennes voies ferrées (Fosseyeux, 2008). Ce même auteur, à partir d'un inventaire des articles parus dans la *Revue d'Histoire des chemins de fer*, montre que 75 % des contributions traitant du patrimoine ferroviaire traitent de l'architecture des gares, du bâti et des ouvrages d'art, alors que seulement 20 % des contributions traitent des chemins de fer, secondaires essentiellement. La constitution des fonds documentaires fait partie des missions dévolues à l'association Rails & Histoire. Plus particulièrement, à travers l'organisation de conférences, colloques et séminaires, l'association pour l'histoire des chemins de fer est alors impliquée dans le champ émergent des humanités numériques en donnant une large place à l'archive orale, en partenariat avec la plate-forme *Huma-Num*. Depuis le début des années 2000, le rassemblement de telles archives a ouvert de vastes questionnements historiographiques sur les conditions qui font que *« leur recueil et leur utilisation [sont] scientifiques »* (Duclert, 2002), à travers des témoignages de cheminots pendant les guerres par exemple ou sur la condition sociale des cheminots dans la seconde moitié du XXème siècle. Parallèlement, en partenariat avec le *Cercle généalogique des cheminots* et *Rails et Mémoire*, l'association participe à la création d'un site web recensant les plaques commémoratives et monuments qui sont le symbole d'une *« appartenance des cheminots à une communauté professionnelle liée à des valeurs et des souvenirs partagés »*[67]. Il s'agit là d'un vaste patrimoine photographique. Si la cartographie des monuments précise la géolocalisation et les détails des inscriptions et personnes associées, le rattachement à une ligne historique n'est pas mentionné.

L'objectif que nous poursuivons ici est de montrer comment l'information géohistorique recueillie et complétée dans la base de données FRANcE peut cohabiter et venir enrichir la production de connaissances dans le champ plus large des humanités numériques, déjà ouvert dans le domaine de l'histoire des chemins de fer. Elle a alors une place privilégiée aux côtés de l'archive orale, textuelle ou photographique, comme porteuse d'une forte valeur historique, comme dans l'étude d'un *« paysage ferroviaire »* (Desmichel, 2012).

Dans la diffusion et la valorisation d'un patrimoine ferroviaire français, l'utilisation de larges librairies digitales et de l'outil Internet semblent avoir le meilleur potentiel pour rendre disponibles les informations géohistoriques accumulées. La constitution d'un outil de *géoweb* ou de *webmapping* semble être une piste déterminante pour répondre aux objectifs de partage et de valorisation de la donnée, et s'accommode alors d'un principe fort des humanités numériques, qui consiste à construire des cyberstructures. Pour autant, cette question semble un écueil à la mise en œuvre rapide d'un tel outil, dans une visée collaborative. En effet, dans le cadre d'une démarche d'inventaire, les aspirations collectives voire collaboratives, en phase avec le développement croissant des humanités numériques, remettent en cause l'utilisation de logiciels monopostes. Ainsi, le développement de projets SIG-H bénéficie de l'apparition des technologies réseaux qui permettent le développement des outils *« client-serveur »* (Costa, 2012). Plus qu'un simple outil SIG, les SIG *« online »* sont la *« base de la constitution de cultures techniques communes qui influent très largement sur nos pratiques de recherche en les rendant plus mutualisées »* (Costa,

[67] www.lieuxdusouvenir.org

2012). Ce passage facilite et incite en outre à la diffusion des données : la création de sites web participe aussi à l'ouverture progressive des humanités aux questionnements patrimoniaux, laissant profiter au plus grand nombre – chercheurs académiques et grand public – les données (géo)-historiques qu'elles collectent et analysent. Ces développements permettent aujourd'hui un large accès aux informations sans pour autant disposer d'un logiciel préalable (Gregory et Ell, 2007) à travers de multiples moyens de diffuser et visualiser l'information. L'intérêt réside également dans la démultiplication des possibilités de superposition d'informations de nature différente.

Dans ce cadre, la structure *TGIR Huma-Num* prévoit la mise à disposition d'un cluster web mutualisé, qui propose une multitude de langages et de bases de données[68]. Alors, *« avec ces briques de base fournies et gérées par Huma-Num, il est possible de gérer un site Web pour diffuser ses données. Des accès interactifs permettent de gérer facilement son arborescence de fichiers et sa base de données »*[69]. On lit dans cette initiative la possibilité d'adosser l'infrastructure de partage des données avec une expertise d'informaticiens et de spécialistes. Et pourtant, *« il revient à la charge de l'utilisateur dans le cadre de son programme de recherche d'assurer la maintenance technique des outils logiciels qu'il met en œuvre »*. Cette première limite nous permet d'identifier le besoin de se référer à des expertises plus locales afin de développer ce type d'outils. C'est dans cette même limite que nous identifions les difficultés de cette infrastructure à s'imposer sur la scène académique nationale, alors que de nombreuses initiatives locales demeurent à l'œuvre, dans le cadre de la Maison des Sciences de l'Homme par exemple. Pendant longtemps, les infrastructures mutualisées se sont concentrées sur les métadonnées, cadre indispensable par lequel l'insertion d'une donnée dans une infrastructure est possible et se sont limitées ainsi à une mise à disposition brute des données. Selon nous, cette entrée peut constituer un frein à la mise en œuvre aisée et rapide de telles initiatives, ainsi qu'à l'instauration d'un dialogue pluridisciplinaire. On notera par exemple que le projet de vectorisation de la carte de Cassini mobilisé dans cette thèse fait l'objet de deux partages déconnectés de la plate-forme *Huma-Num*. Un site de consultation de l'information a été développé par les chercheurs sous une licence libre[70], tandis que les données sont également accessibles sur le *Havard Dataverse*[71], alors qu'elles ont pourtant été collectées par une équipe française.

Alors, une valorisation plus systématique des données peut participer selon nous à une grande visibilité de la géohistoire dans la discipline géographique, ainsi que l'ouverture des projets à d'autres publics, dans les sphères académiques ainsi que pour le grand public. Il s'agit de deux démarches complémentaires : la première consiste à l'hébergement des données sur un site web de manière à pouvoir les partager. La seconde, qui consiste à valoriser la donnée géohistorique, s'adresse en revanche à un public moins averti : dans ce cadre, des efforts de mise en forme et d'ergonomie sont nécessaires à la valorisation du patrimoine ferroviaire français. La collaboration rapprochée entre géographes et informaticiens est requise : il nous semble que le cadre d'une Maison des Sciences de l'Homme, par les services qu'elle peut proposer, est celui à privilégier dans la mise en œuvre de telles infrastructures, parce qu'elle est également un lieu de rencontre en

[68] Parmi ceux-là, PostgreSQL, PostGIS sont déjà des langages utilisés dans la construction et la manipulation de la base de données FRANcE. D'autres sont nécessaires pour une diffusion sur le web : C++, PHP, Python, Java en sont des exemples.
[69] www.huma-num.fr/service-et-outils/diffuser
[70] www.geohistoricaldata.org
[71] https://dataverse.harvard.edu

chercheurs des sciences humaines et sociales dans la phase d'approfondissement du projet que nous défendons ici.

Dans une démarche de collecte hétérogène et diversifiée, le recours aux archives primaires et secondaires dans la phase de collecte pose la question des droits de diffusion et de propriété de données. Dans la base de données FRANcE, une information actuelle de l'IGN constitue la référence géométrique des découpages administratifs et des tracés actuels de voies ferroviaires. Dans le cadre de la politique d'ouverture des données publiques, les bases de données Route500 et GeoFla sont aujourd'hui en *« Licence Ouverte »*.

En questionnant la place des *Spatial Humanities* dans le patrimoine ferroviaire français, nous voyons les nombreuses pistes de réflexion en même temps que l'émergence de nouveaux enjeux. Dans ce domaine, valoriser l'inventaire des lignes ferroviaires sur le temps long ainsi que l'enrichissement de l'information géohistorique revêt un enjeu, traditionnel dans un SIG-H, qui entre en résonance avec le contexte d'une croissante patrimonialisation des friches industrielles et ferroviaires d'une part, et avec celui d'une ouverture de l'histoire des chemins de fer dans les humanités numériques d'autre part.

Dès lors, le recours à davantage de croisements disciplinaires nous permet d'envisager un approfondissement des techniques de collecte et d'analyse, dans une mouvance générale qui fait que la donnée qualitative s'insère de manière croissante dans les projets de recherche (Gregory et Geddes, 2014).

9.2.2. **Un projet dans la mixité des méthodes**

Les perspectives d'élargissement par les échelles et d'approfondissement par les publics et les disciplines passent par une diversification des sources, qui requièrent une évolution dans les méthodes, tant pour la collecte que pour l'analyse des nouvelles informations géohistoriques. La littérature anglo-saxonne parle volontiers de *« Mixed methods »*, qui font d'ailleurs l'objet d'une revue scientifique qui combine articles théoriques, méthodologiques et empiriques dans *Journal of Mixed Methods Research*. Elle pointe également un positionnement scientifique proche de celui adopté tout au long de la thèse (Creswell, 2013) : dans le constructivisme, l'usage des *Mixed Methods* suggère une plus grande collaboration interdisciplinaire ainsi que l'utilisation intensive de l'informatique, respectant ainsi les piliers rappelés au début de la section. La croissance de ces usages passe selon les auteurs par un dépassement du clivage entre quantitatif et qualitatif, qui vient aussi répondre aux évolutions des enjeux des SIG-H, dont les principales problématiques ont migré d'approches quantitatives vers des approches qualitatives et diversifiées. Selon nous, l'utilisation de *Mixed Methods* doit à la fois participer à l'élargissement du corpus ainsi qu'à apporter des réponses à une question de recherche plus précise que celle qui englobe un projet de SIG-H (Creswell, 2013).

Dans le premier cas, la diversité des documents et des archives doit être au service de nouvelles analyses dans notre question de recherche. Là où la cartographie, l'iconographie et la

statistique sont des sources classiques dans un projet de SIG-H, l'utilisation des méthodes mixtes permet leur diversification : depuis la source textuelle à d'autres sources plus originales. Pour en avoir une meilleure connaissance, l'automatisation de la lecture de documents archivés peut contribuer à la prise en compte d'une offre de transport potentielle qui est à la fois basée sur l'infrastructure, mais qui peut l'être aussi sur les services disponibles sur elle. Dans le cadre de la thèse, la qualité de l'infrastructure a été envisagée par sa vitesse. Dans nos perspectives, nous misons sur la prise en compte de la desserte effective de manière géohistorique : les méthodes mixtes sont à la fois au service de l'analyse et de la collecte d'informations à intégrer à la base de données FRANcE. Cette nouvelle information participe au raffinement des analyses qui ont pris corps dans le plan d'expérimentation de la thèse, en venant apporter de nouveaux indicateurs pour mesurer l'effet du réseau sur les dynamiques territoriales. Dans le cadre de ces perspectives, nous prenons l'exemple de *L'indicateur des chemins de fer et de la navigation*. La *Figure 9. 3* montre l'exemple des horaires du service hivernal de 1896[72]. D'une longueur de 147 pages, ce document renferme tous les horaires pour la France entière, avec un découpage par compagnie, ainsi que des grilles de tarification en fonction de la compagnie et de la ligne. Il existe d'ailleurs aussi des indicateurs périodiques, qui se concentrent sur certaines grandes compagnies. La prise en compte de telles informations nous renvoie à la longue et fastidieuse phase de collecte de l'information. Pourtant, dans le cadre de documents plutôt formalisés, il existe des techniques de numérisation, qui consistent en une transformation automatique d'un fichier contenant l'image en un document contenant du texte : c'est une *océrisation*. Cette phase participe à l'attribution de ces nouvelles informations dans le modèle conceptuel de données PONT. Dans le cadre de la prise en compte de la desserte ferroviaire, l'océrisation des indicateurs d'horaires de chemins de fer permettraient une indexation des toponymes, des gares dans notre cas. Alors, les caractéristiques de la desserte peuvent être ajoutées dans le module *« Network »* du MCD, à l'intérieur de l'objet ponctuel *« Gare »*. Pour chaque date où l'information est disponible, on peut alors connaître le nombre d'arrêts dans la gare, ainsi que la répartition des types de desserte, recueillis à partir d'informations d'une nouvelle boîte. Son objet est la description d'un *service*, qui peut être représenté dans le SIG par un lien, dont l'implémentation fait appel à l'horodatage déjà mobilisée auparavant. L'identification du service permet de qualifier le caractère plus ou moins exhaustif de la desserte, entre *« direct »*, *« express »* et *« omnibus »*. L'analyse automatisée du document permet aussi de prendre en compte les temps de parcours effectifs, rendant compte de la longueur des arrêts. Le lien entre la collecte, l'analyse de tels documents et l'outil SIG semble donc évident. La prise en compte de ces nouvelles informations associe le recours à une nouvelle modélisation des données dans le cadre du plan d'expérimentation. Il existe des outils développés par des géographes pour rendre compte de ce type d'informations dans des analyses et la production d'indicateurs à l'aide de la théorie des graphes (Chapelon, 1997). On peut envisager des focus régionaux et l'application de l'offre effective de transport à ces focus, alors que l'on peut tout aussi bien envisager une couverture exhaustive du territoire, à condition d'avoir recours à une longue phase de collecte et son automatisation. Ce nouveau type d'informations peut encourager au dialogue avec d'autres disciplines, autant qu'à la diffusion d'une valeur patrimoniale.

[72] Ce document a été déposé sur une plate-forme collaborative qui répertorie certains indicateurs horaires. http://wikiplm.railsdautrefois.fr. Le fonds documentaire de l'association Rails et Histoire renferme un grand nombre d'indicateurs Chaix qui ne sont toutefois pas encore numérisés.

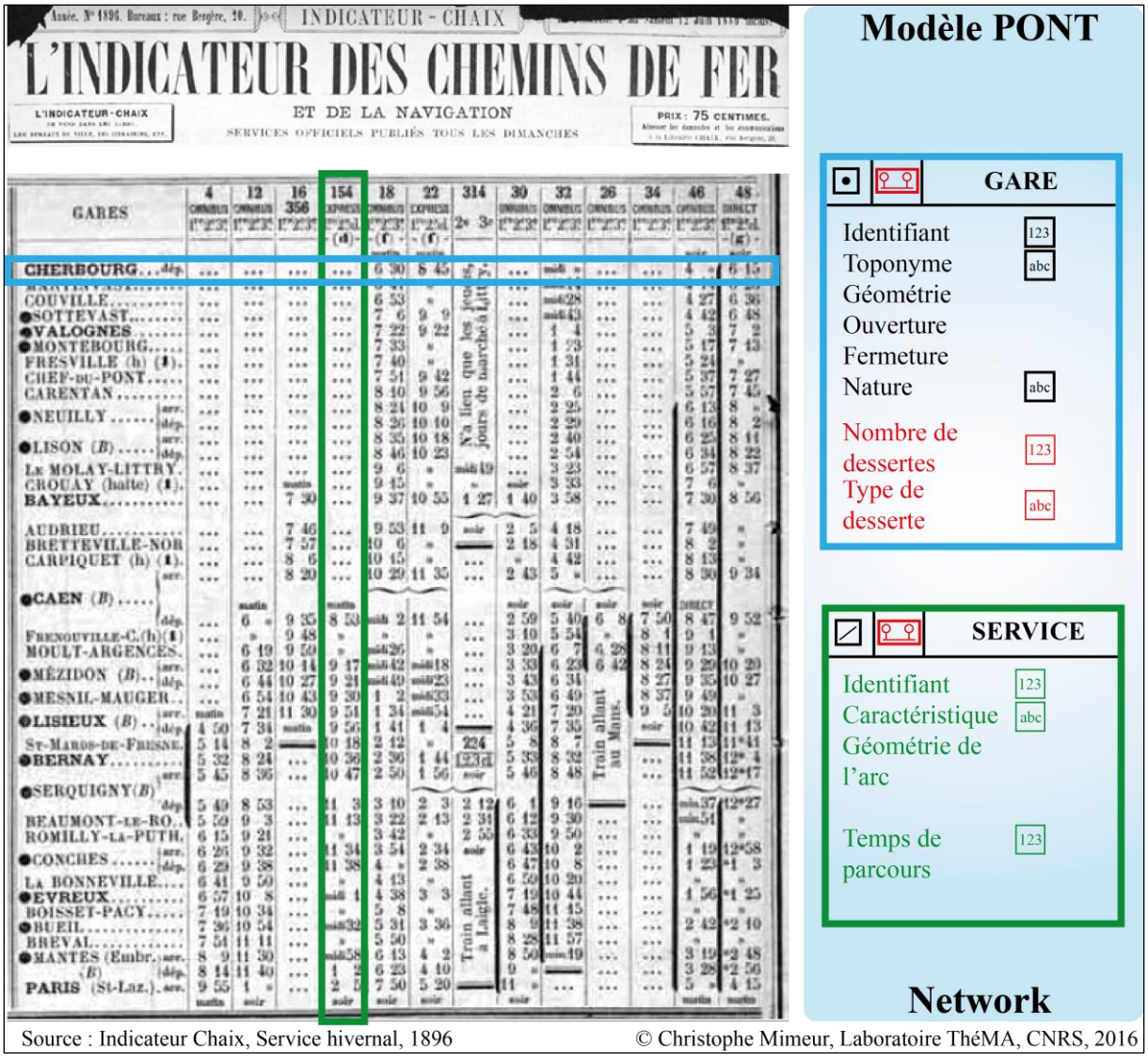


Figure 9. 3. La numérisation de la desserte pour une intégration dans le modèle PONT

L'utilisation de ce type de méthodes ne peut se limiter à la phase de collecte dans un projet de SIG-H. Il semble que l'utilisation du SIG dans une perspective croisée entre quantitatif et qualitatif participe à l'incorporation de perspectives géographiques dans l'analyse de documents autres que cartographiques (Schwartz, 2015). Cette méthode, qui consiste à analyser des sources textuelles, est connue sous le vocable de *« Text Mining »* outre-Atlantique, même si on parle de plus en plus de *« Geographical Text Analysis »* (Porter, Atkinson et Gregory, 2015). Il s'agit d' *« un ensemble de techniques appartenant au domaine de l'intelligence artificielle […] qui permettent d'extraire et de recréer de l'information à partir d'un corpus de textes (classification, analyse, tendance, etc.) »* (Grimmer et Stewart, 2013). Nous évoquons ici les perspectives d'utilisation de *Computer assisted qualitative data analysis* (CAQDA) en lien avec un SIG. Dans une optique constructiviste, il s'agit souvent d'approfondir les analyses pour des précisions à des échelles locales, ou à partir de sources originales. Dans ce cas, nous sommes aussi soumis à la disponibilité

et à la qualité de la donnée, qui nécessite le même type d'océrisation préalable à toute analyse. A partir de là, plusieurs étapes participent à l'analyse :

- *Découpage du document :* il s'agit d'isoler les parties de texte à analyser, de les nommer pour qu'ils soient ensuite interrogés, analysés, interprétés de manière isolée ou croisée ;
- *Fréquence de mots :* il s'agit la plupart du temps d'un processus itératif qui consiste à générer une liste de mots, leur fréquence, leur emplacement pour déterminer et explorer les termes importants pour l'analyse, en supprimant les mots vides de sens ou qui viendraient nuire à l'analyse ainsi qu'à la génération informatique des fréquences de mots dans le texte ;
- *Mots-clés et termes d'intérêt :* à proximité des vocables d'intérêt, il s'agit d'isoler des termes à explorer, qui sont souvent des descripteurs de ces derniers, comme « plus qu'avant », « moins que dans le passé », « augmentation », « croissance », « diminution », « déclin », … ;
- *Système codé :* il s'agit de constituer des balises pour permettre de créer des catégories et des sous-catégories, des regroupements de termes synonymes à l'intérieur. A ce stade, le chercheur contrôle l'analyse en donnant une signification et des clés d'interprétation à l'analyse automatique du texte ;
- *Associations de mots :* il s'agit d'isoler des mots dans une même structure de texte, comme une phrase ou un paragraphe, qui permet de faire des rapprochements entre les catégories mais aussi parfois d'exclure des associations qui ne sont pas fréquentes ou en dehors du champ d'intérêt ;
- *Mots et contexte (Key words in context, KWIC) :* il s'agit d'étudier le contexte de l'emploi des mots pour identifier des thèmes ou des clés d'interprétation en allant analyser des passages importants.

La succession de ces étapes aboutit finalement à une boucle d'analyse dans la mesure où l'étude du contexte au terme de l'étape KWIC participe à la reformulation des questions, la définition de nouveaux codes, et l'attachement de nouveaux commentaires au fil du texte. Par ailleurs, l'indexation systématique de toponymes participe aux perspectives de liens de ces analyses avec l'outil SIG. Il s'agit alors d'imaginer la transposition des informations ressortant de ces analyses dans un format compatible avec l'outil SIG (Bodenhamer, Corrigan et Harris, 2015 ; Gregory et Hardie, 2011).

L'utilisation du *Text Mining* répond alors aux très larges corpus d'archives qui existe et qu'il semble impossible de tous exploiter de manière fine et systématique. Ainsi, dans la mesure où elles sont numérisées, ces sources peuvent être soumises à de telles analyses, afin de produire ou de préciser des connaissances historiques. Alors, dans le cas où l'utilisation de toponymes est large, l'association de ces méthodes avec une représentation cartographique grâce à un SIG est de plus en plus employée (Schwartz, 2015). Derrière ce type d'analyse, nous avons pour perspective d'enrichir les explications quant au choix des tracés, dont on sait qu'il opère d'une combinaison de facteurs. Si le modèle d'évolution du réseau de transport permet d'apporter les éléments de réponse quant à

la prise en compte d'une structure démographique dans le choix des tracés, nous postulons que l'analyse textuelle peut participer de manière isolée à l'explication de certains choix, que ce soit aux frontières des compagnies ou lors de l'arbitrage d'axes concurrents, tant on a vu l'impact des débats politiques au Parlement pour le tracé de la ligne entre Paris et Lyon (Bavoux, 1994 ; Caron, 1997). Il existe également de nombreuses sources quant aux rapports de commissions parlementaires[73] sur l'exploitation des chemins de fer par les grandes compagnies (Caron, 1997), et de telles analyses participent à donner sens à un large corpus qui ne l'aurait pas été par ailleurs (Schwartz, 2015). D'autres sources plus locales peuvent constituer une richesse d'informations considérables quant à l'implantation d'une infrastructure au niveau local : l'analyse de la presse locale[74] et d'époque peut participer conjointement aux approches quantitatives ainsi qu'aux approches qualitatives qui ont trait aux forces en présence pour l'installation d'une voie ferroviaire. Par ailleurs, si nous avons fait le choix de faire du réseau notre objet d'étude, les données que nous possédons peuvent être utilisées au service d'une autre question de recherche, qui requiert le dialogue entre les chercheurs.

Dans ce cadre, l'ouverture croissante des logiciels dans les sphères commerciales et du libre accès arrive en appui du positionnement adopté jusqu'ici dans la thèse, à l'intersection des humanités et de l'utilisation intensive de l'informatique. Ce type d'analyse participe à la diversification des méthodes, nécessaire à l'exploration d'une question de recherche dans un projet de SIG-H et démontre encore une fois les rapprochements disciplinaires qu'il est indispensable d'établir sur la longue durée. En effet, le danger de telles initiatives est une soif invétérée de numérisation de sources, où la collecte l'emporterait une nouvelle fois sur l'analyse : la parcimonie et la collaboration des disciplines sont selon nous le gage d'une expertise du chercheur sur la question qu'il pose, les moyens qu'il se donne pour y répondre et les clés d'analyse qu'il utilise. L'exigence de la contextualisation (Pumain, 2014) s'applique donc aussi bien dans l'utilisation de l'outil SIG que dans l'utilisation des outils de *Text Mining*, dans lesquels les ambiguïtés et les complexités doivent être prises en compte, sans compromettre la recherche de structures dans les relations entre objets de recherche.

L'approfondissement des méthodes, qui requiert d'autres informations géohistoriques, la consultation d'autres types de sources et la modélisation d'autres relations semblent en totale adéquation avec l'esprit des humanités numériques, qui vise l'ouverture des questions, des données, des méthodes et des publics. Finalement, la quête d'une reconnaissance de l'histoire et la géographie de la croissance du réseau ferroviaire français dans la patrimonialisation du chemin de fer en France, ainsi que l'extension des analyses à des sources diversifiées requièrent selon nous une démarche et un dialogue plus étroits sur la scène géohistorique française, qui doivent dépasser le cadre de cette question de recherche pour être généralisés.

[73] Le site de l'Assemblée Nationale propose notamment l'ensemble des comptes rendus des débats des législatures depuis 1881.
[74] Le site Gallica renferme d'importantes archives sur la presse, tant nationale que locale ou régionale.

9.2.3. **Pour un dialogue plus intégré dans les humanités**

Les perspectives d'approfondissement mettent en exergue deux principaux enjeux : le rapprochement avec d'autres disciplines dans la diversification des données et des méthodes ; le besoin d'une infrastructure pour valoriser et partager la donnée. Cela permet selon nous d'interroger la place de la géohistoire sur la scène de la géographie française. Par cette entreprise, nous souhaitons dégager des enjeux pour rapprocher les chercheurs dans les humanités numériques, autour de questions de recherche convergentes, dans l'optique d'une diversification des approches, des données et des méthodes. Finalement, cela rejoint selon nous des perspectives qui sont à la fois didactiques et scientifiques. Des nombreux échanges que nous avons pu avoir au sein de la communauté géohistorique anglo-saxonne[75], il en ressort que le premier défi est celui de susciter l'intérêt des chercheurs, qui ne sont *a priori* pas intéressés par les SIG ou par la géographie (Gregory et Geddes, 2014). Dès lors, c'est par le croisement entre les SIG et d'autres sources que nous pouvons faire émerger l'intérêt d'historiens ou d'autres chercheurs dans les sciences humaines et sociales. C'est d'autant plus valable dans le contexte français, dans la mesure où nous avons vu comment la géohistoire connaît surtout un succès dans la géographie culturelle et dans la géopolitique (Volvey et al., 2005), loin des aspirations de l'analyse spatiale et de la Nouvelle Géographie. Mais il faut également susciter l'intérêt de l'historien, du sociologue, de l'économiste, afin de créer une communauté de *« spatial literacy »* : cet anglicisme traduit la capacité d'un chercheur à utiliser les propriétés de l'espace pour répondre à ses questionnements[76].

L'objectif principal est de fédérer des chercheurs, où le géographe doit imprimer ses capacités de dialogue et de mise en connexion de débats interdisciplinaires, fidèle à la *« science carrefour »* (Bavoux, 2009) que certains auteurs lui prêtent. Alors le projet de SIG-H doit pouvoir combiner différentes approches (Gregory et Ell, 2007) :

- Une base de données géohistorique peut être construite dans l'hypothèse qu'elle peut devenir un support pour de futures recherches ;
- Une base de données géohistorique peut être construite en réponse à une question de recherche spécifique : elle permet d'interroger des approches existantes à partir du SIG cette fois-ci ou permet de poser de nouvelles questions à partir du SIG dont l'objectif est de *« délivrer une histoire, dans une approche plus narrative »* (Gregory et Ell, 2007) ;
- Une base de données géohistorique peut être le support d'une plate-forme collaborative qui conjugue la diversité des sources, des données, ainsi que les moyens pour les partager et les valoriser.

[75] Cette communauté se réunit de manière régulière dans des colloques internationaux annuels, ou dans des *workshop* spécifiques afin d'échanger sur les programmes de recherche et l'évolution des *Spatial Humanities*.
[76] National Academies Press (2006). Executive Summary, in *Learning to Think Spatially : GIS as a Support System in the K-12 Curriculum*

La combinaison de ces trois approches permet selon nous de remettre en perspective trois piliers essentiels de la géohistoire – espace, sources et utilisation intensive de l'informatique – *(Figure 9. 4)* : la succession des trois approches énumérées ci-dessus répond selon nous à une montée en complexité progressive, dans laquelle le géographe peut accompagner le chercheur en sciences humaines et sociales ou à en fédérer pour un projet pluridisciplinaire.

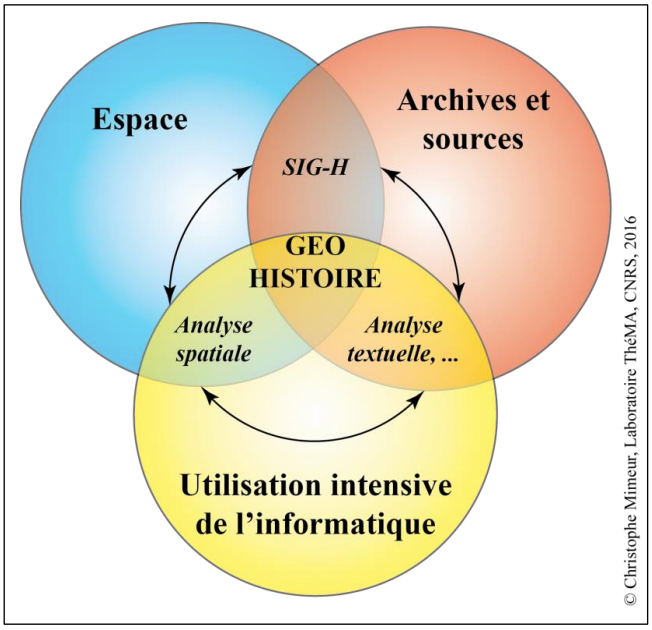


Figure 9. 4. Pour une géohistoire au cœur du dialogue dans les humanités

Selon nous, l'entrée à privilégier pour un meilleur dialogue dans les humanités n'est pas celui que nous avons adopté dans le cadre de la thèse, qui était celle d'une question spécifique pour laquelle nous avons choisi le SIG-H. Ici, l'entrée privilégiée pour retenir l'attention de l'ensemble des sciences humaines et sociales est celle de l'archive et de la source. A chaque discipline, nous pouvons identifier des sources privilégiées, dont il faut rendre intelligible les liens possibles avec le SIG. Dans l'évolution globale des SIG-H, le matériel quantitatif a été celui qui a été le premier incorporé au sein de SIG-H. Pour autant, le passage de l'un à l'autre suggère le plus souvent le recours aux sciences informatiques, au moins dans la phase de formalisation des données. A ce stade, les enjeux de visualisation sont les plus forts : la représentation cartographique, un des piliers des projets des SIG-H, peut montrer de nombreuses opportunités quant à la réponse à certaines questions ainsi qu'à la réflexion sur de nouvelles interrogations.

Ces premières pierres sont le ciment d'une collaboration entre disciplines prises deux à deux, dans laquelle le géographe tient une place centrale. Cette première prise de dialogue a selon nous des fins didactiques : par son expertise sur l'outil SIG, le géographe doit être capable de former le chercheur aux premières approches par le logiciel. Aujourd'hui, de nombreux logiciels SIG, tant commerciaux que du monde libre, proposent des visualisations et de premières analyses de manière intuitive et accompagnée. Le géographe doit en revanche se préoccuper de la manière dont l'espace est envisagé dans les descriptions et les analyses. Nous défendons ici l'idée que c'est à partir des jeux de données des chercheurs extérieurs à la géographie que le géographe doit l'amener à utiliser l'outil SIG : en utilisant une source qui lui est familière et proche, le chercheur en sciences humaines

et sociales aura sans doute davantage de facilité et d'intérêt à se familiariser avec l'outil SIG. Ces premières incursions rejoignent la première approche, dont le but est de montrer la diversité des techniques permises par l'utilisation d'un tel outil, sans toutefois oublier d'insister sur le rôle central de l'espace dans les analyses qui en sont issues, sans anticiper encore les questions de recherche auquel il pourrait permettre de répondre.

Alors, ce premier type de rapprochement peut prendre la forme d'une collaboration isolée, ou alors de formations au sein d'une Maison des Sciences de l'Homme dont l'objectif est d'initier le chercheur aux Systèmes d'Informations Géographiques, et surtout à ces capacités d'adaptabilité face aux sources et aux données qu'ils ont l'habitude de manipuler. Ces rapprochements peuvent également être l'occasion de s'emparer de problématiques patrimoniales et de les appliquer dans un contexte spatialisé : dans ce cadre, la visée se rapproche des structures initiées dans les années 1990 autour de la constitution de SIG nationaux, dont les coûts humains et financiers sont toutefois très importants (Southall, 2014). Cette première approche montre finalement comment l'information peut être davantage valorisée dans un contexte spatialisé et comment l'espace peut rendre compte de nouvelles questions et réponses. Elle peut également susciter de nouvelles aspirations quant à la collecte d'informations qui peuvent être spatialisées. Pour autant, le rôle du géographe ne doit pas s'arrêter là, et il doit selon nous s'impliquer à part entière dans le projet géohistorique, sans quoi il pourrait être relégué au statut de support technique.

Le second type d'approche est complémentaire : il s'agit plutôt de s'insérer dans des problématiques existantes, et d'étudier en quoi un nouveau regard par l'espace peut susciter de nouvelles réponses mais aussi de nouvelles questions. Selon nous, il peut alors s'agir d'adapter les approches dans un contexte spatialisé : cette approche peut naturellement arriver après une première initiation du chercheur au SIG. Les études sont la plupart du temps très précises et sur des terrains très localisés (Gregory et Ell, 2007) mais ce sont elles qui renferment selon nous un potentiel pour l'utilisation de méthodes mixtes. Alors que la première approche consiste à susciter l'intérêt du chercheur, la seconde consiste davantage à l'accompagner dans ses analyses et à développer avec lui un plan d'expérimentation qui l'amènera à la création de nouvelles connaissances géohistoriques. Pour le géographe, l'intérêt réside sans doute dans la multiplication des cas d'études mais aussi à l'application de problèmes spatiaux dans des problématiques historiques. Selon nous, cette approche participe au raffinement des approches par l'introduction progressive de l'espace qui devient une variable clé dans les interprétations. Alors, le géographe a également un rôle dans les analyses, en s'assurant que l'espace est pris en compte dans toute sa complexité.

A ce propos, une étude a été menée par des chercheurs britanniques sur le degré d'utilisation des SIG par la communauté de *« Spatial Literacy »* (Unwin, Tate et Foote, 2012). Elle a permis d'interroger 201 personnes, dont 49 % sont sur le continent américain, 42 % sur le continent européen et 3 % au Japon *(Table 9. 2)*. Il ressort de cette étude que 52 % des répondants se considèrent eux-mêmes comme des experts – qui savent utiliser la plupart des fonctions – ou experts modérés – qui savent en utiliser beaucoup, à propos de l'utilisation de logiciels de SIG installés sur leur propre machine ou sur Internet. Des chiffres semblables décrivent leur utilisation d'applications telles que Google Earth. En revanche, il est intéressant de souligner que près de 40 % des répondants ne s'estiment que débutants dans la gestion de bases de données spatiales. De ce point de vue, il semble donc que le géographe a un rôle important à jouer dans la phase de structuration

de la donnée, inhérente aux futures analyses. L'utilisation de librairies spatiales telles que celles utilisées dans le cadre de la thèse nous semble primordiale pour traiter des données hétérogènes, parfois imprécises, et surtout très diversifiées. L'utilisation de telles structures participe aussi à la diversification des méthodes, dans le sens où elles permettent de faire communiquer des informations quantitatives mais aussi qualitatives avec des attributs spatiaux.

What is your level of knowledge about the following software and Web applications?					
	Expert	Moderate	Average	Beginner	No response
Desktop or Web GIS	23.8	28.1	14.1	27.0	7.0
'Spatial databases (e.g., Oracle Spatial)'	6.0	14.1	16.3	39.1	24.5
'Digital Globes (e.g., Google Earth)'	13.4	35.5	25.3	22.0	3.8
Global Positioning Systems	10.8	20.5	24.9	30.8	13.0
What is your level of expertise with the following analytical functions in GIS?					
	Expert	Proficient	Familiar	Beginner	No response
I can create a spatial database	17.6	23.5	14.7	27.1	17.1
I can run spatial autocorrelations	8.8	11.8	20.0	35.3	24.1
I can perform relational joins	24.9	17.2	15.4	21.3	21.3
I can perform an overlay analysis	24.7	12.4	14.7	30.6	17.6
I can do cluster analysis	17.1	12.4	12.9	34.7	22.9
I can create a publishable map for a journal	32.1	16.1	16.1	22.6	13.1
I can place a map on the Internet	21.3	18.3	17.8	29.6	13.0

Table 9. 2. L'usage des SIG dans la communauté *« Spatial Literacy »*

La deuxième partie de la *Table 9. 2* est encore plus éloquente. Elle permet d'analyser dans quelle mesure le chercheur est capable d'opérer des fonctions spécifiques. Les deux cas pour lesquels les chercheurs se sentent majoritairement experts sont finalement des fonctions simples dans un outil SIG : effectuer une jointure attributaire et mettre en forme une carte pour une publication. Des compétences concernant la création de base de données spatiales ou la réalisation d'une jointure spatiale touchent un public plus restreint. Enfin, pour des fonctions plus avancées, nombre d'entre eux s'estiment encore débutants, comme pour effectuer une analyse de cluster ou effectuer une analyse d'autocorrélation. En revanche, une majorité d'entre eux s'estiment capables de publier une carte sur Internet. Cette étude vient confirmer nos hypothèses sur l'utilisation du SIG par une communauté autre que géographe, et incite aussi le géographe à occuper une place de choix dans les projets qui lient humanités et SIG. La plupart des chercheurs a ainsi développé des compétences quant à la représentation, la visualisation et le croisement de données, mais n'est pas en revanche à l'aise dans les fonctions d'analyse spatiale.

La seconde approche suggère donc une relation encore plus étroite entre le géographe et le chercheur d'une autre discipline des sciences humaines et sociales. Pour autant, dans des questionnements purement géographiques, ce même chercheur peut aussi être l'expert qui permet d'envisager de nouvelles analyses, à partir de sources textuelles qui ont été évoquées dans la sous-

section précédente. Des analyses géohistoriques quantitatives et qualitatives sont donc complémentaires.

Finalement, la troisième approche s'appuie sur les deux premières dans un objectif de productions de connaissances, de partage et de valorisation de données géohistoriques à partir de supports adaptés. Cela nous permet de replacer en perspective les enjeux dégagés tout au long de cette section sur l'approfondissement d'un projet géohistorique, en lien avec l'expérience issue du travail de la thèse. Il apparaît que l'infrastructure est l'entrée par laquelle ces questions peuvent être soulevées. La littérature aborde le plus souvent cette question par le partage des projets sur Internet (Gregory et Geddes, 2014), en couplant articles scientifiques et visualisation des données. Dans un premier temps, nous envisageons plutôt un rapprochement vers l'exploration de données spatio-temporelles à partir d'outils appartenant à la famille *Exploratory Spatial Data Analysis* (Anselin, 1999) *(cf. 4.1.3, p. 147)*. Notre propos consiste à imaginer une plate-forme de confrontation des sources et des données afin d'explorer les relations. La colonne vertébrale qui doit selon nous guider l'architecture d'un tel outil est l'information spatiale-temporelle, dans toute sa diversité et sa complexité.

Parmi les outils existants, le logiciel *GeoDa* connaît un grand succès auprès des économistes (Anselin, Syabri et Kho, 2006). Il permet l'exploitation de fonctions spatiales simples et avancées. Pour autant, notre approfondissement suggère de coupler des fonctions spatiales avancées qui peuvent également intégrer une dimension temporelle. Pour cela, nous pouvons envisager des rapprochements avec un prototype existant, dont les derniers développements se sont concentrés sur l'analyse des mobilités quotidiennes (Thévenin, 2010). Le logiciel *Geographer*, développé au sein du laboratoire ThéMA en langage Java, permet de visualiser des données spatio-temporelles à travers différents outils. L'une de ces principales caractéristiques répond aux enjeux d'un projet géohistorique, à travers l'interopérabilité. Ses capacités de dimensionnement permettent en outre la prise en charge de larges bases de données, dont l'interfaçage avec des librairies spatiales et le système de bases de données Postgres/PostGIS utilisé dans la thèse est prévu par le logiciel. Il contient ainsi déjà des fonctions classiques d'exploration de données *(Table 9. 3)*, auxquelles on peut envisager d'ajouter l'implémentation de fonctions spatiales avancées, de fonctions qui permettent de prendre en compte la dimension temporelle, de même que des fonctions économétriques qui utilisent les attributs contenus dans la base. Selon nous, ce type d'infrastructure est une clé d'entrée fondamentale dans la prise de dialogue entre le géographe et les humanités, parce que cet outil permet une prise en main rapide tout en montrant la palette de ses fonctions. D'autant plus qu'à cela s'ajoutent plusieurs modes de sélection qui s'appuient sur la dimension temporelle (Thévenin, 2010).

Pour autant, les perspectives de développement d'un tel outil sont nombreuses face aux enjeux dégagés jusqu'à maintenant, pour dépasser l'exploration pour aller vers l'analyse et la mise en œuvre d'un plan d'expérimentation. Le géographe, accompagné de l'informaticien, y tient un rôle central dans le dialogue avec les sciences humaines et sociales. Le couplage modéré soutenu dans la thèse peut permettre alors au chercheur d'avoir recours à des langages informatiques qu'il connaît déjà – SQL, R, celui qui s'applique à l'analyse textuelle – et d'explorer les analyses qui en ressortent dans un outil d'exploration spatio-temporelle. Les nombreux liens permis par l'utilisation des librairies spatiales le permettent. Si ce n'avait pas été le choix initial dans la création du logiciel,

il semble aujourd'hui possible d'intégrer à ce genre d'outil des fonctions avancées d'analyse spatiale, ainsi que d'autres fonctions avancées pour modéliser des relations. Selon nous, à travers l'exploration spatio-temporelle de l'information géohistorique, ce type d'outil peut s'ouvrir plus largement au sein des sciences humaines et sociales, en permettant de conduire une mixité de méthodes dans la lignée des enjeux identifiés.

De futurs développements doivent aussi pouvoir par la suite intégrer de nouveaux modes de visualisation de l'information spatio-temporelle, notamment à partir de nouveaux objets, comme le graphe ou le matériel issu des analyses textuelles, à envisager par l'association systématique avec des attributs spatiaux. Dans ce cadre, ces avancées répondent au double enjeu de diffusion et de valorisation de l'information. Ce type de logiciel, à partir d'une structure Java, peut être partagé par un grand nombre de chercheurs, être le lieu de convergence d'une diversité de l'information géohistorique et la scène de nombreuses analyses, tant quantitatives que qualitatives. A plus long terme, cette infrastructure peut aussi être la base d'une autre qui permet la valorisation de la donnée, à partir de sa mise à disposition vers un public plus large et sur la toile.

Catégories	Fonctions
Acquisition de données	Importation de fichier shape (point, ligne, polygone)
	Importation/Exportation de fichier texte
	Création de colonne
	Transformation de variables
Visualisation	Affichage des couches spatiales
	Carte choroplète
Graphique statistique	Histogramme
	Nuage de points
	Boîte à moustache
	Graphe en coordonnées parallèles
Sélection graphique interactive	Flèche de sélection
	Sélection par fenêtre de balayage

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Table 9. 3. Les principales fonctions du logiciel Geographer, d'après Thévenin, 2010

Les enjeux dégagés dans la phase d'approfondissement d'un projet de géohistoire démontrent encore une fois la nécessaire collaboration des chercheurs au sein des sciences humaines et sociales, où l'objectif est de voir converger des thématiques de recherche et d'y associer de manière très étroite la dimension spatiale et temporelle. Alors, la mixité des méthodes participe à la création de connaissances originales, dont le processus réflexif doit pouvoir prendre corps au sein d'une infrastructure qui permet de rassembler les données, les analyses, les chercheurs mais aussi un public plus large.

Conclusion

Ce dernier chapitre a été l'occasion de remettre en perspective la thèse dans une double dimension, celle de la question de recherche ainsi que celle du contexte scientifique dans laquelle elle prend corps. Même s'il est difficile de dissocier les deux, nous avons montré comment les nombreuses avancées en termes de SIG-H dans la collecte d'informations quantitatives sur les réseaux nous permettent aujourd'hui d'envisager l'élargissement des échelles spatiales et temporelles, qui consiste à diversifier les terrains d'études et les réseaux de transport. Par ailleurs, ces perspectives s'inscrivent dans le positionnement scientifique de la thèse, en mettant à l'épreuve le plan d'expérimentation pour la construction de nouvelles perspectives géohistoriques du lien entre réseau et territoire. A partir des enseignements tirés de la croissance du réseau ferroviaire français, caractérisé par un développement rapide qui bouleverse les vitesses de déplacement, nous pouvons désormais envisager d'autres vocables de la controverse scientifique des effets structurants, à travers la concurrence et la substitution des réseaux. L'approche historique peut être une nouvelle entrée des études contemporaines pour affiner les estimations de l'effet du réseau. Par ailleurs, l'extension des échelles questionne directement le concept de réseau, à travers ses permanences, ses ruptures ou l'inversement de ses tendances. Dans cette perspective, nous dégageons quelques pistes pour envisager la décroissance de certains réseaux au profit d'autres à partir d'une modélisation dynamique de l'évolution d'un réseau de transport. Encore une fois, ces perspectives répondent aux enjeux économiques et politiques qui gravitent autour du lien entre réseau et territoire, à travers la forme des réseaux et leurs objectifs.

Pour autant, le travail initié dans la thèse n'est pas une fin en soi, et de nouvelles approches et analyses doivent pouvoir contribuer à une analyse plus fine des sens de la causalité au sein même du processus de croissance du réseau ferroviaire français. De ce point de vue, la base de données FRANcE semble un solide terrain sur lequel nous pouvons éprouver la diversification des approches. Elle nous permet également de défendre une place de la géohistoire dans la géographie française et au cœur des humanités, à l'intersection de l'analyse spatiale, des SIG-H et d'autres types d'analyse. Elle porte en elle une capacité de dialogue avec l'ensemble des sciences humaines et sociales. Le projet géohistorique est alors bien un projet collaboratif, qui utilise de manière intensive l'informatique : il a donc besoin d'infrastructures afin de construire, d'analyser mais à plus de long terme de partager et de valoriser les données qu'il renferme. Ce processus permettra alors l'inscription de ce type de projet dans le cadre plus large des *Digital Humanities*, dans lesquelles les approches qualitatives sont croissantes et dont les liens avec les SIG le sont aussi. La géographie tient alors un statut particulier parce qu'elle permet de mettre l'espace au cœur des raisonnements et d'initier de nouvelles questions auprès de disciplines connexes, dont le recours au SIG n'est pas *a priori* pensé et optimisé. Le recours à la visualisation cartographique et à l'analyse spatiale renferment de nouvelles opportunités pour interroger des archives qui ne l'auraient plus été autrement.

CONCLUSION DE LA TROISIEME PARTIE

Cette troisième partie a permis de mettre à l'épreuve le cheminement qui avait émergé après nos réflexions sur les approches méthodologiques pour envisager le lien entre réseau et territoire, dans une perspective géohistorique *(cf. Figure 6. 20, p. 259)*. Elle vient confirmer l'enjeu de la thèse qui consiste en la valorisation des données géohistoriques de la base de données FRANcE en les confrontant aux principaux vocables de la controverse sur les effets structurants des infrastructures de transport. De cette sorte, nous alimentons les débats sur la boucle d'interaction entre réseau et territoire, en se donnant les moyens de mesurer la force de chacun des deux côtés, venant confirmer l'absence d'effets automatiques qui n'opéreraient que dans un seul sens (Offner, 2014 ; Plassard, 1977).

C'est donc en partant de l'objet graphe, largement évoqué dans la seconde partie, que le septième chapitre cherche à mesurer les effets du réseau sur les dynamiques territoriales, à travers les notions d'homogénéité, hiérarchie et hétérogénéité. Les nouveaux indicateurs d'accessibilité ont permis d'explorer les notions d'homogénéisation et de hiérarchisation progressive du réseau face à ses configurations spatiales. Il apparaît que le processus de diffusion du réseau s'accompagne d'une hiérarchisation progressive observable à toutes les échelles, mais dont les impacts sont encore plus prégnants à l'échelle régionale. Confrontée à l'évolution du système de peuplement, il apparaît que la croissance du réseau ferroviaire français suit une trajectoire parallèle et progressive avec les niveaux de population. L'évolution des temps de parcours montre des relations de structures avec les villes de plus de 2 500 habitants. L'évolution des centralités dans le réseau ferroviaire en montre quant à elle avec l'ensemble des communes françaises. Ces conclusions encouragent à l'exploration des rapports de causalité entre le réseau et l'évolution des niveaux de population.

Les techniques économétriques soulevées dans la seconde partie ont été appliquées afin de mesurer les effets du réseau prenant en compte l'hétérogénéité des communes françaises et les effets temporels. Ainsi, derrière l'amélioration des temps de parcours qui laisse à penser à une homogénéisation, l'accès à une gare reste un facteur déterminant dans les dynamiques territoriales. Cet effet se propage dans un rayon d'une heure de marche à pied. Il est d'autant plus fort que la connexion au réseau est ancienne, démontrant les effets de très long terme d'une infrastructure qui peut aussi renfermer des effets d'anticipation. Cette clé de lecture globale démontre la présence d'un rapport causal dans des temporalités longues que nos perspectives prévoient d'explorer davantage par l'économétrie spatiale et la création de typologies.

Ces efforts nous ont ensuite incités à explorer l'autre côté de la boucle d'interaction entre réseau et territoire à partir de l'étude de chaque lien dans le réseau. Ce huitième chapitre interroge les processus de sélection, de hiérarchisation et de concentration des investissements pour la croissance du réseau ferroviaire. La confrontation d'une structure vidée de choix d'investissements et d'une autre qui intègre les vitesses effectives, a permis de constater le double processus de hiérarchisation et de concentration des flux sur la longue durée, qui donne une prime aux radiales à

partir de Paris, ainsi qu'à des axes structurants le territoire des grandes compagnies. Ces choix se font au détriment des transversales, attestant d'une structure fortement hiérarchisée et en arbre, confirmée et renforcée par les décisions d'aménagement et d'investissement, qui tendent à suivre une structure démographique préexistante (Pumain, 1982). La visualisation de ces différentiels participe à l'exploration des facteurs d'explication des tracés, entre efficience économique, décision politique et contraintes topographiques.

Ces observations ont ensuite été explorées à l'aide du modèle d'évolution d'un réseau de transport. Le modèle montre sa capacité à reproduire la structure stellaire du réseau ferroviaire à partir d'une structure démographique qui ne prend en compte que les très grandes villes, démontrant une prise en charge partielle de la structure initiale dans les premières décisions du plan de Legrand et privilégiant les tracés directs.

Pour autant, la sinuosité de certains tracés participe à la prise en compte d'une structure plus maillée de la population, nous permettant de dégager des différentiels entre est et ouest de la France, en fonction des contextes. En revanche, il semble que le modèle proposé tend à favoriser les radiales au détriment des transversales : le modèle, prenant en compte des critères d'efficacité, privilégie la connexité à la connectivité. Une démarche par les potentiels nous a permis de dégager des sous-systèmes à l'intérieur du réseau, préfigurant des découpages plus régionaux. Enfin, des approches plus localisées nous ont permis de montrer comment la densité du réseau est bien reproduite par le modèle, même si les décisions d'aménagement ont tenté de les rééquilibrer. Par ailleurs, la prise en compte de décisions exogènes intégrées dans le modèle d'évolution a montré leur force dans le dessin d'un réseau à l'échelle d'une grande compagnie, en faisant fi d'un tracé qui aura été plus efficient d'un point de vue démographique et économique.

Finalement, ces résultats nous ont permis d'évoquer dans le neuvième chapitre des perspectives de recherche au-delà du travail de cette thèse. La mise en œuvre du plan d'expérimentation dans cette troisième partie nous conforte à chercher à élargir le cadre spatial et le cadre temporel de l'approche, par une étude européenne intégrée ou la confrontation de situations nationales. L'objectif est de questionner l'universalité du concept de réseau face aux contextes dans lesquels il naît, croît et parfois tend à décroître, face à la concurrence et la substitution des autres réseaux. Plus largement, ces investigations tendent également à vouloir mobiliser d'autres chercheurs dans les sciences humaines et sociales, en imaginant une scène de dialogue favorisant la diversification des sources, des méthodes et des publics.

CONCLUSION GENERALE

Cette thèse s'inscrit sans conteste dans la littérature portant sur les interactions entre réseau et territoire, dont les réminiscences et les permanences ont été démontrées. Dans le cadre de ce travail, l'élargissement des échelles tant spatiales que temporelles participe à une démarche de modélisation géohistorique, dont il convient de rappeler les différentes étapes et les différents apports qui ont émergé tout au long de cette recherche.

Notre démarche compréhensive aboutit à la conceptualisation du terme de réseau. L'émergence d'une théorie des réseaux nous permet de replacer les principales caractéristiques de la doctrine saint-simonienne ainsi que les principaux vocables qui caractérisent les interactions entre réseau et territoire, depuis maintenant deux siècles. Ce travail démontre que la profondeur temporelle ne peut être négligée pour dégager des structures dans ces interactions, en mobilisant une large base de données spatio-temporelles. La complémentarité entre question de recherche et moyens d'analyses nous fait avoir recours à une démarche globale constructiviste et abductive, où chacune des approches procèdent de ses propres mécanismes inductifs ou hypothético-déductifs.

La démarche modélisatrice permet la construction d'un instrument, qui mobilise l'archive jusqu'à la constitution d'un plan d'expérimentation. Cette démarche a montré comment le formalisme des graphes permet de pallier une information imprécise et souvent lacunaire à de larges échelles. La construction d'un graphe théorique assorti d'un graphe ferroviaire nous permet d'explorer les interactions sur l'ensemble du territoire. Par ailleurs, cette méthode est susceptible d'être adaptée à d'autres réseaux de transport, à d'autres époques : il s'agit d'une démarche généralisable qui permet la multiplication des indicateurs décrivant la structure d'un réseau et ses performances, en s'appuyant sur le modèle comportemental des déplacements de l'époque.

Pour répondre à notre question de recherche, le premier sens de la boucle d'interaction a interrogé un modèle analytique pour caractériser les effets du réseau, dans des dimensions spatiale et temporelle. L'effet d'une desserte ferroviaire a un effet direct sur les dynamiques territoriales, dont la force d'attraction se propage dans un rayon d'une heure de marche à pied. Cette force est fonction de la centralité dans le réseau mais est aussi cumulative et non-linéaire sur le temps long. Ces résultats démontrent la complexité d'effets croisés dans de multiples dimensions, démontrant que derrière les volontés d'homogénéisation, la croissante hiérarchie du réseau et l'inertie du temps long aboutissent à une hétérogénéisation des dynamiques.

Le second sens de la boucle d'interaction a interrogé quant à lui un modèle exploratoire pour caractériser les règles élémentaires qui ont régi la construction du réseau. La hiérarchisation du réseau et l'émergence d'inégalités dans la structure du réseau apparaissent alors nécessaire à sa plus large diffusion. Le modèle reproduit les règles d'une efficience économique par une structure stellaire qui privilégie les radiales au détriment des transversales, démontrant que de tels axes répondent davantage à des enjeux d'aménagement qu'à des enjeux d'exploitation. La complexification progressive des interprétations permet d'enrichir les analyses à partir du réseau simulé, de même qu'à réintroduire des facteurs exogènes dans le modèle.

Ce bilan montre comment les différents degrés de modélisation participent tous à une meilleure compréhension des interactions réseau et territoire à de larges échelles. Cette thèse montre aussi le potentiel de larges bases de données spatio-temporelles construites au prix d'une phase fastidieuse de collecte. Surtout, ce travail démontre les multiples dimensions dans lesquelles nous pouvons inscrire les analyses jusqu'à la valorisation des données. D'ailleurs, les résultats produits participent au dégagement de nouvelles perspectives de recherche, vers d'autres pistes de développement, qui concilient l'ensemble de la démarche sur le temps, de la formalisation des questions jusqu'à la conduite des analyses en passant par la structuration géohistorique, comme une étape cruciale qui conditionne la pertinence des premières et la véracité des secondes, tout en apportant des clés pour des questionnements plus contemporains. Plutôt que de revenir sur des perspectives déjà abordées, nous revenons dans les prochaines lignes sur le cheminement plus global qui a conduit ce travail de recherche, à travers nos lectures, nos rencontres ainsi que notre expérience initiée tout au long de ces trois dernières années voire davantage, défi tout aussi académique, que didactique et personnel.

Une trajectoire de l'accumulation à la capitalisation

Notre incursion dans cette thématique de recherche est à attribuer à une convergence d'intérêt et d'attraits. Les premiers s'appuient sur un solide attrait pour la dimension historique, doublés d'une formation dans le premier cycle universitaire. Les seconds se forgent dans notre formation en géographie, dans laquelle nous portons l'espace comme une variable significative dans des perspectives d'aménagement et de politiques de transport. Cette dimension s'est très tôt concrétisée comme une expérience de recherche, au contact des spécialistes des interactions entre transport et territoire. Les troisièmes résident enfin dans notre fortuite rencontre avec les travaux de collecte du réseau ferroviaire français et des populations, au sein du laboratoire ThéMA et dans le cadre du projet européen déjà évoqué. Ce projet de thèse devait donc s'inscrire naturellement aux confins des sciences géographiques et historiques.

Tout aussi naturellement, une telle entreprise nous a plutôt orienté vers le recours à l'archive. Cette démarche visait plutôt l'accumulation des données pour appréhender les relations entre réseau et territoire : la quête d'une réalité historique est ainsi plutôt celle de l'historien, dans laquelle l'exactitude de la source est la seule qui permet l'interprétation valide dans le seul contexte duquel elle est extraite. Dans cette optique, force est de constater qu'aujourd'hui, ces démarches initient davantage d'historiens qu'elles ne suscitent l'intérêt du géographe (Bodenhamer, Corrigan et Harris, 2010 ; Gregory et Ell, 2007), et cela est d'autant plus vrai dans la géographie française. Nous étions ainsi plutôt confortés dans cette approche par notre inscription dans les réseaux géohistoriques européens : elle devait recourir à l'exhaustivité de la donnée. Elle est pourtant porteuse d'une ambivalence, qui remet selon nous en cause le travail de recherche du géographe : l'archive est hétérogène, rare, porteuse d'imprécision (Knowles, 2005) tandis que notre question de recherche vise à s'affranchir de la multitude des monographies et des études de cas pour profiter de

la combinaison de larges échelles spatiales et temporelles pour mieux appréhender l'effet du réseau sur le territoire. Selon nous, l'un des principaux freins à de telles initiatives portées par le géographe tient à la nature de la donnée, qui fait du recours à l'archive un puissant obstacle à ces projets, et qui consolide la très timide place faite à la géohistoire en France.

Définitivement, la donnée issue de tels projets a un statut particulier. Mais nos questionnements ont muté de la place de l'archive vers la place de la donnée. Elle est déjà étudiée par l'historien, mais souvent à partir de fonctions spatiales qui restent simples. Elle est aussi étudiée par le géomaticien, qui s'est emparé de la dimension temporelle dans le SIG depuis déjà longtemps, mais dont les modes de représentation de cette information restent encore en retrait par rapport aux avancées de la recherche informatique par ailleurs (Thévenin, 2010). Résolument, notre trajectoire de recherche nous a conduit à l'interface de ces deux préoccupations : ce n'est plus la donnée face à l'archive, ni la donnée face au SIG qui nous importaient, mais la donnée issue de l'archive face à notre question de recherche. Là est le raisonnement constructiviste et abductif. Alors, si les phases de collecte et de formalisation de la donnée sont nécessaires, notre questionnement s'est tourné vers le nouvel enjeu des *Spatial Humanities*, centré sur les approches qui conduisent aux analyses, où seule la géographie, même accompagnée de l'histoire ne peut suffire. Il s'agit désormais de capitaliser l'information qui existe pour l'enrichir.

Aussi naturellement que nous nous sommes appuyés sur l'expertise de l'historien pour identifier les sources nécessaires à la construction de la base de données spatio-temporelle, nous devions nous rapprocher d'autres disciplines, nous projetant non seulement sur la scène des humanités numériques, mais surtout sur celles des interactions entre transport et territoire. Notre trajectoire nous permettait à nouveau d'échanger avec des chercheurs issus de la socio-économie des transports, en même temps qu'elle venait nous assurer du statut de notre questionnement dans ceux plus contemporains des effets des infrastructures de transport. Parce que nous n'étudions pas un phénomène révolu dans le passé, cette double projection est sans nulle doute à l'origine de l'affirmation de notre positionnement au sein duquel le géographe a une place de choix.

Le géographe, et l'informaticien, l'économiste, l'historien…

L'évolution de notre trajectoire n'est pas synonyme de moins d'histoire, elle permet surtout plus d'espace, tant pour la géographie que pour d'autres types d'approches, mais toujours par un regard spatialisé. Pour autant, la visée reste celle d'étudier les doubles effets temporels et spatiaux qui touchent l'interaction entre réseau et territoire. Cette conciliation ne devait pas prendre corps dans un outil intégré, parce qu'il compromet selon nous l'adhésion des autres disciplines. Notre instrument de recherche devait se trouver à l'interface entre différentes fonctionnalités, outils et approches qui mêlent l'espace et le temps. Il fait de la donnée un matériau empirique au service de théories d'arrière-plan. Il peut aussi faire de la donnée une nouvelle information destinée à de nouvelles interprétations. Si l'outil se trouve à l'interface, alors nous nous y trouvons également,

par l'élargissement de notre réseau de recherche. Pour autant, ce dialogue croissant avec d'autres disciplines ne devait pas être segmenté : le danger est la discussion entre deux disciplines de manière isolée. Or notre inscription dans les humanités numériques nous pousse à des collaborations qui doivent également se nourrir entre elles : nous y avons joué le rôle de coordination et d'impulsion dans le cadre de la thèse. De ce point de vue, nous défendons l'idée que le recours à l'interdisciplinarité en géographie renforce la vigilance que nous devons conserver voire renforcer quant à la construction d'un dispositif méthodologique cohérent, mais qui se justifie dans le cadre de la combinaison des échelles spatiales et temporelles. Il implique en revanche un net élargissement de nos compétences.

Cette cohérence s'est d'abord trouvée confortée dans une tendance de la géographie à se tourner vers l'informaticien de manière croissante depuis les années 1990. Notre ancrage dans la géographie théorique et quantitative montre comment l'étude des dynamiques territoriales fait appel à l'analyse spatiale (Pumain et Saint-Julien, 2010), et s'est emparée depuis longtemps des capacités informatiques pour exploiter des outils nécessaires à la réponse à ses questions sur la scène de la *Geocomputation*. Par la granularité imposée par notre question et notre approche, le recours à la puissance de l'ordinateur est indispensable pour modéliser l'interaction spatiale, facteur sur lequel nous mettons au centre des interactions entre transport et territoire. A partir d'une structure générique pour héberger les données, confrontation statistique, modélisation économétrique et modèle de simulation ont pris corps dans un espace commun au sein duquel chacune des approches a su puiser les informations nécessaires.

Par l'apprentissage de différents langages informatiques, nous nous sommes mis en capacité de dialoguer avec l'informaticien, l'économiste mais aussi l'historien. Le premier cherche à maximiser l'automatisation des procédures pour davantage de performances, le second s'attache à la structuration des données et la traduction du langage économétrique dans un logiciel, le troisième est davantage partisan de son incursion dans des logiciels déjà existants, et faciles d'accès. Dans le cadre de la thèse, nous nous sommes positionnés en faveur d'une automatisation des procédures pour enrichir la donnée déjà existante. La frontière entre les sciences de l'informatique et de la géographie peut alors paraître poreuse quand nous avons fait le choix du formalisme des graphes : le graphe du mathématicien et de l'informaticien est proche du graphe spatial, mais le langage de sa formalisation diffère parfois, alors qu'il s'avère un formidable objet pour prendre en compte une richesse de l'information dans des capacités qui restent raisonnables. Il s'est montré indispensable pour explorer la simulation d'un réseau de transport, à partir de situations théoriques, et d'adapter le code produit à des fins empiriques. Cette première interface permet au géographe de tester ses hypothèses, de les adapter à différentes structures qui lui sont familières, puis de les confronter à la réalité du terrain. Cette croissance dans la modélisation n'est pas nouvelle, mais la collaboration avec d'autres disciplines permet l'élargissement du dimensionnement de telles entreprises : c'est ce que nous avons porté dans le cadre de la thèse.

De manière complémentaire, le rapprochement vers la socio-économie des transports a mis à l'épreuve la permanence des questionnements sur les interactions entre réseau et territoire. Notre travail contribue aux réflexions méthodologiques sur l'évaluation d'un effet du réseau, qui s'appuie sur l'interaction spatiale initiée avec l'informaticien, mais aussi sur des techniques économétriques qui rendent compte d'effets spatiaux couplés à des effets temporels. Ces nouveaux croisements ont enrichi nos questionnements, nos compétences. En cohérence avec l'interfaçage des outils, approches et données, cette deuxième voie nous a permis de valider des hypothèses et des parti-pris

quant aux choix de la structuration des données, qu'il conviendra de valider encore sur d'autres terrains et à d'autres époques.

La boucle peut-elle se boucler ? En tout cas, elle démontre que le passage de l'accumulation de données à leur capitalisation et valorisation ne se fait pas au détriment de l'histoire. Alors l'historien peut continuer d'y être associé, à condition que la dimension spatiale reste au cœur du raisonnement, dans lequel nous nous inscrivons, en tant que géographe, à l'interface entre les disciplines, les outils, les approches et les données. Les choix que nous avons effectués dans la thèse nous permettent aujourd'hui d'envisager le passage d'une donnée quantitative à une donnée qualitative, sans oublier la première. Ces choix continuent d'alimenter le dialogue entre le géomaticien, l'informaticien, l'économiste et l'historien : nous défendons alors encore davantage le rôle pivot du géographe dans les *Spatial Humanities,* dans lesquels il doit se trouver encore plus légitime que dans les timides expériences de la géohistoire en France. Le rapprochement avec l'association *Rails & Histoire* démontre le passage d'une source classique, iconographique et textuelle mais des sources plus innovantes, comme l'archive orale, ou l'information spatiale, sur laquelle nous misons pour davantage mettre en valeur le patrimoine ferroviaire français sur le long terme.

Notre expérience de recherche, qui a élargi le cercle des disciplines au fur et à mesure de notre cheminement, nous conduit à fédérer les chercheurs en sciences humaines et sociales, dont nous posons l'hypothèse que leur matériau de recherche ont de potentielles nouvelles connaissances à délivrer quand il est confronté à la dimension spatiale et quand nous promouvons le partage et la valorisation de ces données. Par notre rapprochement déjà forgé avec les sciences de l'informatique, nous défendons l'idée que les *Spatial Humanities* doivent être le vecteur d'un élargissement de l'audience, où les données n'ont pas vocation à rester dans les mains du chercheur qui les a collectées, mais plutôt à être partagées auprès d'une audience académique autant qu'auprès d'un public plus large.

Le géographe, un vecteur pour penser le réseau par l'interaction spatiale

Applicable tant dans l'étude des interactions entre transport et territoire que dans la science géographique dans ce qu'elle a de plus large, notre trajectoire de recherche nous a permis de donner un meilleur statut à la donnée, en la sanctionnant systématiquement par une évaluation des critères de précision et de pertinence qui ont permis sa mesure (Desrosières, 2010), de sa compréhension à sa structuration. C'est selon nous une condition nécessaire pour susciter l'intérêt du chercheur en sciences humaines et sociales, parce qu'un certain nombre d'entre eux reste suspicieux quant au statut de la donnée et à l'approche quantitative.

Notre démarche peut alors entrer en résonance avec des perspectives didactiques. Notre expérience dans l'enseignement nous a permis de nous imprégner des évolutions de la discipline, à travers l'utilisation d'outils diversifiés, au croisement de la géomatique et de la statistique. La démarche de modélisation croissante adoptée au cours de la thèse est selon nous celle qui peut dicter nos pratiques d'enseignements, par la compréhension des objets et des concepts, l'évaluation de leurs mesures respectives, afin d'envisager les modalités qui permettront de les faire entrer en interaction d'un point de vue spatial, mais aussi pourquoi pas dans des perspectives temporelles. Parmi ces dispositifs, nous misons sur la force du *« souvenir statistique »* croissant depuis la stabilisation des régimes politiques, en France comme sur d'autres terrains

Selon nous, l'interaction spatiale doit être portée par le géographe. Elle est pour nous une clé d'ancrage déterminante dans notre discipline pour diversifier les approches, les méthodes, alors que cela passe par la mobilisation de divers outils. Alors, l'espace reste au cœur de nos réflexions et l'instrument de recherche construit pendant la thèse prend en compte les deux métriques différentes engendrées par l'invention de la vitesse (Lévy et Lussault, 2003 ; Studeny, 1995). Sur le temps long, la combinaison des échelles participe à une meilleure compréhension du rapport à l'espace entretenu par l'Homme, au gré des innovations, qui révolutionnent successivement le monde des transports, sur le temps long. L'interaction spatiale peut aussi être mobilisée dans l'introduction de nouvelles métriques, tant on voit que les moyens de communication se sont diversifiés par l'apparition des réseaux immatériels et virtuels, dont la dimension scalaire n'est toutefois pas vidée de son sens.

Cette réflexion sur l'interaction spatiale doit aussi à notre sens guider les questionnements contemporains qui interrogent sans cesse les rapports entre transport et territoire, encore davantage parce que les projets d'infrastructures nécessitent des investissements publics de plus en plus importants. Par conséquent, la puissance publique se trouve au centre de nombreuses procédures complexes, à travers les dossiers de maîtrise d'ouvrages, l'évaluation des infrastructures, les débats publics. Alors l'approche que nous défendons dans cette thèse participe aux réflexions méthodologiques qui prennent en compte de manière croissante la nature de l'infrastructure et qui sont devenus indispensables dans l'évaluation des politiques publiques. La diversification de l'information, de même que la place croissante dévolue à l'information qualitative participe selon nous à une meilleure compréhension et intégration des contextes locaux dans lesquels l'infrastructure s'implante, et qui permettent de mêler les effets spatiaux et les effets temporels. Parmi ces réflexions, l'utilisation d'une large base de données spatio-temporelles paraît un puissant allié économétrique dans la mise en œuvre d'estimations par la méthode des variables instrumentales (Bahoken et al., 2016). Elle est aussi la scène sur laquelle nous devons imaginer l'ensemble des modalités de l'interaction des individus, des acteurs, des territoires d'un point de vue spatial, à partir de la diversification des données, par leur nature, leur couverture, leur richesse et leur limite.

Dans des perspectives tout aussi contemporaines, l'inscription de cette recherche dans la *Geocomputation* participe aux réflexions sur la modélisation dynamique des interactions entre transport et territoire. Malgré un vocabulaire commun avec les démarches de la modélisation classique des interactions entre utilisation du sol et transport, notre approche privilégie en revanche le développement endogène d'un réseau de transport et peut participer à l'analyse d'interactions

dont les perspectives peuvent aussi ajouter graduellement des facteurs historiques, politiques et sociaux. La nature de ces informations est alors différente : elle participe à l'explication de situations passées, et peuvent permettre d'envisager la force de certains facteurs dans des situations présentes et futures. A une échelle intermédiaire, le redécoupage des grandes régions questionne aujourd'hui la stratégie infrastructurelle de ces nouvelles institutions, au regard des pertes de parts de marché du ferroviaire au profit d'autres modes de transport. Même si le rail reste un puissant vecteur d'aménagement, la question du déclassement et de l'arrêt de l'exploitation de certaines lignes est plus que jamais au cœur des problématiques actuelles.

Alors, la multiplicité des pistes de développement montre qu'au-delà de besoins initiaux, s'ouvrent *« des potentiels de partage et de réutilisation jusqu'ici inenvisageables »* (Costa, 2012). Ces démarches favorisent l'interdisciplinarité dans les humanités, faisant prendre conscience de l'importance du facteur spatial couplé à la force du temps long dans les analyses des interactions entre réseau et territoire : dans ces conditions, de la constitution de bases de données spatio-temporelles à leur traitement, le géographe tient une place centrale.

BIBLIOGRAPHIE

ADAMATZKY Andrew, **MARTÍNEZ** Genaro J., **CHAPA-VERGARA** Sergio V., **ASOMOZA-PALACIO** René, **STEPHENS** Christopher R., 2011, « Approximating Mexican highways with slime mould », *Natural Computing*, 10(3), p. 1195–1214.

ADAMS John S., 1970, « Residential Structure of Midwestern Cities », *Annals of the Association of American Geographers*, 60(1), p. 37‑62.

ADRIENKO N., **ADRIENKO** G., 2011, « Spatial Generalization and Aggregation of Massive Movement Data », *IEEE Transactions on Visualization and Computer Graphics*, 17(2), p. 205‑219.

ALBERT Réka, **BARABÁSI** Albert-László, 2000, « Topology of evolving networks: local events and universality », *Physical review letters*, 85(24), p. 5234.

ALVAREZ Eduard, **FRANCH** Xavi, **MARTÍ-HENNEBERG** Jordi, 2013, « Evolution of the Territorial Coverage of the Railway Network and its Influence on Population Growth: The Case of England and Wales, 1871–1931 », *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(3), p. 175‑191.

ANDRIENKO Gennady L., **ANDRIENKO** Natalia V., 1999, « Interactive maps for visual data exploration », *International Journal of Geographical Information Science*, 13(4), p. 355‑374.

ANSELIN Luc, 1995, « Local Indicators of Spatial Association—LISA », *Geographical Analysis*, 27(2), p. 93–115.

ANSELIN Luc, 1999, « Interactive techniques and exploratory spatial data analysis », *Geographical Information Systems: Principles, Techniques, Management and Applications, eds., P. Longley, M. Goodchild, D. Maguire, and D. Rhind. Cambridge: Geoinformation Int*.

ANSELIN Luc, **SYABRI** Ibnu, **KHO** Youngihn, 2006, « GeoDa: an introduction to spatial data analysis », *Geographical analysis*, 38(1), p. 5–22.

ANTONI Jean-Philippe, **KLEIN** Olivier, **MOISY** Stéphane, 2004, « Cartographie interactive et multimédia : vers une aide à la réflexion géographique », *Cybergeo : European Journal of Geography*.

APPARICIO Philippe, **SEGUIN** Anne-Marie, **LELOUP** Xavier, 2007, « Modélisation spatiale de la pauvretéà Montréal: apport méthodologique de la régression géographiquement pondérée », *Canadian Geographer / Le Géographe canadien*, 51(4), p. 412–427.

ARBELLOT Guy, 1973, « La grande mutation des routes de France au milieu du XVIIIe siècle », *Annales. Histoire, Sciences Sociales*, 28(3), p. 765‑791.

ARBELLOT Guy, 1990, « Les problèmes de la route française à l'entrée du XIXe siècle », *Histoire, économie et société*, 9(1), p. 9‑17.

ARCAND Jean-Louis, **LABART** Kelly, 2009, « Santé et salaires : une estimation en variables

instrumentales sur un panel de travailleurs chinois », *Economie & prévision*, 186, p. 89‑100.

ASCHAUER David Alan, 1989, « Is public expenditure productive? », *Journal of Monetary Economic*, 23(2), p. 177‑200.

ATACK Jeremy, 2013, « On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited », *The Journal of Economic History*, 73(2), p. 313–338.

ATACK Jeremy, **BATEMAN** Fred, **HAINES** Michael, **MARGO** Robert A., 2010, « Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850–1860 », *Social Science History*, 34(2), p. 171‑197.

AUPHAN Etienne, 1989, *Obsolescence ou renaissance des réseaux ferrés pour le transport des voyageurs en Europe occidentale?(France, Grande-Bretagne, Allemagne fédérale)*, Ph.D. thesis, Marseille, Aix-Marseille II, 1491 p.

AXHAUSEN Kay, 2008, « Accessibility: long-term perspectives »,.

BAHOKEN Françoise, **KONING** Martin, **MIMEUR** Christophe, **OLARTE-BACARES** Carlos, **THEVENIN** Thomas, 2016, « Les temps de parcours interurbains en France : Une analyse géo-historique », *Transports : Economie, politique, société*, 495, p. 17‑25.

BAILLY Antoine S., 1973, « Les théories de l'organisation de l'espace urbain », *Espace géographique*, 2(2), p. 81‑93.

BANISTER David, **BERECHMAN** Joseph, 2004, *Transport Investment and Economic Development*, Routledge, 383 p.

BANISTER David, **BERECHMAN** Yossi, 2001, « Transport investment and the promotion of economic growth », *Journal of transport geography*, 9(3), p. 209–218.

BAPTISTE Hervé, 2003, « Modélisation de l'évolution d'un système de transport et impacts sur un système départemental de villes », *in Graphes et réseaux, modélisation multiniveau*, Paris, Lavoisier, Information Géographique et Aménagement du Territoire, p. 113‑135.

BARABÁSI Albert-László, **ALBERT** Réka, 1999, « Emergence of scaling in random networks », *science*, 286(5439), p. 509–512.

BARABÁSI Albert-László, **FRANGOS** Jennifer, 2014, *Linked: the new science of networks science of networks*, Basic Books.

BARGE Olivier, **RODIER** Xavier, **DAVTIAN** Gourguen, **SALIGNY** Laure, 2004, « L'utilisation des systèmes d'information géographique appliquée à l'archéologie française », *Revue d'archéométrie*, 28(1), p. 15–24.

BARRO Robert J., 1990, « Government Spending in a Simple Model of Endogenous Growth », *Journal of Political Economy*, 98(5), p. 103‑126.

BARTHÉLEMY Marc, **FLAMMINI** Alessandro, 2006, « Optimal traffic networks », *Journal of Statistical Mechanics: Theory and Experiment*, 2006(7), p. L07002.

BAUDET Jean Claude, 2015, *Les plus grandes controverses de l'Histoire de la science*, Paris, La boite à pandore, 279 p.

BAUM-SNOW Nathaniel, 2007, « Did highways cause suburbanization? », *The Quarterly Journal of Economics*, p. 775–805.

BAUM-SNOW Nathaniel, **BRANDT** Loren, **HENDERSON** J. Vernon, **TURNER** Matthew A., **ZHANG** Qinghua, 2012, « Roads, railroads and decentralization of Chinese cities », Working paper.

BAVOUX Jean-Jacques, 1994, *Le« carrefour » bourguignon: analyse d'un espace de circulation*, Paris, CNRS editions, 179 p.

BAVOUX Jean-Jacques, 2000, « Les réseaux ferroviaires dans les pays en développement : une structuration encore souvent déficiente », *Flux*, 16(41), p. 17‑27.

BAVOUX Jean-Jacques, 2009, *La géographie: objet, méthodes, débats*, Paris, Armand Colin, 288 p.

BAVOUX Jean-Jacques, **BEAUCIRE** Francis, **CHAPELON** Laurent, **ZEMBRI** Pierre, 2005, *Géographie des transports*, Paris, A. Colin, 230 p.

BAZIN S., **BECKERICH** C., **BLANQUART** C., **DELAPLACE** M., **VANDENBOSSCHE** L., 2011, « Grande vitesse ferroviaire et développement économique local: une revue de la littérature », *Recherche Transports Sécurité*, 27(3), p. 215–238.

BEAUCIRE Francis, 2014, « De l'effet structurant ... au projet structuré », *L'Espace géographique*, 43(1), p. 51‑67.

BEAUGUITTE Laurent, 2012, « Analyser les réseaux avec R (packages statnet, igraph et tnet) », *Groupe FMR - Flux, matrices, réseaux*.

BEAUQUIER Danièle, **BERSTEL** Jean, **CHRETIENNE** Philippe, 1992, *Eléments d'algorithmique*, Paris, Masson, 463 p.

BEAVER S. H., 1941, « Railways in the Balkan Peninsula », *The Geographical Journal*, 97(5), p. 273‑294.

BÉDARD Yvan, **LARRIVÉE** Suzie, **PROULX** Marie-Josée, **NADEAU** Martin, 2004, « Modeling Geospatial Databases with Plug-Ins for Visual Languages: A Pragmatic Approach and the Impacts of 16 Years of Research and Experimentations on Perceptory », *in* **Wang** Shan, **Tanaka** Katsumi, **Zhou** Shuigeng, **Ling** Tok-Wang, **Guan** Jihong, **Yang** Dong-qing, **Grandi** Fabio, **Mangina** Eleni E., **Song** Il-Yeol, **Mayr** Heinrich C. (dir.), *Conceptual Modeling for Advanced Application Domains*, Springer Berlin Heidelberg, Lecture Notes in Computer Science, p. 17‑30.

BEGUIN Michèle, **PUMAIN** Denise, 1994, *La représentation des données géographiques: Statistique et cartographie*, Paris, Armand Colin, 256 p.

BERGER Peter, **LUCKMANN** Thomas, 1966, *La Construction sociale de la réalité*, New York, Doubleday & Company Inc., 344 p.

BERION Pascal, **JOIGNAUX** Guy, **LANGUMIER** Jean-François, 2007, « L'évaluation socio-économique des infrastructures de transport: Enrichir les approches du développement territorial », *Revue d'Économie Régionale & Urbaine*, 4, p. 651–676.

BERQUE Augustin, 1981, « Les lendemains de la cliométrie », *Espace géographique*, 10(3), p. 239‑240.

BERRY Brian JL, 1964, « Approaches to regional analysis: a synthesis », *Annals of the Association of American Geographers*, 54(1), p. 2–11.

BERRY David M., 2011, « The computational turn: Thinking about the digital humanities », *Culture Machine*, 12(0), p. 2.

BERTHELOT Jean-Michel, 2002, « Pour un programme sociologique non réductionniste en étude des sciences », *Revue européenne des sciences sociales. European Journal of Social Sciences*, XL-124, p. 233‑251.

BERTRAND Marianne, **DUFLO** Esther, **MULLAINATHAN** Sendhil, 2004, « How much should we trust differences-in-differences estimates? », *The Quarterly Journal of Economics*, 119(1), p. 249‑275.

BEYER Antoine, 2004, « La numérotation des routes françaises », *Flux*, 55, mars 2004, p. 17‑29.

BEYER Antoine, 2011, « La Belgique à la croisée des chemins. Une géohistoire des grandes infrastructures de transport du Plat Pays à l'aune de ses frontières », *Annales de géographie*, n° 681(5), p. 465‑485.

BEYZATLAR Mehmet Aldonat, **KARACAL** Müge, **YETKINER** Hakan, 2014, « Granger-causality between transportation and GDP: A panel data approach », *Transportation Research Part A: Policy and Practice*, 63, p. 43–55.

BIVAND Roger, 2011, « Geocomputation and open source software: components and software stacks », *NHH Dept. of Economics Discussion Paper*, 23.

BIVAND Roger, **NETELER** Markus, 2000, « Open Source Geocomputation : using the R data analysis language integrated with GRASS GIS and PostgreSQL data base systems »,.

BLACK William R., 1971, « An Iterative Model for Generating Transportation Networks* », *Geographical Analysis*, 3(3), p. 283‑288.

BLANCHARD Marcel, 1942, *Géographie des chemins de fer*, Paris, NRF, 230 p.

BLIER Gérard, 1996, *Nouvelle géographie ferroviaire de la France*, Paris, Vie du rail, 272 p.

BOARNET Marlon G., 1995, « Highways and Economic Productivity: Interpreting Recent Evidence »,.

BOARNET Marlon G., **HAUGHWOUT** Andrew F., 2000, « Do highways matter? Evidence and policy implications of highways' influence on metropolitan development », *University of California Transportation Center*.

BODENHAMER David J., **CORRIGAN** John, **HARRIS** Trevor M., 2010, *The Spatial Humanities: GIS and the Future of Humanities Scholarship*, Indiana University Press, 222 p.

BODENHAMER David J., **CORRIGAN** John, **HARRIS** Trevor M. (dir.), 2015, *Deep Maps and Spatial Narratives*, Bloomington, Indiana University Press, 256 p.

BOITEUX Marcel, 2001, « Transports : choix des investissements et coût des nuisances », Paris,

Commissariat général du Plan.

BOL Peter, **GE** Jianxiong, 2005, « China historical GIS », *Historical Geography*, 33, p. 150–152.

BONNAFOUS Alain, 1993, « Réseaux secondaires et transport intégré : promouvoir un aménagement du territoire équilibré »,.

BONNAFOUS Alain, 2014, « Permanent Observatories as Tools for Ex-Post Assessment: The French Case Study », International Transport Forum Discussion Paper, 2014/10, OECD Publishing.

BOULANGER Philippe, **TROCHET** Jean-René, 2005, *Où en est la géographie historique?*, Paris, Editions L'Harmattan, 347 p.

BOUNEAU Christophe, 1990, « Chemins de fer et développement régional en France de 1852 à 1937 : la contribution de la Compagnie du Midi », *Histoire, Économie et Société*, 9(1), p. 95‑112.

BOURDIEU Pierre, 1976, « Le champ scientifique », *Actes de la recherche en sciences sociales*, 2(2), p. 88‑104.

BOUTINET Jean-Pierre, 1990, *Anthropologie du projet*, Paris, Presses Universitaires de France.

BRAUDEL Fernand, 1949, *La Méditerranée et le monde méditerranéen à l'époque de Philippe II*, Paris, Armand Colin.

BRAUDEL Fernand, 1980, *Civilisation matérielle, économie et capitalisme, XV-XVIII siecle: Les structures du Quotidien: le possible et l'impossible*, Armand Colin, 736 p.

BRETAGNOLLE Anne, 2003, « Vitesse des transports et sélection hiérarchique entre les villes françaises », *in Données urbaines*, Anthropos, Paris, p. 433.

BRETAGNOLLE Anne, 2009, *Villes et réseaux de transport: des interactions dans la longue durée (France, Europe, Etats-Unis)*, Ph.D. thesis, Paris, Université Panthéon-Sorbonne-Paris I, 204 p.

BRETAGNOLLE Anne, 2014, « Les effets structurants des transports, une question d'échelles ? », *L'Espace géographique*, 43(1), p. 51‑67.

BRETAGNOLLE Anne, **GIRAUD** Timothée, **VERDIER** Nicolas, 2010, « Modéliser l'efficacité d'un réseau », *L'Espace géographique*, Vol. 39(2), p. 117‑131.

BRETAGNOLLE Anne, **VERDIER** Nicolas, 2005, « Images d'un réseau en évolution: les routes de poste dans la France préindustrielle (XVIIe siécle–début XIXe siècle) », *Mappemonde*, 79(3), p. 1‑8.

BRUINSMA Frank, **RIETVELD** Pieter, 1998, « The accessibility of European cities: theoretical framework and comparison of approaches », *Environment and Planning A*, 30(3), p. 499–521.

BRUNSDON Chris, **FOTHERINGHAM** A. Stewart, **CHARLTON** Martin E., 1996, « Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity », *Geographical Analysis*, 28(4), p. 281–298.

CAMAGNI Roberto, 1996, *Principes et modèles de l'économie urbaine*, Paris, Economica, 382 p.

CAPOT-REY Robert, 1946, *Géographie de la circulation*, Paris, Gallimard.

CARALP-LANDON Raymonde, 1959, *Les chemins de fer dans le massif central: étude des voies ferrées régionales*, Impr. nationale.

CARON François, 1992, « Introduction : l'évolution des transports terrestres en Europe (vers 1800-vers 1940) », *Histoire, économie et société*, 11(1), p. 5‑11.

CARON Francois, 1997, *Histoire des chemins de fer en France: 1740-1883*, Paris, Fayard.

CARON François, 2002, « Les réseaux et les politiques d'aménagement du territoire : l'exemple des chemins de fer », *in La politique d'aménagement du territoire. Racines, logiques, résultats*, Rennes, Presses universitaires de Rennes, p. 17‑29.

CASTELLS Manuel, 1998, *La Société en réseaux*, Paris, Fayard, 671 p.

CASTELLS Manuel, **DELAMARE** Philippe, 1998, *La société en réseaux*, Fayard Paris.

CHANG Stephanie, 2004, « Transportation geography: the influence of Walter Isard and regional science », *Journal of Geographical Systems*, 6, p. 55‑69.

CHAPELON Laurent, 1997, *Offre de transport et aménagement du territoire*, Ph.D. thesis, Tours, Université François Rabelais, 558 p.

CHARLOT Sylvie, **LAFOURCADE** Miren, 2000, « Infrastructures publiques, coûts de transport et croissance régionale », *Economie géographique*, p. 143–177.

CHATELAIN Abel, 1947, « Démogéographie du grand tronc ferré sud-est (Paris-Lyon-Méditerranée) », *Les Études rhodaniennes*, 22(1), p. 35–82.

CHENG Tao, **HAWORTH** James, **MANLEY** Ed, 2012, « Advances in geocomputation (1996-2011) », *Computers, Environment and Urban Systems*, 36(6), p. 481–487.

CHESNAIS Michel, 1975, « Géographie des transports et analyse régionale », *Espace géographique*, 4(3), p. 183–195.

CHEVALIER Michel, 1832, *Système de la Méditerranée*, Paris, Aux bureaux du Globe, 66 p.

CHEVALLIER Jacques, 2012, *Le service public*, 9e édition, PRESSES UNIVERSITAIRES DE FRANCE - PUF.

CHEVALLIER Raymond, 1972, *Les voies romaines*, Paris, Armand Colin.

CHI Guangqing, 2012, « The impacts of transport accessibility on population change across rural, suburban and urban areas: a case study of Wisconsin at sub-county levels », *Urban studies*, 49(12), p. 2711–2731.

CHI Guangqing, **VOSS** Paul R., **DELLER** Steven C., 2006, « Rethinking highway effects on population change », *Public Works Management & Policy*, 11(1), p. 18–32.

CHRISTALLER Walter, 1933, *Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*, University Microfilms.

CLARK Andrew Hill, 1949, « The invasion of New Zealand by people, plants and animals. », *The invasion of New Zealand by people, plants and animals.*

CLAVAL Paul, 1981, *La logique des villes: essai d'urbanologie*, **Claval** Françoise (dir.), Paris, France, LITEC, 633; xx p.

CLAVAL Paul, 1985, « Causalité et géographie », *Espace géographique*, 14(2), p. 109‑115.

CLAVAL Paul, 2005, « Géographie et temporalités », *Où en est la géographie historique ?*, p. 345.

CLOZIER René, 1963, *Géographie de la circulation*, Paris, Genin.

COLLET Claude, 2005, « Analyse spatiale, géomatique et systèmes d'information géographique. », *Revue internationale de géomatique*, 15(4), p. 393–414.

COLLIOT-THELENE Catherine, 2004, « Expliquer/comprendre : relecture d'une controverse », *Espaces Temps*, 84(1), p. 6‑23.

COMBES Pierre-Philippe, **LAFOURCADE** Miren, 2012, « Revue de la littérature académique quantifiant les effets d'agglomération sur la productivité et l'emploi », Paris, Société du Grand Paris.

COMBES Pierre-Philippe, **LAFOURCADE** Miren, **THISSE** Jacques-François, **TOUTAIN** Jean-Claude, 2011, « The rise and fall of spatial inequalities in France: A long-run perspective », *Explorations in Economic History*, 48(2), p. 243–271.

COMMENGES Hadrien, 2013, *L'invention de la mobilité quotidienne. Aspects performatifs des instruments de la socio-économie des transports*, Ph.D. thesis, Paris, Paris VII - Paris-Diderot, 392 p.

CORNILLY Jean-Yves, **DANZANVILLIERS** P, **FURGAUT** C, 1988, « Les effets socio-économiques des grandes infrastructures routières », Bagneux, SETRA.

COSTA Laurent, 2012, « La construction de référentiels géohistoriques : un enjeu pour l'interdisciplinarité dans les sciences historiques », *L'Espace géographique*, Tome 41(4), p. 340‑351.

COUCLELIS Helen, 1998, « Geocomputation in context », *in Geocomputation: A primer*, Chichester, John Wiley and Sons New York, p. 17–30.

CRAFTS Nicholas, **LEUNIG** Timothy, 2005, « The historical significance of transport for economic growth and productivity », Research Annex 1.

CRESCENZI Riccardo, **RODRÍGUEZ-POSE** Andrés, 2008, « Infrastructure endowment and investment as determinants of regional growth in the European Union », *Eib Papers*, 13(2), p. 62–101.

CRESSWELL Tim, 2001, « The production of mobilities », *New Formations*, 43, p. 11‑25.

CRESSWELL Tim, 2013, *Place: A short introduction*, John Wiley & Sons.

CRESSWELL Tim, **DIXON** Deborah P., **BOL** Peter K., **ENTRIKIN** J. Nicholas, 2015, « Editorial », *GeoHumanities*, 1(1), p. 1‑19.

CRESWELL John W., 2013, *Research design: Qualitative, quantitative, and mixed methods approaches*, Sage publications, 273 p.

CROZET Matthieu, **LAFOURCADE** Miren, 2009, *La nouvelle économie géographique*, Paris, France, La Découverte, 128 p.

CSARDI Gabor, **NEPUSZ** Tamas, 2006, « The igraph software package for complex network research », *InterJournal, Complex Systems*, 1695(5), p. 1–9.

CURIEN Nicolas, 2000, *Economie des réseaux*, La découverte.

CUYALA Sylvain, 2014, *Analyse spatio-temporelle d'un mouvement scientifique. L'exemple de la géographie théorique et quantitative européenne francophone.*, phdthesis, Université Paris 1 Panthéon-Sorbonne.

DACOS Marin, **MOUNIER** Pierre, 2014, « Humanités numériques. Etat des lieux et positionnement de la recherche française dans le contexte international », Paris, Institut Français.

DANCOISNE Pascale, 1984, *Théorie des graphes et constitution du réseau ferré francais*, Ph.D. thesis, Paris, Paris I - Sorbonne, 280 p.

DAVID Paul A., 1969, « Transport innovation and economic growth: Professor Fogel on and off the rails », *The Economic History Review*, 22(3), p. 506–525.

DEAR Michael, 2015, « Practicing Geohumanities », *GeoHumanities*, 1(1), p. 20‑35.

DEBRIE Jean, 2008, « Géohistoire d'un tracé technique: frontières et réseaux en Afrique de l'ouest continentale «francophone» », *Flux*, 70(4), p. 49–58.

DELAPLACE Marie, 2014, « Dessertes ferroviaires à grande vitesse : des effets structurants toujours ... la pensée et l'actioin publique », *L'Espace géographique*, 43(1), p. 51‑67.

DESMARAIS Gaëtan, **RITCHOT** Gilles, 1998, « La modélisation dynamique en géographie humaine », *Cahiers de géographie du Québec*, 42(117), p. 329–332.

DESMICHEL Pascal, 2012, « Le patrimoine ferroviaire bâti en Limousin rural », *Histoire & Sociétés Rurales*, 37(1), p. 149‑171.

DESROSIERES Alain, 1994, « Le territoire et la localité. Deux langages statistiques », *Politix*, 7(25), p. 46‑58.

DESROSIERES Alain, 2008, *Pour une sociologie historique de la quantification*, Paris, Presses de l'Ecole des mines, 328 p.

DESROSIERES Alain, 2010, *La politique des grands nombres*, Paris, La Découverte, 462 p.

DEZERT Bernard, **VANDERMOTTEN** Christian, 2008, *L'identité de l'Europe: Histoire et géographie d'une quête d'unité*, Armand Colin, Paris, Armand Colin, 336 p.

DIEBOLT Claude, 2001, « La cliométrie : une discipline de synthèse », *Association Française de Cliométrie*, p. 2.

DIEBOLT Claude, 2004, « La cliométrie se rebiffe », *Association Française de Cliométrie*, p. 6.

DIEBOLT Claude, 2005, « Quelle place pour les recherches empiriques en économie ? », *Association Française de Cliométrie*, p. 4.

DIEBOLT Claude, 2016, « Cliometrica after 10 years: definition and principles of cliometric research », *Cliometrica*, 10(1), p. 1–4.

DIJKSTRA Edsger W., 1959, « A note on two problems in connexion with graphs », *Numerische mathematik*, 1(1), p. 269–271.

DION Roger, 1936, *Le Val de Loire. Etude de géographie régionale*, Tours, Arrault et cie, maîtes imprimeurs, 752 p.

DOBRUSZKES Frédéric, **LANNEAUX** Marie-Agnès, 2010, « Conclusions : comme un reflet de la géographie des transports à la française », *Belgeo. Revue belge de géographie*, 1‑2, p. 245‑246.

DONALDSON Dave, 2010, « Railroads of the Raj: Estimating the impact of transportation infrastructure », National Bureau of Economic Research.

DORMONT Brigitte, 1989, « Petite apologie des données de panel », *Économie & prévision*, 87(1), p. 19‑32.

DUCLERT Vincent, 2002, « Archives orales et recherche contemporaine », *Sociétés & Représentations*, 13, p. 69‑86.

DUCRUET César, 2010a, « Les mesures globales d'un réseau »,.

DUCRUET César, 2010b, « Les mesures locales d'un réseau »,.

DUCRUET César, **NOTTEBOOM** Theo, 2012, « The worldwide maritime network of container shipping: spatial structure and regional dynamics », *Global Networks*, 12(3), p. 395‑423.

DUPUY Gabriel, 1987a, « Les réseaux techniques sont-ils des réseaux ? », *Espace géographique*, 16(3), p. 175‑184.

DUPUY Gabriel, 1987b, « Vers une théorie territoriale des réseaux: une application au transport urbain », *Annales de Géographie*, 96, p. 658–679.

DUPUY Gabriel, 1989, « Réseaux, philosophie de l'organisation », *Encyclopædia Universalis*.

DUPUY Gabriel, 1991, *L'urbanisme des réseaux: théories et méthodes*, Paris, Armand Colin, 198 p.

DUPUY Gabriel, 1993, « Géographie et économie des réseaux », *Espace géographique*, 22(3), p. 193‑209.

DUPUY Gabriel, 1997, « Offner J.-M. et Pumain D. (dir.), ~~Réseaux et territoires, significations croisées~~ », *Annales de Géographie*, 106(597), p. 549‑551.

DUPUY Gabriel, 1999, *La dépendance automobile: symptômes, analyses, diagnostic, traitements*, Anthropos.

DUPUY Gabriel, 2011, « Fracture et dépendance: l'enfer des réseaux? », *Flux*, n° 83(1), p. 6‑23.

DUPUY Gabriel, **CREWS** Judith, 1991, « Revue de la thèse de Pascale Dancoisne », *Flux*, 7(3), p. 58‑60.

DUPUY Gabriel, **OFFNER** Jean-Marc, 2005, « Réseau: bilans et perspectives », *Flux*, 4, p. 38–46.

DUPUY Gabriel, **STRANSKY** Vaclav, 1996, « Cities and highway networks in Europe », *Journal of Transport Geography*, 4(2), p. 107–121.

DURAND-DASTES François, 1990, « La mémoire de Gaïa », *Actes de Géopoint*, p. 147‑151.

DURAND-DASTES François, 2001, « Les concepts de la modélisation en analyse spatiale », *in Modèles en analyse spatiale*, Paris, Lavoisier, p. 31‑59.

DURANTON Gilles, **TURNER** Matthew A., 2012, « Urban Growth and Transportation », *The Review of Economic Studies*, 79(4), p. 1407‑1440.

EDDINGTON Rod, 2006, *The Eddington Transport Study. Main Report: Transport's Role in Sustaining the UK's Productivity and Competitiveness*, Londres, Her Majesty's Stationery Office, 52 p.

EDELBLUTTE Simon, 2008, « Paysages et territoires du patrimoine industriel au Royaume-Uni », *Revue Géographique de l'Est*, 48(1‑2).

EGENHOFER Max J., **FRANZOSA** Robert D., 1991, « Point-set topological spatial relations », *International Journal of Geographical Information System*, 5(2), p. 161–174.

EHLEN Judy, **CALDWELL** Douglas R., **HARDING** Stephen, 2002, « GeoComputation: what is it? », *Computers, Environment and Urban Systems*, 26(4), p. 257–265.

ELISSALDE Bernard, 2000, « Géographie, temps et changement spatial », *Espace géographique*, 29(3), p. 224–236.

EWING Reid, **CERVERO** Robert, 2010, « Travel and the built environment: a meta-analysis », *Journal of the American planning association*, 76(3), p. 265–294.

FACCHINETTI-MANNONE Valérie, **BAVOUX** Jean-Jacques, 2010, « L'implantation des gares TGV en France : tensions interscalaires, jeux d'acteurs et recompositions spatiales », *Belgeo. Revue belge de géographie*, 1‑2, p. 9‑22.

FAVORY François, **NUNINGER** Laure, **SANDERS** Lena, 2012, « Intégration de concepts de géographie et d'archéologie spatiale pour l'étude des systèmes de peuplement », *L'Espace géographique*, Tome 41(4), p. 295‑309.

FERNANDEZ Enrique, **DE CEA** Joaquin, **FLORIAN** Michael, **CABRERA** Enrique, 1994, « Network equilibrium models with combined modes », *Transportation Science*, 28(3), p. 182–192.

FINNEGAN Diarmid A., 2007, « The Spatial Turn: Geographical Approaches in the History of Science », *Journal of the History of Biology*, 41(2), p. 369‑388.

FISHLOW Albert, 1965, *American Railroads and the Transformation of the Ante-bellum Economy*, Harvard University Press Cambridge, MA.

FLONNEAU Mathieu, **GUIGUENO** Vincent, 2009, *De l'histoire des transports à l'histoire de la mobilité*, Rennes, Presses universitaires de Rennes, 332 p.

FOGEL Robert William, 1962, « A Quantitative Approach to the Study of Railroads in American Economic Growth: A Report of Some Preliminary Findings », *The Journal of Economic History*, 22(2), p. 163‑197.

FOSSEYEUX Jean, 2008, « Vingt ans d'histoire du patrimoine architectural et du paysage ferroviaire », *Revue d'histoire des chemins de fer*, 39, p. 245‑253.

FOTHERINGHAM A. Stewart, **BRUNSDON** Chris, **CHARITON** Martin, 1998, « Geographically weighted regression-modelling spatial non-stationarity », *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47(3), p. 431–443.

FOX Michael, 1995, « Transport planning and the human activity approach », *Journal of Transport Geography*, 3(2), p. 105‑116.

FRANCH Xavi, **MORILLAS-TORNÉ** Mateu, **MARTÍ-HENNEBERG** Jordi, 2013, « Railways as a Factor of Change in the Distribution of Population in Spain, 1900–1970 », *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(3), p. 144‑156.

FRANK Andrew U., 1992, « Spatial concepts, geometric data models, and geometric data structures », *Computers & Geosciences*, 18(4), p. 409–417.

FRANKHAUSER Pierre, 1997, « L'approche fractale. Un nouvel outil de réflexion dans l'analyse spatiale des agglomérations urbaines », *Population*, 52(4), p. 1005‑1040.

FREMONT Antoine, 2014, « Au-delà du mythe, les transports : un levier pour des projets de territoires ? », *L'Espace géographique*, 43(1), p. 51‑67.

FRISTCH B., 2007, « Infrastructures de transport, densification et étalement urbains: quelques enseignements de l'expérience nantaise », *Cahiers Scientifiques du Transport*, 51, p. 37–60.

GAHEGAN Mark, 1999, « Guest Editorial: What is Geocomputation? », *Transactions in GIS*, 3(3), p. 203–206.

GARMY Pierre, **KADDOURI** Lahouari, **ROZENBLAT** Céline, **SCHNEIDER** Laurent, 2005, « Logiques spatiales et «systèmes de villes» en Lodévois de l'Antiquité à la période moderne », *Temps et espaces de l'homme en société, analyses et modèles spatiaux en archéologie. XXVème rencontres internationales d'archéologie et d'histoire d'Antibes*, p. 1–12.

GARRISON William L., 1960, « Connectivity of the Interstate Highway System », *Papers in Regional Science*, 6(1), p. 121‑137.

GARRISON William L., 1990, « Networks : reminiscence and lessons », *Flux*, 6(1), p. 5‑12.

GARRISON William L., **LEVINSON** David M., 2014, *The Transportation Experience: Policy, Planning, and Deployment*, OUP USA, 635 p.

GARRISON William L., **SOULEYRETTE** Reginald R., 1996, « Transportation, innovation, and development: the companion innovation hypothesis », *Logistics and Transportation review*, 32(1), p. 5.

GASTNER Michael T., **NEWMAN** Mark EJ, 2006, « Shape and efficiency in spatial distribution networks », *Journal of Statistical Mechanics: Theory and Experiment*, 2006(1), p. P01015.

GENRE-GRANDPIERRE Cyrille, **FRANKHAUSER** Pierre, 1998, « La géométrie fractale, un nouvel outil pour évaluer le rôle de la morphologie des réseaux de transport public dans l'organisation spatiale des agglomérations », *Les Cahiers scientifiques du transport*, 33, p. 41‑78.

GEURS Karst T., **RITSEMA VON ECK** J. R., 2001, « Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact », RIVM Report 408505 006, Bilthoven, National Institute of Public Health and the Environment.

GEURS Karst T., **VAN WEE** Bert, 2004, « Accessibility evaluation of land-use and transport strategies: review and research directions », *Journal of Transport Geography*, 12(2), p. 127‑140.

GILLES Henri, **GUIRONNET** Jean-Pascal, **PARENT** Antoine, 2014, « Géographie économique des morts de 14-18 en France », *Revue économique*, 65(3), p. 519‑532.

GIRVAN Michelle, **NEWMAN** Mark EJ, 2002, « Community structure in social and biological networks », *Proceedings of the national academy of sciences*, 99(12), p. 7821–7826.

GLEYZE Jean-François, 2007, « Effets spatiaux et effets réseau dans l'évaluation d'indicateurs sur les nø euds d'un réseau d'infrastructure », *Cybergeo: European Journal of Geography*.

GOLD Matthew, 2012, *Debates in the digital humanities*, U of Minnesota Press, 513 p.

GONJO Yasuo, 1972, « Le « plan Freycinet », 1878-1882 : un aspect de la « grande dépression » économique en France », *Revue Historique*, 248(1 (503)), p. 49‑86.

GOODCHILD M.F, **LONGLEY** P.P.A, 2014, « The Practice of Geographic Information Science », *in Handbook of Regional Science*, Springer Berlin Heidelberg, Berlin, p. 1107‑1122.

GOODCHILD Michael F., 1995, « Future Directions for Geographic Information Science », *Geographic Information Sciences*, 1(1), p. 1‑7.

GRAMLICH Edward M., 1994, « Infrastructure Investment: A Review Essay », *Journal of Economic Literature*, 32(3), p. 1176‑1196.

GRAS Alain, 1998, *Les macro-systèmes techniques*, Paris, Presses Universitaires de France - PUF.

GRATALOUP Christian, 1994, *L'espace de la transition : essai de géohistoire chorématique*, Ph.D. thesis, Paris, Université de Paris 1.

GRATALOUP Christian, 1996, *Lieux d'histoire. Essai de géohistoire systématique*, Paris, Reclus, 200 p.

GRATALOUP Christian, 2015, *Introduction à la géohistoire*, Paris, Armand Colin, 224 p.

GREGORY Ian N., 2005, « The great Britain historical GIS », *Historical Geography*, 33, p. 136–138.

GREGORY Ian N., **ELL** Paul S., 2006, « Error-sensitive historical GIS: Identifying areal interpolation errors in time-series data », *International Journal of Geographical Information Science*, 20(2), p. 135‑152.

GREGORY Ian N., **ELL** Paul S., 2007, *Historical GIS: technologies, methodologies, and scholarship*, Cambridge University Press, 227 p.

GREGORY Ian N., **GEDDES** Alistair, 2014, *Toward Spatial Humanities: Historical GIS and Spatial History*, Bloomington, Indiana University Press, 205 p.

GREGORY Ian N., **HARDIE** Andrew, 2011, « Visual GISting: bringing together corpus linguistics and Geographical Information Systems », *Literary and linguistic computing*, 26(3), p. 297–314.

GREGORY Ian N., **HEALEY** Richard G., 2007a, « Historical GIS: structuring, mapping and analysing geographies of the past », *Progress in Human Geography*, 31(5), p. 638‑653.

GREGORY Ian N., **HEALEY** Richard G., 2007b, « Historical GIS: structuring, mapping and analysing geographies of the past », *Progress in Human Geography*, 31(5), p. 638‑653.

GRENON Pierre, **SMITH** Barry, 2004, « SNAP and SPAN: Towards dynamic spatial ontology », *Spatial cognition and computation*, 4(1), p. 69–104.

GRIMMER Justin, **STEWART** Brandon M., 2013, « Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts », *Political Analysis*, p. mps028.

GROSSETTI Michel, 2011, « L'espace à trois dimensions des phénomènes sociaux », *SociologieS*.

GUELKE Leonard, 1997, « The Relations Bwetween Geography and History Reconsidered », *History and Theory*, 36(2), p. 191‑234.

GUERIN-PACE France, 1993, *Deux siècles de croissance urbaine: la population des villes françaises de 1831 à 1990*, Pairs, Anthropos, 205 p.

GUERIN-PACE France, **PUMAIN** Denise, 1990, « 150 ans de croissance urbaine », *Economie et statistique*, 230(1), p. 5–16.

GUERMOND Yves, 2005, *Modélisations en géographie. Déterminismes et complexités*, Paris, Hermes Science Publications, 389 p.

GUIGUENO Vincent, 2009, « Les paysages de la vitesse », *in De l'histoire des transports à l'histoire de la mobilité*, Rennes, Presses universitaires de Rennes, p. 332.

GUILLERME André, 1991, « Réseau: Genèse d'une catégorie dans la pensée de l'ingénieur sous la Restauration », *Flux*, 7(6), p. 5–17.

GUTIÉRREZ Javier, **CONDEÇO-MELHORADO** Ana, **MARTÍN** Juan Carlos, 2010, « Using accessibility indicators and GIS to assess spatial spillovers of transport infrastructure investment », *Journal of transport geography*, 18(1), p. 141–152.

GUTIÉRREZ Javier, **URBANO** Paloma, 1996, « Accessibility in the European Union: the impact of the trans-European road network », *Journal of Transport Geography*, 4(1), p. 15‑25.

HACKING Ian, 2001, *Entre science et réalité : la construction sociale de quoi ?*, La Découverte, Paris, 308 p.

HAGERSTRAND Torsten, 1967, *Innovation diffusion as a spatial process*, Chicago, University of Chicago Press.

HAGGETT Peter, **FRECHOU** Hubert, 1973, *L'analyse spatiale en géographie humaine*, Armand Colin, Paris, 390 p.

HAGGETT Peter, **FREY** Allan E., **CLIFF** A. D., 1977, *Locational Analysis in Human Geography*, New York, John Wiley & Sons, 605 p.

HAKLAY Muki, **SINGLETON** Alex, **PARKER** Chris, 2008, « Web Mapping 2.0: The Neogeography of the GeoWeb », *Geography Compass*, 2(6), p. 2011‑2039.

HAMAINA Rachid, **LEDUC** Thomas, **MOREAU** Guillaume, 2012, « Caractérisation des tissus urbains à partir de l'analyse structurelle des réseaux viaires », *Cybergeo : European Journal of Geography*.

HANSEN Walter G., 1959, « How accessibility shapes land use », *Journal of the American Institute of Planners*, 25(2), p. 73–76.

HANSON Gordon, 2005, « Market potential, increasing returns and geographic concentration », *Journal of International Economics*, 67(1), p. 1‑24.

HARMELLE Claude, **ELIAS** Gabrielle, 1982, « Saint-Antonin et sa région (1850-1940): révolutions des transports et changement social », *Revue Recherches*, 47.

HARRIS Chauncy D., 1954, « The Market as a Factor in the Localization of Industry in the United States », *Annals of the Association of American Geographers*, 44(4), p. 315‑348.

HASKINS Conaway B., 2002, *The influence of highways on rural economic development: Evidence from North Carolina*, Ph.D. thesis, University of North Carolina at Chapel Hill.

HEAD Keith, **MAYER** Thierry, 2004, « Market potential and the location of Japanese investment in the European Union », *Review of Economics and Statistics*, 86(4), p. 959–972.

HEALEY Richard G., **THOMAS** William G., **LAHMAN** Katie, 2013, « Railroads and regional labor markets in the mid-nineteenth-century United States: a case study of the Baltimore and Ohio Railroad », *Journal of Historical Geography*, 41, p. 13‑32.

HEFFER Jean, s. d., « Nouvelle histoire économique », *Encyclopædia Universalis*.

HERTZOG Anne, 2011, « Les géographes et le patrimoine », *EchoGéo*, 18.

HILAL Mohamed, 2012, « Cinquante ans de mobilités résidentielles et quotidiennes : quelles formes territoriales pour un mode de vie durable ? », Consultation de recherche, Dijon, Ministère de l'Egalité des territoires et du Logement Plan Urbanisme Construction Architecture.

HOLL Adelheid, 2007, « Twenty years of accessibility improvements. The case of the Spanish motorway building programme », *Journal of Transport Geography*, 15(4), p. 286‑297.

HORNUNG Erik, 2012a, « Railroads and micro-regional growth in prussia », Ifo Working Paper.

HORNUNG Erik, 2012b, « Railroads and micro-regional growth in prussia », Ifo Working Paper.

HOWKINS Trevor J, 1996, « Railway geography and the demarcation of Poland's borders 1918–1930 », *Journal of Transport Geography*, 4(4), p. 287‑299.

HUGHES Thomas Parke, 1993, *Networks of power: electrification in Western society, 1880-1930*, JHU Press.

JACOB-ROUSSEAU Nicolas, 2010, « Géohistoire/géo-histoire: quelles méthodes pour quel récit? », *Géocarrefour*, 84(4), p. 211–216.

JANELLE Donald G., 1969, « Spatial reorganization: a model and concept », *Annals of the*

Association of American Geographers, 59(2), p. 348–364.

JEFFERSON Mark, 1928, « The Civilizing Rails », *Economic Geography*, 4(3), p. 217‑231.

JENKS Leland H., 1944, « Railroads as an Economic Force in American Development », *The Journal of Economic History*, 4(1), p. 1‑20.

JIMÉNEZ Jorge Solanas, **PALAU** Eduard Alvarez, **HENNEBERG** Jordi Martí, 2015, « Estación ferroviaria y ciudades intermedias: lectura geo-espacial del crecimiento urbano mediante indicadores SIG vectoriales. El caso de Cataluña (1848 – 2010) », *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 0(16), p. 253‑280.

KADDOURI Lahouari, 2007, *Structures spatiales et mises en réseaux de villes pour la régionalisation des territoires*, Ph.D. thesis, Montpellier, Université Paul Valéry - Montpellier III, 311 p.

KANSKY Karel Joseph, 1963, *Structure of transportation networks*, Ph.D. thesis, Chicago, University of Chicago, 155 p.

KASRAIAN Dena, **MAAT** Kees, **STEAD** Dominic, **VAN WEE** Bert, 2016, « Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies », *Transport Reviews*, p. 1–21.

KEELER Theodore E., **YING** John S., 1988, « Measuring the benefits of a large public investment », *Journal of Public Economics*, 36(1), p. 69‑85.

KEELING David J., 2007, « Transportation geography: new directions on well-worn trails », *Progress in Human Geography*, 31(2), p. 217.

KEELING David J., 2009, « Transportation geography: local challenges, global contexts », *Progress in Human Geography*, 33(4), p. 516.

KEELING David J., **BLOMLEY** Nick, 2007, « Transportation geography-new regional mobilities », *Progress in Human Geography*.

KESSLER Marie-Christine, 1986, « Les grands corps de l'État », *La Revue administrative*, 39(231), p. 221‑228.

KITCHIN Rob, 2014a, *The data revolution: Big data, open data, data infrastructures and their consequences*, Sage.

KITCHIN Rob, 2014b, « Big Data, new epistemologies and paradigm shifts », *Big Data & Society*, 1(1), p. 2053951714528481.

KIVELÄ Mikko, **ARENAS** Alex, **BARTHELEMY** Marc, **GLEESON** James P., **MORENO** Yamir, **PORTER** Mason A., 2014, « Multilayer networks », *Journal of Complex Networks*, 2(3), p. 203‑271.

KNOWLES Anne Kelly, 2005, « Emerging trends in historical GIS », *Historical Geography*, 33(1), p. 7–13.

KOENIG J. G., 1980, « Indicators of urban accessibility: Theory and application », *Transportation*, 9(2), p. 145‑172.

KOKALJ Ziga, **PEHANI** Peter, **HVALA** Sneza Tecco, **OŠTIR** Krištof, 2006, « Application of Internet GIS tools for heritage management. ARKAS case study. », Vol.1, p. 267‑272.

KONDRATIEFF Nicolaï Dimitrievitch, 1992, *Les grands cycles de la conjoncture*, Paris, Economica, 562 p.

KRUGMAN Paul, 1991, « Increasing returns and economic geography », *Journal of Political Economy*, 99(3), p. 483‑499.

KRUGMAN Paul, **VENABLES** Anthony J., 1995, « Globalization and the Inequality of Nations », Working Paper, 5098, National Bureau of Economic Research.

KUNZ Andreas, 1992, « Voies navigables et développement économique », *Histoire, économie et société*, 11(1), p. 13‑17.

KWAN Mei-Po, 2004, « GIS Methods in Time-Geographic Research: Geocomputation and Geovisualization of Human Activity Patterns », *Geografiska Annaler: Series B, Human Geography*, 86(4), p. 267‑280.

KWAN Mei-Po, **JANELLE** Donald G., **GOODCHILD** Michael F., 2003, « Accessibility in space and time: A theme in spatially integrated social science », *Journal of Geographical Systems*, 5(1), p. 1‑3.

LAFOURCADE Miren, **MAYER** Thierry, 2009, « Des routes pavées de bonnes intentions », *Références économiques pour le développement durable*, 9, p. 5.

LALANNE Léon, 1863, « Comptes rendus hebdomadaires des séances de l'Académie des sciences », *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 57, p. 206‑210.

LANGLOIS Patrice, **REGUER** Daniel, 2005, « La place du modèle et de la modélisation en Sciences Humaines et Sociales », *in Modélisations en géographie, déterminismes et complexité*, Paris, Hermes Science Publications, p. 389.

LANGRAN Gail, 1992, *Time In Geographic Information Systems*, CRC Press, 204 p.

LE DEUFF Olivier, 2014, *Le temps des humanités digitales. La mutation des sciences humaines et sociales*, FYP Editions, Limoges, Société de la connaissance, 175 p.

LE GALLO Julie, 2002, *Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale*, phdthesis, Université de Bourgogne.

LE SUEUR Bernard, 1989, « L'évolution de la navigation intérieure et de ses métiers, XIXème - XXème siècle », *Culture Technique*, 19, p. 230‑237.

LEMERCIER Claire, 2005, « Analyse de réseaux et histoire »,.

LENNTORP B., 1976, « Paths in time-space environments », *A time-geographic study of movement possibilities of individuals. Meddelanden fr\a an Lunds universitets Geografiska institutioner, Avhandlingar LXXVII, Lund*.

LEPETIT Bernard, 1984, *Chemins de terre et voies d'eau, Réseaux de transports, Organisation de l'espace*, Editions de l'Ecole des Hautes Etudes en Sciences Sociales, Paris, 148 p.

LEPETIT Bernard, 1986, « L'impensable réseau. Les routes françaises avant les chemins de fer », *Cahier / Groupe Réseaux*, 2(5), p. 11‑29.

LEVINSON David, **YERRA** Bhanu, 2006, « Self-organization of surface transportation networks », *Transportation Science*, 40(2), p. 179–188.

LEVY Jacque, **LUSSAULT** Michel, 2003, *Dictionnaire de la géographie et de l'espace des sociétés*, Belin, Paris.

L'HOSTIS Alain, 2003, « Théorie des graphes et représentations des distances : chronocartes et autres représentations », *in Graphes et réseaux, modélisation multiniveau*, Paris, Lavoisier, Information Géographique et Aménagement du Territoire, p. 215‑230.

LICHTER Daniel T., **FUGUITT** Glenn V., 1980, « Demographic response to transportation innovation: The case of the interstate highway », *Social Forces*, 59(2), p. 492–512.

LIU Dan, **MEISSNER** Christopher M., 2015, « Market potential and the rise of US productivity leadership », *Journal of International Economics*, 96(1), p. 72‑87.

LIVET Pierre, 1983, « La fascination de l'auto-organisation », *in L'auto-organisation, de la physique au politique*, Paris, Le Seuil, p. 556.

LONGLEY P. A., 1998, « Foundations », *in Geocomputation: A primer*, Chichester, John Wiley and Sons New York, p. 3–15.

LOUF Rémi, **JENSEN** Pablo, **BARTHELEMY** Marc, 2013, « Emergence of hierarchy in cost-driven growth of spatial networks », *Proceedings of the National Academy of Sciences*, 110(22), p. 8824‑8829.

LOZANO Angélica, **STORCHI** Giovanni, 2002, « Shortest viable hyperpath in multimodal networks », *Transportation Research Part B: Methodological*, 36(10), p. 853‑874.

LUSSAULT Michel, 2007, *L'homme spatial: la construction sociale de l'espace humain*, Paris, Le Seuil, 363 p.

MACKINDER Halford John, 1902, *Britain and the British seas*, New York: D. Appleton.

MAGUIRE David J., 1991, « An overview and definition of GIS », *in Geographical Information Systems: Overview, Principles and Applications*, Harlow, Longmans, p. 9‑20.

MARCEL Émerit, 1975, « Saint-Simon père de la sociologie. », *Annales. Économies, Sociétés, Civilisations*, 30(2), p. 309‑316.

MARTÍ-HENNEBERG Jordi, 2011, « Geographical information systems and the study of history », *Journal of Interdisciplinary History*, 42(1), p. 1–13.

MARTÍ-HENNEBERG Jordi, 2013, « European integration and national models for railway networks (1840–2010) », *Journal of Transport Geography*, 26, p. 126–138.

MATHIS Philippe, 2003, *Graphes et réseaux. Modélisation multiniveau*, Paris, Hermes Science Publications, 361 p.

MATTELART Armand, 1994, *L'invention de la communication*, La Decouverte, Paris, 380 p.

MAYER Thierry, **TREVIEN** Corentin, 2012, « Transports publics urbains et attractivité internationale. Enjeux théoriques et évaluation empirique de l'impact du Réseau express régional sur les investissements étrangers en Île de France », Paris, Société du Grand Paris.

MCCLOSKEY Donald N., 1985, « The problem of audience in historical economics: Rhetorical thoughts on a text by Robert Fogel », *History and Theory*, 24(1), p. 1–22.

MCMASTER Robert B., **NOBLE** Petra, **MASTER** Robert B. Mc, 2005, « The U.S. National Historical Geographic Information System », *Historical Geography*, 33.

MENSCH Gerhard, 1980, *Stalemate in Technology: Innovations Overcome the Depression*, Cambridge, Mass, Ballinger Pub Co, 241 p.

MERCIER Aurélie, 2008, *Accessibilité et évaluation des politiques de transport en milieu urbain : le cas du tramway strasbourgeois*, phdthesis, Université Lumière - Lyon II.

MERENNE Emile, 2014, *Géographie des transports : Contraintes et enjeux*, Rennes, PU Rennes, 327 p.

MERLIN Pierre, 2007, *L'aménagement du territoire*, Paris, PUF, 448 p.

MILLER Harvey J., **GOODCHILD** Michael F., 2014, « Data-driven geography », *GeoJournal*, 80(4), p. 449‑461.

MONTOUT Sylvie, **ROBIN** Fany, 2013, « La localisation des investissements des pays émergents en Europe », *Revue française d'économie*, me XXVII(3), p. 71‑104.

MORILLAS-TORNÉ Mateu, 2012, « Creation of a Geo-Spatial Database to Analyse Railways in Europe (1830-2010). A Historical GIS Approach », *Journal of Geographic Information System*, 4(2), p. 176‑187.

MORRIS J. M., **DUMBLE** P. L., **WIGAN** M. R., 1979, « Accessibility indicators for transport planning », *Transportation Research Part A: General*, 13(2), p. 91‑109.

MOTTE Claude, **SEGUY** Isabelle, **THERE** Christine, 2003, *Communes d'hier, communes d'aujourd'hui: les communes de la France métropolitaine, 1801-2001: dictionnaire d'histoire administrative*, INED.

MUSSO Pierre, 1987, « Aux origines du concept moderne : corps et réseau dans la philosophie de Saint Simon », *Quaderni*, 3(1), p. 11‑29.

MUSSO Pierre, 1999, *Saint Simon et le Saint-Simonisme*, Presses Universitaires de France - PUF, Que sais-je ?, 128 p.

NIIZATO Takayuki, **SHIRAKAWA** Tomohiro, **GUNJI** Yukio-Pegio, 2010, « A model of network formation by Physarum plasmodium: interplay between cell mobility and morphogenesis », *Biosystems*, 100(2), p. 108–112.

NIJKAMP P., 1986, « Infrastructure and regional development: A multidimensional policy analysis », *Empirical Economics*, 11(1), p. 1‑21.

NIJKAMP Peter, **DE GRAAFF** Hans, **SIGAR** Edward, 1981, « A multidimensional analysis of regional infrastructure and economic development », p. 30.

NOUCHER Matthieu, **GAUTREAU** Pierre, 2013, « Le libre accès rebat-il les cartes ? », *Les Cahiers du numérique*, 9(1), p. 57‑83.

OFFNER Jean-Marc, 1993a, « Les «effets structurants» du transport: mythe politique, mystification scientifique », *Espace géographique*, 22(3), p. 233–242.

OFFNER Jean-Marc, 1993b, « Le développement des réseaux techniques : un modèle générique », *Flux*, 9(13), p. 11‑18.

OFFNER Jean-Marc, 1996, « " Réseaux" et" Large Technical System": concepts complémentaires ou concurrents? », *Flux*, 12(26), p. 17–30.

OFFNER Jean-Marc, 2014, « Les "effets structurants" du transport : vingt ans après », *L'Espace géographique*, 43(1), p. 51‑67.

OGBORN Miles, 1999, « The relations between geography and history: work in historical geography in 1997 », *Progress in Human Geography*, 23(1), p. 97‑108.

OLLIVRO Jean, 2000, *L'homme à toutes vitesses: de la lenteur homogène à la rapidité différenciée*, Rennes, Presses universitaires de Rennes, 180 p.

OLLIVRO Jean, 2009, « Celui qui court plus vite avance-t-il davantage ? », *in De l'histoire des transports à l'histoire de la mobilité*, Rennes, Presses universitaires de Rennes, p. 332.

OPENSHAW Stan, **ABRAHART** Robert J., 1996, *GeoComputation*, CRC Press, New York, Taylor & Francis, 443 p.

OPENSHAW Stanley, **TAYLOR** Peter, 1979, « A million or so correlation coefficients », *Statistical methods in the spatial sciences*, p. 127–144.

O'SULLIVAN David, 2005, « Geographical information science: time changes everything », *Progress in Human Geography*, 29(6), p. 749.

OTTAVIANO Gianmarco IP, **PINELLI** Dino, 2006, « Market potential and productivity: evidence from Finnish regions », *Regional Science and Urban Economics*, 36(5), p. 636–657.

PARENT Christine, **SPACCAPIETRA** Stefano, **ZIMANYI** Esteban, **DONINI** Pier, **PLAZANET** Corinne, **VANGENOT** Christelle, **ROGNON** Nadia, **POULIOT** J., **CRAUSAZ** P.-A., 1997, « MADS: un modèle conceptuel pour des applicationsspatio-temporelles », *Revue internationale de géomatique*, 7(LBD-ARTICLE-1997-002), p. 317–352.

PARROCHIA Daniel, 2005, « Quelques aspects historiques de la notion de réseau », *Flux*, n° 62(4), p. 10‑20.

PASSERON Jean-Claude, 1991, *Le raisonnement sociologique: un espace non poppérien de l'argumentation*, Paris, France, Nathan, 671 p.

PELLETIER Monique, 2013, *Les Cartes de Cassini, La science au service de l'Etat et des provinces*, Paris, France, Editions du CTHS, 383 p.

PERRET Julien, **GRIBAUDI** Maurizio, **BARTHELEMY** Marc, 2015, « Roads and cities of 18th century France », *Scientific Data*, 2, p. 150048.

PERROUX François, 1964, *L'économie du XXe siècle*, Paris, Presses Universitaire de France.

PEUQUET Donna J., 2002, *Representations of Space and Time*, Guilford Press, 396 p.

PICKLES John, 1995, *Ground truth: The social implications of geographic information systems*, Guilford Press, 251 p.

PLASSARD Fran\ccois, 1977, *Les autoroutes et le développement régional*.

PLASSARD François, 2003, *Transport et territoire*, Paris, La Documentation Fran\ccaise, Paris, 97 p.

PORTER Catherine, **ATKINSON** Paul, **GREGORY** Ian, 2015, « Geographical Text Analysis: A new approach to understanding nineteenth-century mortality », *Health & Place*, 36, p. 25‑34.

PROULX Marie-Josée, **LARRIVEE** Suzie, **BEDARD** Yvan, 2002, « Représentation multiple et généralisation avec UML et l'outil Perceptory », *in Généralisation et représentation multiple*, Hermès, Londres, p. 113–139.

PUMAIN Denise, 1982, « Chemin de fer et croissance urbaine en France au XIXe siècle », *Annales de Géographie*, 91(507), p. 529‑550.

PUMAIN Denise, 1997, « Pour une théorie évolutive des villes », *Espace géographique*, 26(2), p. 119‑134.

PUMAIN Denise, 1998, « Les modèles d'auto-organisation et le changement urbain », *Cahiers de géographie du Québec*, 42(117), p. 349–366.

PUMAIN Denise, 2003, « Une approche de la complexité en géographie », *Géocarrefour*, 78(1), p. 25‑31.

PUMAIN Denise, 2005, « Cumulativité des connaissances », *Revue européenne des sciences sociales. European Journal of Social Sciences*, XLIII-131, p. 5‑12.

PUMAIN Denise, 2014, « Les effets structurants ou les raccourcis de l'explication géographique », *L'Espace géographique*, 43(1), p. 51‑67.

PUMAIN Denise, **BRETAGNOLLE** Anne, **DEGORGE-LAVAGNE** Melina, 1999, « La ville et la croissance urbaine dans l'espace-temps », *Mappemonde*, 55(3), p. 38–42.

PUMAIN Denise, **LEPETIT** Bernard, 1993, *Temporalités urbaines*, Paris, Economica, 317 p.

PUMAIN Denise, **SAINT-JULIEN** Thérèse, 2001, *Les interactions spatiales*, Paris, Armand Colin, Cursus, 191 p.

PUMAIN Denise, **SAINT-JULIEN** Thérèse, 2010, *Analyse spatiale: Les localisations*, Paris, Armand Colin, 167 p.

QUEYROI François, 2013, *Partitionnement de grands graphes : mesures, algorithmes et visualisation*, Ph.D. thesis, Bordeaux, Bordeaux I, 154 p.

QUINE Willard Van Orman, 1977, *Relativité de l'ontologie et autres essais*, Aubier-Montaigne, Paris, 187 p.

QUINET Emile, **PRAGER** Jean-Claude, 2013, « Les effets des infrastructures sur la répartition spatiale des population et des emplois », Paris, France, Commissariat général à la stratégie et à la prospective.

RAFFESTIN Claude, 1980, *Pour une géographie du pouvoir*, LITEC, Paris, 250 p.

RAYNAUD Dominique, 2003, « Sociologie des controverses scientifiques »,.

REVERDY Georges, 1993, *L'histoire des routes de France du XIXe siècle*, Paris, Presses de l'Ecole Nationale des Ponts et Chaussées, 263 p.

REVERDY Georges, 1995, *Histoire des routes de France*, Paris, Presses Universitaires de France, Que sais-je ?, 128 p.

REY Sergio J., 2009, « Show me the code: spatial analysis and open source », *Journal of Geographical Systems*, 11(2), p. 191‑207.

REYNAUD Alain, 1992, *Une géohistoire: la Chine des printemps et des automnes*.

RIBEILL Georges, 1984, *Les cheminots*, Paris, La Découverte.

RIBEILL Georges, 1985a, « Aspects du développement du réseau ferré français sur la longue durée. L'approche historique », *Cahier / Groupe Réseaux*, 1(1), p. 10‑25.

RIBEILL Georges, 1985b, « Aspects du développement du réseau ferré français sur la longue durée. L'approche historique », *Cahier / Groupe Réseaux*, 1(1), p. 10‑25.

RIBEILL Georges, 1993, *La révolution ferroviaire*, Paris, Belin.

RIETVELD Piet, 1989, « Infrastructure and regional development », *The Annals of Regional Science*, 23(4), p. 255‑274.

RIETVELD Piet, **VICKERMAN** Roger, 2004, « Transport in regional science: The "death of distance" is premature* », *Papers in regional science*, 83(1), p. 229–248.

RIOUX Jean-Pierre, 1989, *La Révolution industrielle, 1780-1880*, Paris, Seuil, 273 p.

ROBERT Sandrine, **COSTA** Laurent, 2009, *Guide de lecture des cartes anciennes*, Paris, Errance, 104 p.

ROBIN Jean-Michel, 2000, « Endogénéité et variables instrumentales dans les sciences sociales », INSEE Methodes, Paris.

ROSEAU Nathalie, 2012, *Aerocity : quand l'avion fait la ville*, Marseille, Parenthèses, 300 p.

ROSTOW Walt Whitman, 1962, *The process of economic growth*, New York, W.W. Norton & Co.

ROUGE Maurice-François, 1953, « L'organisation de l'espace et les réseaux », *in Eventail de l'histoire vivante. Hommage à Lucien Febvre*, Armand Colin, Paris, p. 401‑405.

ROUGE Maurice-François, **DUPUY** Gabriel, 1989, « L'organisation de l'espace et les « réseaux » », *Flux*, 5(1), p. 79‑92.

ROUSSEAU Jean-Jacques, 1762, *Emile, ou de l'Education*, Paris, France, Nicolas Bonaventure

Duchesne.

ROZENBLAT Céline, 2010, « Opening the Black Box of Agglomeration Economies for Measuring Cities' Competitiveness through International Firm Networks », *Urban Studies*, 47(13), p. 2841‑2865.

SAINT-GERAND Thierry, 2005, « Comprendre pour mesurer ... ou mesurer pour comprendre ? HBDS : pour une approche conceptuelle de la modélisation géographique du monde réel », *in Modélisations en géographie : déterminisme et complexités*, Paris, Lavoisier, p. 261‑298.

SALSBURY Stephen, 1995, « Grands réseaux techniques, modèles de développement dans le temps : l'exemple des chemins de fer et de l'électricité », *Flux*, 11(22), traduit par **Waton** E., p. 31‑42.

SANDERS L., **PUMAIN** D., **MATHIAN** H., **GUÉRIN-PACE** F., **BURA** S., 1997, « SIMPOP: A Multiagent System for the Study of Urbanism », *Environment and Planning B: Planning and Design*, 24(2), p. 287‑305.

SANDERS Lena, 2001, « Modèles en analyse spatiale: introduction », *in Modèles en analyse spatiale*, p. 17‑29.

SANTOS Milton, 1972, « DIMENSION TEMPORELLE ET SYSTÈMES SPATIAUX: DANS LES PAYS DU TIERS MONDE », *Revue Tiers Monde*, 13(50), p. 247‑268.

SASSEN Saskia, 1991, *The Global City - New York, London, Tokyo*, Princeton, N.J, Princeton University Press, 480 p.

SCHOPENHAUER Arthur, 1998, *L'art d'avoir toujours raison*, Editions Mille et une nuits, Paris, 89 p.

SCHREIBMAN Susan, **SIEMENS** Ray, **UNSWORTH** John, 2008, *A companion to digital humanities*, Londres, Wiley-Blackwell, 611 p.

SCHWARTZ R., 2015, « Digital Partnership: Combining Text Mining and GIS in a Spatial History of Sea Fishing in the United Kingdom, 1860 to 1900 », *International Journal of Humanities and Arts Computing*, 9(1), p. 36‑56.

SCHWARTZ R., **GREGORY** Ian N., **MARTI-HENNEBERG** Jordi, 2011, « History and GIS: Railways, population change, and agricultural development in late nineteenth-century Wales », *in GeoHumanities: Art, History, Text at the Edge of Place. New York: Routledge*, New York, Routledge.

SHORT Jack, **KOPP** Andreas, 2005, « Transport infrastructure: investment and planning. Policy and research aspects », *Transport Policy*, 12(4), p. 360–367.

SIERRA Alexis, 2012, « La géographie et l'histoire : l'association de deux disciplines qui s'ignorent ? », *EchoGéo*, 22.

SILVEIRA Luís Espinha da, **ALVES** Daniel, **PAINHO** Marco, **COSTA** Ana Cristina, **ALCÂNTARA** Ana, 2013, « The Evolution of Population Distribution on the Iberian Peninsula: A Transnational Approach (1877–2001) », *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(3), p. 157‑174.

SIMON Guillaume, 2009, « L'évolution de l'inventaire du patrimoine ferroviaire immobilier en

France, des années 1970 à nos jours », *In Situ. Revue des patrimoines*, 10.

SINGH Puyam S., **LYNGDOH** Rosly B., **CHUTIA** Dibyajyoti, **SAIKHOM** Victor, **KASHYAP** Bhargav, **SUDHAKAR** S., 2015, « Dynamic shortest route finder using pgRouting for emergency management », *Applied Geomatics*, 7(4), p. 255‑262.

SINGLETON Alex, 2014, « Bid Data in the Real World », Université Catholique de Louvain, 2 décembre 2014.

SMITH Barry, 2001, « Objects and their environments: from Aristotle to ecological ontology », *The life and motion of socio-economic units: GISDATA*, 8.

SOUTHALL Humphrey R., 2014, « Applying Historical GIS beyond the Academy : Four Use Cases for the Great Britain HGIS », *in Toward Spatial Humanities*, Bloomington, Indiana University Press, p. 205.

STATHOPOULOS Nikolas, 1997, *La performance territoriale des réseaux de transport*, Paris, Presses de l'École Nationale des Ponts et Chaussées, 228 p.

STEINIGER Stefan, **BOCHER** Erwan, 2009, « An overview on current free and open source desktop GIS developments », *International Journal of Geographical Information Science*, 23(10), p. 1345–1370.

STEINIGER Stefan, **HUNTER** Andrew J. S., 2013, « The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption », *Computers, Environment and Urban Systems*, 39, p. 136‑150.

STUDENY Christophe, 1995, *L'invention de la vitesse: France, XVIIIe-XXe siècle*, Paris, Gallimard, 412 p.

STUDENY Christophe, 2009, « La révolution des transports et l'accélération de la France (1770-1870) », *in De l'histoire des transports à l'histoire de la mobilité*, Rennes, Presses universitaires de Rennes, p. 332.

SWENSON David, **EATHINGTON** Liesl, **OTTO** Daniel, 1998, « Economic Growth, Property Valuation Change, and Transportation Investments », *Crossroad 2000 papers and proceedings, Iowa Department of Transportation and Iowa State University*.

TERO Atsushi, **KOBAYASHI** Ryo, **NAKAGAKI** Toshiyuki, 2006, « Physarum solver: a biologically inspired method of road-network navigation », *Physica A: Statistical Mechanics and its Applications*, 363(1), p. 115–119.

TERO Atsushi, **TAKAGI** Seiji, **SAIGUSA** Tetsu, **ITO** Kentaro, **BEBBER** Dan P., **FRICKER** Mark D., **YUMIKI** Kenji, **KOBAYASHI** Ryo, **NAKAGAKI** Toshiyuki, 2010, « Rules for biologically inspired adaptive network design », *Science*, 327(5964), p. 439–442.

THEVENIN Thomas, 2010, *Transport, espace, temps: regard croisé entre time geography et géo-histoire. Mémoire de synthèse*, Ph.D. thesis, Dijon, Université de Bourgogne, 180 p.

THÉVENIN Thomas, **MIMEUR** Christophe, **SCHWARTZ** Robert, **SAPET** Loïc, 2016, « Measuring one century of railway accessibility and population change in France. A historical GIS approach », *Journal of Transport Geography*, 56, p. 62–76.

THÉVENIN Thomas, **SCHWARTZ** Robert, **SAPET** Loïc, 2013, « Mapping the Distortions in Time and

Space: The French Railway Network 1830–1930 », *Historical Methods*, 46(3), p. 134‑143.

TIRADO Daniel A., **PALUZIE** Elisenda, **PONS** Jordi, 2002, « Economic integration and industrial location: the case of Spain before World War I », *Journal of Economic Geography*, 2(3), p. 343‑363.

TO W. M., 2016, « Centrality of an Urban Rail System », *Urban Rail Transit*, 1(4), p. 249‑256.

TOBLER Waldo R., 1970, « A computer movie simulating urban growth in the Detroit region », *Economic geography*, 46, p. 234–240.

TOPALOV Christian, 2015, « La boîte de Pandore et l'histoire sociale des sciences humaines », *Genèses*, 100‑101, p. 238‑246.

TRANOS Emmanouil, 2012, « The causal effect of the internet infrastructure on the economic development of European city regions », *Spatial Economic Analysis*, 7(3), p. 319–337.

TUKEY John, 1977, *Exploratoru Data Analysis*, Addison-Wesley, Redding, Californie, 688 p.

TURNOCK David, 1979, « The Romanian railway debate: a theme in political geography », *The Journal of Transport History*, 5(2), p. 105.

UNWIN David, **TATE** Nicholas, **FOOTE** Kenneth, 2012, *Teaching Geographic Information Science and Technology in Higher Education*, John Wiley & Sons, 497 p.

UNY Daniel, 2014, « Pourquoi pas un petit regard dans le rétroviseur ...? »,.

VAN DER HERTEN Bart, **VAN MEERTEN** Michelangelo, **VERBEURGT** Greta, 2001, *Le temps du train: 175 ans de chemins de fer en Belgique, 75e anniversaire de la SNCB*, Presses Univ. de Louvain.

VANDENBULCKE Grégory, **STEENBERGHEN** Thérèse, **THOMAS** Isabelle, 2009, « Mapping accessibility in Belgium: a tool for land-use and transport planning? », *Journal of Transport Geography*, 17(1), p. 39‑53.

VANHAUTE Eric, 2005, « The Belgium Historical GIS », *Historical Geography*, 33, p. 136–139.

VARLET Jean, 1990, « Réseau ferré et relations interrégionales en France », *Revue d'histoire des chemins de fer*, 2, p. 87‑101.

VELTZ Pierre, 1996, *L'économie d'archipel*, Paris, Editions de l'Aube.

VERDIER Nicolas, **BRETAGNOLLE** Anne, 2007, « L'extension du réseau des routes de poste en France, de 1708 à 1833 », *Histoire des réseaux postaux en Europe du XVIIIe au XXIe siècle*, p. 155–193.

VESCHAMBRE Vincent, 2007, « Patrimoine : un objet révélateur des évolutions de la géographie et de sa place dans les sciences sociales », *Annales de géographie*, 656, p. 361‑381.

VICKERMAN R. W., 1995, « The regional impacts of Trans-European networks », *The Annals of Regional Science*, 29(2), p. 237‑254.

VIRILIO Paul, 1977, *Vitesse et politique : Essai de dromologie*, Paris, Editions Galilée, 151 p.

VIRILIO Paul, 1995, *La vitesse de libération*, Paris, Galilée, 179 p.

VOLVEY Anne, **BRETAGNOLLE** Anne, **DJAMENT** Géraldine, **HOYAUX** André-Frédéric, **VRAC** Michel, 2005, *Echelles et temporalités*, Paris, Atlande, 240 p.

VON THÜNEN Johann Heinrich, 1826, *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalokönomie*, Düsseldorf, Wissenschaft und Finanzen GmhH.

VRANA Ric, 1989, « Historical data as an explicit component of land information systems », *International journal of geographical information systems*, 3(1), p. 33‑49.

WARF Barney, **ARIAS**, 2008, *The Spatial Turn: Interdisciplinary Perspectives*, Routledge, Londres, 231 p.

WEBER Max, 1965, *Essais sur la théorie de la science*, Plon, Paris.

WHITE Richard, 2010, « What is Spatial History ? », *Spatial History Lab, Stanford University Spatial History*, p. 1‑6.

WILSON John P., **FOTHERINGHAM** A. Stewart, 2008, *The handbook of geographic information science*, John Wiley & Sons.

WOLKOWITSCH Maurice, 1980, « Réseau ferré et structuration de l'espace français », *Eventail de la spatiologie*, p. 103‑121.

WOLKOWITSCH Maurice, 2004, « Conclusion », *Revue d'histoire des chemins de fer*, 30, p. 465‑470.

WYNNE Graerme, 2003, « Review of H.C. Darby », *H-Environment*.

XIE Feng, **LEVINSON** David, 2008, « The weakest link: The decline of the surface transportation network », *Transportation Research Part E: Logistics and Transportation Review*, 44(1), p. 100–113.

XIE Feng, **LEVINSON** David, 2009, « Modeling the growth of transportation networks: a comprehensive review », *Networks and Spatial Economics*, 9(3), p. 291–307.

XIE Feng, **LEVINSON** David, 2011, *Evolving transportation networks*, Springer Science & Business Media.

XUONG Nguyen Huy, 1992, *Mathématiques discrètes et informatique*, Paris, Masson, 412 p.

YERRA Bhanu M., **LEVINSON** David M., 2005, « The emergence of hierarchy in transportation networks », *The Annals of Regional Science*, 39(3), p. 541–553.

1992, « Géographie et histoire », *Encyclopédie de géographie*, p. 1132.

2000, *Barrington Atlas of the Greek and Roman World*, Princeton, New Jersey, Princeton University Press, 280 p.

2014, « Les controverses de L'Espace géographique », *L'Espace géographique*, Tome 40(1), p. 76‑93.

TABLE DES MATIERES

INTRODUCTION GENERALE 10

Réseau et territoire : une dualité d'actualité différentiée par la vitesse 11
Trace et tracé dans une approche par la donnée géohistorique 12
La base de données FRANcE au cœur des humanités numériques 14
Le cheminement de la thèse : une modélisation croissante 15

PREMIERE PARTIE : EFFET STRUCTURANT ET ECHELLES SPATIO-TEMPORELLES 18

INTRODUCTION DE LA PREMIERE PARTIE 19
CHAPITRE 1. L'EFFET STRUCTURANT DES INFRASTRUCTURES DE TRANSPORT : DE L'ENGOUEMENT A LA CONTROVERSE 22
Introduction 22
1.1. Les vertus de la mise en réseau 23
1.1.1. Le saint-simonisme : structurer par la connexité 23
1.1.2. L'unité nationale : structurer le contrôle politique 27
1.1.3. Du mythe à la théorie des réseaux 30
1.2. Qualifier les effets des infrastructures de transport 36
1.2.1. Infrastructure et développement : un lien global et peu spatialisé 37
a. Une nouvelle infrastructure de transport a-t-elle un effet significatif dans le développement économique ? .. 38
b. Une nouvelle infrastructure de transport a-t-elle des effets homogènes sur un territoire ? 39
1.2.2. Un rôle croissant de l'espace, mais sans géographes 42
a. Une nouvelle infrastructure de transport génère-t-elle des effets de proximité ? 43
b. La nouvelle infrastructure génère-t-elle des effets de dispersion et de nouvelles inégalités ? 45
1.2.3. Les enseignements de la profondeur temporelle 47
a. Quelle place dans l'économie historique ? 47
b. Quelle place dans la géographie et l'histoire ? 49
1.3. De la remise en cause à la controverse 51
1.3.1. Une mainmise de l'aménagement du territoire 52
1.3.2. Acteurs et échelles dans la complexification du territoire 54
1.3.3. Une controverse de permanente actualité scientifique 57
Conclusion 62
CHAPITRE 2. LA CONSTRUCTION DE NOUVELLES ECHELLES DANS UNE PERSPECTIVE GEOHISTORIQUE 64
Introduction 64
2.1. Une géohistoire pour des réflexions à de larges échelles 65
2.1.1. Les vicissitudes des liens entre histoire et géographie, et pourtant … 65
2.1.2. La construction de systèmes spatiaux dans la longue durée 68
2.1.3. La longue durée, le parent pauvre de la géographie des transports ? 73
2.2. Le « spatial turn », entre questionnement et instrument 76
2.2.1. Les humanités à l'heure du *« digital turn »* et du *« spatial turn »* 76
2.2.2. SIG-H : de l'information géographique à l'information spatio-temporelle 81
2.2.3. Les projets de SIG-H : expériences et enjeux 84
2.3. Vers la définition de cadres d'exploration du lien réseau/territoire 87
2.3.1. Du mythe à la science, du positivisme au constructivisme 87
2.3.2. Quantifier le lien entre réseau et territoire : des critères de précision entre convention et mesure 91
2.3.3. Des cadres de validité dans des critères de pertinence 95
Conclusion 100
CHAPITRE 3. LES TRACES DU LIEN RESEAU / TERRITOIRE PAR LA BASE DE DONNEES FRANCE 102
Introduction 102

3.1. L'information géohistorique face à l'hétérogénéité des sources *103*
3.1.1. L'information démographique, un « classique mondial » 103
3.1.2. L'information ferroviaire, reflet de la complexité du système ferroviaire 106
3.1.3. La construction du modèle conceptuel PONT 111
3.1.4. La vitesse dans FRANcE 114
3.2. De l'exploitation à la complexification de l'information *119*
3.2.1. La mise en connexité du réseau ferroviaire dans la base de données FRANcE 119
3.2.2. Le cycle de vie du réseau ferroviaire dans la base de données FRANcE 123
3.2.3. Se déplacer en train : quelle approche pour le déplacement dans la base de données FRANcE 128
Conclusion *131*
CONCLUSION DE LA PREMIERE PARTIE 132

DEUXIEME PARTIE : L'ACCESSIBILITE DANS UN CONTEXTE GEOHISTORIQUE : UNE MODELISATION CROISSANTE DE LA DONNEE, DE L'EFFET DU RESEAU A SA DYNAMIQUE DE CROISSANCE 136

INTRODUCTION DE LA DEUXIEME PARTIE 137
CHAPITRE 4. CAPITALISER L'INFORMATION GEOHISTORIQUE : ENTRE PRECISION ET GENERALISATION 140
Introduction *140*
4.1. Introduire la base de données FRANcE dans la Geocomputation *141*
4.1.1. Des enjeux de la science de l'information géographique … 141
4.1.2. … à la *GeoComputation* dans une perspective géohistorique 144
4.1.3. Une chaîne géohistorique dans un couplage modéré 147
4.2. La théorie des graphes pour le réseau ferroviaire *151*
4.2.1. Les principes de la modélisation des réseaux de transport 151
4.2.2. Du SIG-H au graphe : processus de simplification et d'abstraction 154
4.2.3. Des indices globaux aux indices locaux dans une démarche géohistorique 160
4.3. L'accessibilité des 36 000 communes dans FRANcE : entre précision et généralisation *164*
4.3.1. Cheminer vers la gare 164
4.3.2. Modéliser le cheminement pédestre : quelle généralisation du réseau ? 167
4.3.3. Les différentiels de réseaux dans un multigraphe 172
Conclusion *176*
CHAPITRE 5. MESURER DES RELATIONS DE LONG TERME ENTRE INFRASTRUCTURE ET TERRITOIRE 178
Introduction *178*
5.1. Caractériser la dotation en réseau par l'accessibilité *179*
5.1.1. Des indicateurs fonctionnels … 179
5.1.2. … aux indicateurs multipolaires 183
5.2. Un rapprochement vers l'économie historique *191*
5.2.1. L'économie historique : entre cliométrie et empirisme 191
5.2.2. Précautions et méthodes pour modéliser la croissance par le réseau 194
5.2.3. Les grands principes de l'économétrie de panel 197
5.3. Mobiliser des informations géohistoriques de long terme pour la base de données FRANcE *201*
5.3.1. Intégrer des variables de contexte local 201
5.3.2. Les « voies » de l'instrumentation : principes et stratégies 209
5.3.3. Route et réseaux historiques pour instrumenter FRANcE 211
Conclusion *217*
CHAPITRE 6. MODELISER LES INTERACTIONS DANS UN CONTEXTE GEOHISTORIQUE 220
Introduction *220*
6.1. La modélisation des interactions entre transport et territoire *221*
6.1.1. Un nouveau degré dans la modélisation des interactions géohistoriques 221
6.1.2. Les familles de modélisation de la croissance des réseaux 224
a. Les approches traditionnelles des interactions entre transport et espace 226
b. La modélisation de l'optimisation dans la science des réseaux 227
6.2. La construction d'un modèle de croissance d'un réseau de transport pour FRANcE *230*
6.2.1. La définition et le positionnement du modèle 231
6.2.2. L'initialisation du modèle : d'une situation préexistante à l'interaction spatiale 234
6.2.3. Le module dynamique pour générer le réseau de transport 238

6.3. Explorer le modèle d'évolution de réseau de transport ... *240*
6.3.1. Une exploration théorique du modèle par des espaces stylisés ... 241
6.3.2. De la structure à la hiérarchie de l'évolution du réseau de transport dans les espaces stylisés ... 243
a. Le modèle face à une distribution homogène ... 244
b. Le modèle face à une distribution centre/périphérie ... 247
c. Le modèle face à une distribution polycentrique ... 250
6.3.3. Vers une définition des paramètres pour une exécution empirique du modèle ... 254
Conclusion ... *256*
CONCLUSION DE LA DEUXIEME PARTIE ... 257

PARTIE 3 : GEOHISTOIRE DE L'INTERACTION ENTRE RESEAU ET POPULATION : DE LA MESURE DE L'EFFET DU RESEAU A LA SIMULATION DE SA CROISSANCE ... 262

INTRODUCTION DE LA TROISIEME PARTIE ... 263
CHAPITRE 7. DE L'ACCELERATION DU RESEAU A LA DEFINITION DES EFFETS SUR LES DYNAMIQUES DEMOGRAPHIQUES ... 266
Introduction ... *266*
7.1. Expansion et accélération : une réelle homogénéisation ? ... *267*
7.1.1. L'ouverture de l'horizon spatio-temporel : continuités et discontinuités des déformations de l'espace ... 267
7.1.2. Une accélération polynucléaire ? ... 273
7.2. Les effets structurants du réseau sur la structure démographique ... *278*
7.2.1. Accessibilité et niveaux de population : une coévolution ... 279
7.2.2. Centralité et niveau de population : une double coévolution ... 285
7.2.3. La double hiérarchisation du réseau face à la croissance démographique locale ... 287
7.3. Portée spatiale et temporelle de l'effet du réseau ... *293*
7.3.1. Effets directs et indirects sur la croissance de long terme : une homogénéisation ? ... 293
7.3.2. Une portée hétérogène des réseaux dans l'espace et dans le temps ... 297
7.3.3. Vers l'instrumentation des effets directs et indirects du réseau ... 303
Conclusion ... *309*
CHAPITRE 8. DES CAPACITES CIRCULATOIRES A LA SIMULATION DE LA CROISSANCE DU RESEAU FERROVIAIRE ... 312
Introduction ... *312*
8.1. L'émergence géohistorique de corridors par les capacités circulatoires du réseau ... *313*
8.1.1. Les capacités circulatoires sur le temps long : hiérarchisation et concentration ... 313
8.1.2. Gains et pertes d'intermédiarité sur le temps long ... 317
8.1.3. Corridors géohistoriques et niveaux de population ... 327
8.2. De l'efficacité radiale à la suggestion de transversales ... *332*
8.2.1. La structure globale du réseau : une prise en compte directe mais partielle de la structure urbaine ... 333
8.2.2. L'émergence de carrefours intermédiaires pour desservir la population ... 338
8.2.3. De potentielles transversales dans un arbre hiérarchisé ... 342
8.3. Enrichir les comparaisons par des explorations locales ... *346*
8.3.1. Le réseau du Nord : entre frontière, littoral et bassin industriel ... 346
8.3.2. Le réseau du PLM : la force du politique pour éviter le Morvan ... 351
Conclusion ... *358*
CHAPITRE 9. LA BASE DE DONNEES FRANCE, UN PROJET ENTRE ELARGISSEMENT ET APPROFONDISSEMENT. 360
Introduction ... *360*
9.1. L'élargissement : une théorie des réseaux « à la française » ? ... *361*
9.1.1. Elargir le cadre spatial ... 361
9.1.2. Elargir le cadre temporel ... 366
9.1.3. Questionner la permanence du concept de réseau ... 370
9.2. L'approfondissement : méthodes, outils et publics ... *375*
9.2.1. Un projet pour le partage et la valorisation du patrimoine ferroviaire ... 375
9.2.2. Un projet dans la mixité des méthodes ... 379
9.2.3. Pour un dialogue plus intégré dans les humanités ... 384
Conclusion ... *390*
CONCLUSION DE LA TROISIEME PARTIE ... 391

CONCLUSION GENERALE 394

Une trajectoire de l'accumulation à la capitalisation 395
Le géographe, et l'informaticien, l'économiste, l'historien… 396
Le géographe, un vecteur pour penser le réseau par l'interaction spatiale 398

BIBLIOGRAPHIE 402

TABLE DES MATIERES 428

TABLE DES FIGURES 432

TABLE DES TABLEAUX 436

ANNEXES 438

TABLE DES FIGURES

Figure 1. 1. La distance au chemin de fer en Europe dans le premier tiers du XXème siècle 26
Figure 1. 2. Approche croisée du passage du réseau virtuel au réseau réel, du réseau technique au réseau territorial 33
Figure 1. 3. Les cycles de réseaux : plus d'accumulation que de destruction … 34
Figure 1. 4. Le schéma tendanciel des réseaux techniques 35
Figure 1. 5. Infrastructures de transport et développement régional Approche typologique et temporelle 41
Figure 1. 6. Les modèles d'organisation de l'espace, base théorique de l'étude des effets structurants 43
Figure 1. 7. Les effets des infrastructures de transport : du mythe à la controverse 61

Figure 2. 1. Echelles de temporalités 69
Figure 2. 2. De la construction de la démarche géohistorique au processus de construction de connaissances géohistoriques 71
Figure 2. 3. GeoHumanities 79
Figure 2. 4. Les Spatial Humanities dans la géohistoire contemporaine 80
Figure 2. 5. La matrice d'information géographique 83
Figure 2. 6. Le projet du SIG-H : dépasser la collecte pour l'analyse 86
Figure 2. 7. Le processus de construction de connaissances géohistoriques 90
Figure 2. 8. Evolution du réseau ferroviaire français entre 1860 et 1930 93
Figure 2. 9. Evolution de la vitesse moyenne de déplacement entre 1815 et 2000 94
Figure 2. 10. La scène de l'aménagement du territoire face à notre enjeu de recherche 97
Figure 2. 11. Les trois dimensions des phénomènes sociaux 98

Figure 3. 1. Trois exemples de représentation spatio-temporelle dans un SIG 105
Figure 3. 2. L'information démographique dans la base de données FRANcE 106
Figure 3. 3. Sources primaires et secondaires pour la spatialisation du réseau dans le SIG-H 108
Figure 3. 4. De la source au SIG-H : sources pour la chronologie du réseau 109
Figure 3. 5. L'évolution du réseau ferroviaire français en fonction du type de voie entre 1840 et 2000 110
Figure 3. 6. L'information ferroviaire dans la base de données FRANcE 110
Figure 3. 7. Le modèle PONT dans la base de données FRANcE 114
Figure 3. 8. La vitesse dans la base de données FRANcE 117
Figure 3. 9. La vitesse dans le modèle PONT 118
Figure 3. 10. De l'initiative ponctuelle à la mise en connexité du réseau ferroviaire français entre 1842 et 1860 121
Figure 3. 11. Différenciation des vitesses en 1860 selon les territoires des grandes compagnies 122
Figure 3. 12. Estimation de la courbe tendancielle du réseau ferroviaire entre 1840 et 1930 124
Figure 3. 13. Evolution de la densité ferroviaire entre 1860 et 1930 125
Figure 3. 14. Estimation des courbes tendancielles du réseau ferroviaire entre 1840 et 1930 selon la classe de vitesses 127

Figure 3. 15. La complexification des contraintes de déplacement pour mesurer l'accessibilité dans la base de données FRANcE 130

Figure 4. 1. Du SIG utilisateur au SIG serveur 147
Figure 4. 2. Trois intensités de couplage entre SIG et outils d'analyse 148
Figure 4. 3. Une chaîne géo-méthodologique dans la GeoComputation 150
Figure 4. 4. A l'origine de la théorie des graphes … 152
Figure 4. 5. Graphe planaire 155
Figure 4. 6. Du SIG-H au graphe G 157
Figure 4. 7. Du graphe ressemblant au graphe fonctionnel 158
Figure 4. 8. Un sous-graphe de G' de 1880 159
Figure 4. 10. La centralité d'intermédiarité des liens dans la théorie des graphes 163
Figure 4. 11. Une approche classique par l'accessibilité directe 165
Figure 4. 12. La fonction de déplacement pour mesurer l'accessibilité piétonne aux gares ferroviaires 166
Figure 4. 13. L'application des contraintes en fonction de la pente définie par le MNT 168
Figure 4. 14. Un réseau en espalier par les chemins de moindre coût sur le MNT 168
Figure 4. 15. Comparaison des chemins optimaux vers Dijon 170
Figure 4. 16. La triangulation des 36 000 communes françaises : approche comparée avec BD CARTO 171
Figure 4. 17. Le multigraphe F pour les déplacements dans FRANcE 174
Figure 4. 18. Temps de parcours depuis Perros-Guirec jusqu'à Paris de 1860 à 1930 175
Figure 4. 19. Les contours de l'instrument de recherche pour la base de données FRANcE 176

Figure 5. 1. Temps de parcours vers la gare la plus proche en 1860 et en 1930 181
Figure 5. 2. Accessibilité fonctionnelle en 1900 182
Figure 5. 3. Accessibilité multipolaire en 1860, 1880, 1900 et 1920 185
Figure 5. 4. Dispersion des temps de parcours moyens et pondérés 186
Figure 5. 5. L'effet vitesse en 1860, 1880, 1900 et 1920 188
Figure 5. 6. La construction du market potential pour la base de données FRANcE 190
Figure 5. 7. Des moindres carrés ordinaires aux données de panel 198
Figure 5. 8. L'hétérogénéité inobservée dans la modélisation en données de panel 199
Figure 5. 9. Les voies de communication issues de la Carte de Cassini 204
Figure 5. 10. La navigation intérieure depuis le XVIIème siècle 205
Figure 5. 11. Les gisements houillers sur le territoire français 206
Figure 5. 12. Les principales voies romaines entre 550 av. JC et 640 ap. JC 212
Figure 5. 13. Le multigraphe ROM 214
Figure 5. 14. Le réseau postal en 1833 215
Figure 5. 15. Le multigraphe POS 216
Figure 5. 16. Les degrés de mobilisation de la base de données pour étudier l'effet du réseau 218

Figure 6. 1. La modélisation dynamique des interactions entre réseau de transport et territoire – Carte d'identité des approches et application aux cinq tendances dans la modélisation des réseaux de transport 225
Figure 6. 2. La reproduction in-situ du réseau métropolitain de Tokyo à l'aide du Physarum Plasmodium 229

Figure 6. 3. Le processus de modélisation dynamique du réseau de transport ferroviaire 231
Figure 6. 4. La carte d'identité du modèle d'évolution du réseau ferroviaire 234
Figure 6. 5. La phase d'initialisation du modèle : entre graphe théorique et empirique 235
Figure 6. 6. Modéliser le besoin existant dans le modèle d'évolution du réseau de transport 237
Figure 6. 7. Le modèle d'évolution du réseau de transport : du modèle statique au modèle dynamique 240
Figure 6. 8. Trois espaces stylisés pour l'expérimentation du modèle 243
Figure 6. 9. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution homogène 244
Figure 6. 10. Génération du réseau de transport avec α = 0,5 245
Figure 6. 11. Distribution des flux selon le nombre d'itérations avec α = 0,5 246
Figure 6. 12. Génération du réseau de transport après 8 itérations 247
Figure 6. 13. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution centre/périphérie 248
Figure 6. 14. Génération du réseau de transport dans un espace monocentrique après 8 itérations avec α = 0,5 249
Figure 6. 15. Génération du réseau de transport dans un espace monocentrique après 20 itérations 250
Figure 6. 16. Confrontation des flux sur le réseau et des flux distribués par le modèle gravitaire avec une distribution polycentrique 251
Figure 6. 17. Génération du réseau de transport dans un espace polycentrique après 10 itérations avec α = 0,5 252
Figure 6. 18. Génération du réseau de transport dans un espace polycentrique après 10 itérations 253
Figure 6. 19. Une signification géohistorique des paramètres du modèle d'évolution du réseau de transport 255
Figure 6. 20. Le plan d'expérimentation pour l'étude des interactions entre réseau et population 259

Figure 7. 1. Evolution des distorsions spatiales par l'accessibilité à Paris entre 1860 et 1930 268
Figure 7. 2. La trajectoire des villes françaises dans la distorsion de l'espace fonctionnel pour rejoindre Paris entre 1860 et 1930 270
Figure 7. 3. La trajectoire des villes dans la distorsion multipolaire de l'espace fonctionnel entre 1860 et 1930 272
Figure 7. 4. Une comparaison à une situation contrefactuelle sur le temps et à trois échelles 274
Figure 7. 5. Evolution des rythmes d'accélération selon l'échelle et le temps face à la construction de voies nouvelles selon le type de voies 275
Figure 7. 6. La contribution de la hiérarchie du réseau aux gains d'accessibilité aux trois échelles en 1910 276
Figure 7. 7. La diffusion de la desserte des populations selon quatre indicateurs d'accessibilité entre 1860 et 1930 279
Figure 7. 8. La confrontation des indicateurs d'accessibilité en 1910 280
Figure 7. 9. Les niveaux de population de villes supérieures à 2 500 habitants en fonction des quartiles d'accessibilité 282
Figure 7. 10. Les niveaux de population des communes inférieures à 2 500 habitants en fonction des quartiles d'accessibilité 284
Figure 7. 11. Carrefours et niveaux de population pour les villes ayant plus de 2 500 habitants entre 1860 et 1930 286
Figure 7. 12. Caractériser la desserte par la vitesse et la centralité 288
Figure 7. 13. Centralité, niveau de population et croissance démographique des communes de plus de

2 500 habitants entre 1866 et 1911 290
Figure 7. 14. Effets de la mémoire de la connexion et de la centralité dans le réseau dans la modélisation en panel entre 1866 et 1911 302
Figure 7. 15. Les corridors issus des lois de1842 et 1878 305
Figure 7. 16. La nature des effets du réseau par une complexification des estimations 308

Figure 8. 1. La distribution des flux par la structure du réseau en 1860 et en 1910 315
Figure 8. 2. La distribution des flux selon la structure du réseau et les vitesses en 1870 316
Figure 8. 3. Les capacités circulatoires du réseau ferroviaire français en 1860 319
Figure 8. 4. Les capacités circulatoires du réseau ferroviaire français en 1880 322
Figure 8. 5. Les capacités circulatoires du réseau ferroviaire français en 1900 324
Figure 8. 6. Les capacités circulatoires du réseau ferroviaire français en 1930 326
Figure 8. 7. Potentiel de circulation et population entre 1870 et 1930 328
Figure 8. 8. Potentiel de circulation et structure démographique préexistante 330
Figure 8. 9. Potentiel de circulation et structure démographique préexistante pour les communes de moins de 2 500 habitants 331
Figure 8. 10. Les paramètres du modèle pour caractériser le réseau de premier ordre 334
Figure 8. 11. Simulation du réseau pour une structure de 40 000 habitants 335
Figure 8. 12. Les paramètres du modèle pour caractériser la nodalité 338
Figure 8. 13. Simulation du réseau pour une nodalité dans le réseau 339
Figure 8. 14. Les paramètres du modèle pour caractériser le réseau secondaire 342
Figure 8. 15. Simulation du réseau pour une structure plus complète 343
Figure 8. 16. De potentielles transversales dans le réseau national 345
Figure 8. 17. Les paramètres du modèle pour simuler l'infrastructure sur le territoire de la Compagnie des Chemins de Fer du Nord 347
Figure 8. 18. Simulation de l'infrastructure sur le territoire de la Compagnie des Chemins de Fer du Nord 349
Figure 8. 19. Comparaison des densités d'infrastructure entre réseau réel de 1930 et réseau simulé. 350
Figure 8. 20. Les paramètres du modèle pour simuler le réseau du PLM 352
Figure 8. 21. Simulation de l'infrastructure sur le territoire du PLM 353
Figure 8. 22. Simulation de l'infrastructure sur le territoire du PLM avec une contrainte dans le modèle 355
Figure 8. 23. Les paramètres de simulation avec une contrainte dans le modèle 355

Figure 9. 1. La succession et la diffusion des réseaux aux Etats-Unis 370
Figure 9. 2. De la croissance à la décroissance des réseaux de transport 373
Figure 9. 3. La numérisation de la desserte pour une intégration dans le modèle PONT 381
Figure 9. 4. Pour une géohistoire au cœur du dialogue dans les humanités 385

TABLE DES TABLEAUX

Table 1. 1. Les priorités de la politique publique de transport. 53

Table 2. 1. De la géographie historique à la géohistoire contemporaine … appliquée au lien réseau/territoire 69
Table 2. 2. Les nouveaux enjeux de la géographie des transports, peu confrontés à la longue durée ... 75
Table 2. 3. Du SIG au SIG-H 82

Table 3. 1. La grammaire UML dans le modèle PONT 112
Table 3. 2. La correction des erreurs de topologie dans FRANcE 116
Table 3. 3. Comparaison des temps de parcours entre les sources secondaires et la base de données FRANcE 118
Table 3. 4. La mise en connexité du réseau par classes de vitesses entre 1842 et 1860 120

Table 4. 1. L'analyse spatiale en géographie 144
Table 4. 2. La simplification du graphe en chiffres 160
Table 4. 4. Comparaison des cheminements pédestres sur le MNT et à partir de BD CARTO 169
Table 4. 5. Comparaison des cheminements pédestres à partir de BD CARTO et de la triangulation 172

Table 5. 1. La complexification de l'indicateur d'accessibilité au réseau 180
Table 5. 2. Panorama de l'utilisation des données de panel 200
Table 5. 3. Répartition des voies romaines dans la France actuelle 213

Table 7. 1. Centralité dans le réseau et croissance démographique locale des communes de moins de 2 500 habitants 291
Table 7. 2. Croissance de la population entre 1866 et 1911 et accès au réseau 294
Table 7. 3. Croissance de la population entre 1866 et 1911 et temps de parcours 295
Table 7. 4. Croissance de la population entre 1866 et 1911 et accessibilité potentielle 297
Table 7. 5. Effet de l'accessibilité potentielle entre 1866 et 1911 299
Table 7. 6. Effet de la proximité de la gare entre 1866 et 1911 300
Table 7. 7. Centralité dans le réseau entre 1866 et 1911 301
Table 7. 8. Mémoire de la connexion entre 1866 et 1911 302
Table 7. 9. Instrumentation des effets directs et indirects du réseau entre 1866 et 1911 306
Table 7. 10. Instrumentation des effets directs et indirects du réseau par les réseaux historiques entre 1876 et 1906 307

Table 9. 1. Nombre de kilomètres de voies ferrées par pays dans le corpus européen entre 1870 et 2000 362
Table 9. 2. L'usage des SIG dans la communauté « Spatial Literacy » 387
Table 9. 3. Les principales fonctions du logiciel Geographer, d'après Thévenin, 2010 389

ANNEXES

Annexe 1. Résultats économétriques

1. Complexification croissante de la modélisation

Variable dépendante : Croissance de la population 1866-1911	(1) *Accès au réseau en 1860*			(2) *Accès au réseau en 1860*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Dotation en réseau	0,0068	***	0,0003	0,0039	***	0,0003
Croissance 1856-1866				0,1246	***	0,0099
Accès aux voies navigables				0,0029	***	0,0002
Ressources minières				0,0006	***	0,0002
Accès à la route				0,0006	***	0,0001
Hiérarchie administrative				0,0044	***	0,0002
Spécialisation agricole						
Spécialisation industrielle						
Diversité sectorielle						
Constante	-0,0043	***	0,0000	-0,0051	***	0,0001
Observations	30450			30450		
R^2	**0,0371**			**0,1502**		
Effets fixes	Non			Non		

Variable dépendante : Croissance de la population 1866-1911	(3) *Accès au réseau en 1860*			(4) *Accès au réseau en 1860*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Dotation en réseau	0,0038	***	0,0003	0,0035	***	0,0002
Croissance 1856-1866	0,1293	***	0,0110	0,0736	***	0,0096
Accès aux voies navigables	0,0028	***	0,0002	0,0023	***	0,0002
Ressources minières	0,0007	***	0,0002	0,0005	***	0,0002
Accès à la route	0,0005	***	0,0001	0,0006	***	0,0001
Hiérarchie administrative	0,0043	***	0,0002	0,0036	***	0,0002
Spécialisation agricole	-0,0014	***	0,0003	0		0,0010
Spécialisation industrielle	-0,0012	***	0,0002	-0,0001		0,0003
Diversité sectorielle	0,0005	***	0,0001	0,0014	***	0,0002
Constante	-0,0045	***	0,0002	-0,0025	***	0,0005
Observations	30450			30450		
R^2	**0,1547**			**0,322**		
Effets fixes	Non			Département		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Croissance de la population entre 1866 et 1911 et accès au réseau

2. Choix de la modélisation des effets dans la structure de panel

La spécification globale du modèle propose un R² de 14 % pour l'ensemble de l'échantillon *(Colonne 1)*, confirmant la forte hétérogénéité des données, qui doit être prise en compte dans les régressions à partir de la structure en panel. Alors, le pouvoir explicatif du modèle est presque doublé[77], mais le coefficient mesurant l'effet de l'infrastructure diminue fortement, attestant d'une forte hétérogénéité inobservée, qui est désormais contrôlée par les effets individuels. Si les estimations ne dérivent pas très fortement selon que l'on assigne des effets individuels de manière fixe *(Colonne 2)* ou aléatoire *(Colonne 3)*, le test de Hausman[78] nous permet de désigner lequel est le plus approprié. Ce test, très significatif, nous pousse à privilégier le modèle à effets fixes. Alors, cette première estimation montre la significativité de l'effet de la connexion au réseau ferroviaire sur une période longue : le coefficient est de 6,9 % pour chaque période intercensitaire, qui signifie que la connexion augmente le niveau de la population de 0,6 % par an en moyenne. Dans ce modèle, on voit bien également que les départements les plus spécialisés en agriculture sont ceux les plus touchés par le phénomène de l'Exode Rural, alors que la croissance bénéficie davantage aux territoires spécialisés dans l'industrie, la prime à la diversité étant légèrement négative.

On peut aussi croiser des effets individuels à des échelles plus agrégées, qui entrent en interaction directe avec les effets temporels *(Colonne 4)* : on traduit des effets que l'on ne maîtrise pas par la donnée, telle la découverte de nouvelles ressources minérales susceptibles d'introduire un choc dans l'économie à une échelle agrégée, ou encore d'autres chocs exogènes comme les épidémies ou des phénomènes météorologiques qui affecteraient la productivité agricole. On voit que l'effet de la possession d'une gare s'en trouve légèrement affaibli, mais reste néanmoins significatif, attestant d'un réel effet dans la variabilité des niveaux de population. Dès lors, nous concluons à ce stade que la modélisation à effets fixes avec des effets sur les périodes et sur les individus peut être complexifiée par les temps de parcours moyens.

Variable dépendante : Niveau de population entre 1866-1911	**(1)** *Pool.*			**(2)** *Fixed effects*			**(3)** *Random effects*			**(4)** *Effets croisés*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accès au réseau	0,794	***	0,007	0,069	***	0,002	0,080	***	0,002	0,0586	***	0,0016
Accès aux voies navigables	0,434	***	0,008	0,012		0,009	0,117	***	0,008	0,0333	***	0,0086
Spécialisation agricole	-0,076	***	0,016	-0,221	***	0,006	-0,183	***	0,006	-0,2530	***	0,0119
Spécialisation industrielle	-0,211	***	0,014	0,105	***	0,006	0,076	***	0,006	0,0021		0,0074
Diversité sectorielle	-0,128	***	0,007	-0,086	***	0,003	-0,083	***	0,003	-0,0540	***	0,0033
Constante	6,102	***	0,011	6,260	***	0,004	6,253	***	0,006	6,2943	***	0,0043
Observations	30450			30450			30450			30450		
R²	**0,1415**			**0,2419**			**0,2401**			**0,3119**		
Effets temporels	Oui			Oui			Oui			Oui		
Effets individuels	Non			Oui			Oui			Oui		
Effets croisés	Non			Non			Non			Oui		

Source : Base de données FRANcE

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Effet de la connexion au réseau entre 1866 et 1911

[77] Il faut noter que le pouvoir explicatif possède les mêmes ordres de grandeur que ceux rencontrés par ailleurs dans la bibliographie des effets de long terme.

[78] Le test de Hausman consiste en la comparaison des variances entre les deux spécifications. En l'absence de différences significatives, le modèle à effets fixes est privilégié.

3. Effets directs et indirects du réseau dans la spécification en panel

Le modèle en panel montre une significativité de l'impact des temps de parcours moyen sur l'évolution des populations entre 1866 et 1911. Le pouvoir explicatif reste stable par rapport à la seule spécification de la présence d'une gare. Ainsi, même quand elle n'est pas desservie par une gare, une commune qui bénéficie de la vitesse ferroviaire connaît une évolution de sa population supérieure *(Colonne 1)*. Ainsi, pour un même niveau de population, la diminution de 1 % des temps de parcours signifie une augmentation de la population de 1,31 % par an en moyenne. Cet effet doit être relativisé par la prise en compte d'une hétérogénéité supplémentaire fixée à l'échelle des départements *(Colonne 2),* qui fait baisser ce coefficient à 1,04 % par an en moyenne, en restant tout autant significatif mais avec un pouvoir explicatif de 30,82 %. De la même manière que précédemment, nous distinguons les effets directs et indirects de la possession d'une gare dans la commune. Alors, l'effet marginal est significatif que l'on ajoute ou non un élément fixé au département *(Colonnes 3 et 4)* : un gain de 1 % dans les temps de parcours moyen signifie une augmentation du niveau de population de l'ordre de 0,6 % à 1 % par an en moyenne, avec une prime à la possession d'une gare qui s'élève à 0,1 % par an en moyenne. Il est intéressant de noter que cet effet marginal est stable quelle que soit la dimension de l'hétérogénéité ajoutée. Ainsi, assez logiquement, ce modèle en panel vient confirmer avec des estimations plus robustes le double effet direct et indirect du réseau, par les temps de parcours et un accès direct à l'infrastructure.

Variable dépendante : Niveau de population entre 1866-1911	**(1)**			**(2)**			**(3)**			**(4)**		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	-0,149	***	0,005	-0,117	***	0,005	-0,095	***	0,005	-0,063	***	0,006
Effet marginal "gare"							0,008	***	0,000	0,007	***	0,000
Accès aux voies navigables	0,021	**	0,009	0,038	***	0,009	0,016	*	0,009	0,035	***	0,009
Spécialisation agricole	-0,220	***	0,006	-0,265	***	0,012	-0,223	***	0,006	-0,261	***	0,012
Spécialisation industrielle	0,118	***	0,006	0,002		0,007	0,111	***	0,006	0,002		0,007
Diversité sectorielle	-0,100	***	0,003	-0,061	***	0,003	-0,094	***	0,003	-0,057	***	0,003
Constante	7,344	***	0,038	7,156	***	0,039	6,954	***	0,040	6,349	***	0,027
Observations	30450			30450			30450			30450		
R^2	**0,2379**			**0,3082**			**0,2421**			**0,3115**		
Effets temporels	Oui			Oui			Oui			Oui		
Effets individuels	Oui			Oui			Oui			Oui		
Effets croisés	Non			Oui			Non			Oui		

Source : Base de données FRANcE

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Effet des temps de parcours moyen entre 1866 et 1911

4. L'instrumentation de l'accès au réseau et des temps de parcours moyens par des variables retardées

La première étape de l'instrumentation consiste à vérifier que la variable instrumentale a un coefficient significatif dans l'estimation de l'accès au réseau. Cette première étape reprend par ailleurs les autres variables de contrôle utilisées jusqu'à maintenant, les effets individuels et temporels rendant compte de l'hétérogénéité inobservée. Le pouvoir explicatif présente une valeur correcte, toujours étant donnée la taille de l'échantillon, et comparée aux autres résultats de la littérature. Par ailleurs, le test F de Kleibergen-Paap confirme sa robustesse[79]. Dès lors, nous pouvons examiner la seconde étape de la spécification, où la variable d'accès au réseau est instrumentée. Avec un pouvoir explicatif de 22 %, et en reprenant toutes les variables de contrôle, nous constatons que l'effet de l'accès au réseau reste significatif et stable sur la période entre 1866 et 1911. Pour autant, notre étude nous pousse à instrumenter les temps de parcours moyen.

Variable dépendante : Niveau de population entre 1866-1911	**(1)** *First-stage*			**(2)** *Second-stage*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accès au réseau	-	-	-	0,174	***	0,032
Corridor des grands plans	0,068	***	0,004			
Observations	30450			30450		
R^2	**0,1415**					
R^2 Second-stage				**0,2226**		
First stage F-Test	**332,07**					
Over-id test p-value						
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		
Effets croisés	Non			Non		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Instrumentation de la variable d'accès au réseau entre 1866 et 1911

Pour prendre en compte les temps de parcours, nous faisons appel aux contraintes topographiques et à une situation retardée. L'objectif est d'utiliser les difficultés à rejoindre la gare la plus proche comme emblématiques des temps de parcours moyens, en supposant ainsi que ces contraintes de cheminement n'ont pas d'incidence sur la croissance démographique locale. La prise en compte de cette difficulté est retardée de deux périodes. Nous croisons alors deux instruments pour de meilleures estimations. Ainsi, nous prenons en compte la distance euclidienne à la gare la plus proche, à laquelle nous ajoutons la difficulté de cheminer[80].

[79] Comme prévu, nous n'arrivons toutefois pas à valider l'exogénéité de l'instrument. Mais à la manière de Atack et al. (2010), nous avons effectué les régressions de cette variable sur la croissance de la population, portant sur les périodes contemporaines à l'adoption des plans : le coefficient d'estimation de la variable « corridor » n'étant pas significatif, nous continuons nos investigations.

[80] Cette variable consiste à rapporter le temps de parcours à pied vers la gare la plus proche selon le graphe P au temps de parcours théorique à partir de la distance euclidienne, à une vitesse de 5 km/h.

La première étape suggère une relation significative où les difficultés topographiques constatées vingt ans auparavant permettent de prédire les temps de parcours moyens, avec un F-Test très élevé *(Colonne 1)*. Alors la *Colonne 2* suggère un effet significatif des temps de parcours moyens sur la croissance des niveaux de population entre 1866 et 1911, quand ils sont instrumentés. Le pouvoir explicatif s'en trouve même renforcé. La robustesse de l'instrument est par ailleurs confirmée par un test de sur-identification des variables instrumentales, dont une *p-value* élevée permet de rejeter l'hypothèse nulle de cette sur-identification et donc de valider les instruments utilisés.

Variable dépendante : Niveau de population entre 1866-1911	**(1)** *First-stage*			**(2)** *Second-stage*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	-	-	-	-0,212	***	0,029
Distance à vol d'oiseau à la gare la plus proche	0,017	***	0,001			
Rugosité vers la gare la plus proche	0,002	**	0,001			
Observations	30450			30450		
R^2	**0,94**					
R^2 *Second-stage*				**0,3117**		
F-Test	**1731,25**					
Over-id test p-value				**0,47**		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		
Effets croisés	Non			Non		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Instrumentation des temps de parcours entre 1866 et 1911

Pour modéliser les effets directs et indirects, la première étape de l'instrumentation est répétée pour chacune des variables endogènes *(Colonnes 1 et 2)*. Il est intéressant de noter la significativité de chacune des variables d'interaction à 5 % d'intervalle de confiance. Ces premières étapes passent l'épreuve de la robustesse des instruments quand le test de sur-identification permet de les valider, à un seuil qui est respecté mais proche du refus. Ainsi, la *Colonne 3* montre qu'avec l'utilisation des variables instrumentales, l'effet des temps de parcours moyens conserve un rôle significatif, alors même que l'effet direct de la gare a un coefficient significatif lui aussi au seuil de 5 %. Notons que les coefficients sont renforcés quand on utilise les variables instrumentales, attestant de la confirmation d'un effet du réseau sur la croissance démographique. La spécification avec des effets croisés pour renforcer la prise en compte d'une éventuelle hétérogénéité à des échelons plus agrégés montre des coefficients tout aussi stables, voire légèrement supérieurs, qui viennent aussi renforcer la robustesse des estimations. La démarche des deux étapes est également répétée pour l'instrumentation prend en compte les réseaux historiques.

Variable dépendante : Niveau de population entre 1866-1911	(1) First-stage			(2) First-stage			(2) Second-stage		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	-	-	-				-0,242	***	0,027
Effet marginal "gare"				-	-	-	0,010	**	0,005
Distance à vol d'oiseau à la gare la plus proche	0,016	***	0,001						
Rugosité vers la gare la plus proche	0,002	**	0,001						
Interaction 1				0,040	***	0,003			
Interaction 2				0,002	**	0,001			
Observations	30450			30450			30450		
R^2	**0,94**			**0,05**					
R^2 Second-stage							**0,3089**		
F-Test	**1365,37**			**58,8**					
Over-id test p-value							**0,17**		
Effets temporels	Oui			Oui			Oui		
Effets individuels	Oui			Oui			Oui		
Effets croisés	Non			Non			Non		

Source : Base de données FRANcE

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Instrumentation des effets directs et indirects du réseau entre 1866 et 1911

5. L'instrumentation de l'accès au réseau et des temps de parcours moyens par les réseaux historiques

Les tests statistiques nous permettent de mieux évaluer la validité des instruments quand ils sont plusieurs. La première étape *(Colonne 1)* montre le caractère significatif des deux variables sur la variable endogène avec un F-Test qui reste largement haut ***(Erreur ! Source du renvoi introuvable.)***. Alors, la seconde étape montre encore une relation significative de l'accès direct au réseau, avec un pouvoir explicatif qui reste correct *(Colonne 2)*. Le test de sur-identification nous permet de valider la qualité des instruments utilisés.

Variable dépendante : Niveau de population entre 1876-1906	**(1)** *First-stage*			**(2)** *Second-stage*		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Accès au réseau	-	-	-	0,268	***	0,063
Corridor des grands plans	0,067	***	0,007			
Temps d'accès aux réseaux historiques	-0,004	***	0,001			
Observations	34624			34624		
R^2	**0,1288**					
R^2 Second-stage				**0,1812**		
First stage F-Test	**49,00**					
Over-id test p-value				**0,5572**		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		
Effets croisés	Non			Non		

Source : Base de données FRANcE © Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Instrumentation de l'accès au réseau par les réseaux historiques entre 1876 et 1906

En revanche, il semble plus aisé d'instrumenter les temps de parcours moyens par les temps de parcours calculés sur les réseaux historiques. Pour renforcer la robustesse des instruments, et dans un raisonnement géohistorique, nous prenons en compte également la connexion aux réseaux historiques. Ainsi, la première étape montre une relation significative avec un pouvoir explicatif très fort. Plus le temps de parcours moyen sur les réseaux historiques est fort, plus le temps de parcours moyen sur le réseau ferroviaire est élevé *(Colonne 1)*. Le F-Test est largement significatif et nous conduit à analyser la seconde étape. Alors l'effet des temps de parcours moyen est significatif et garde un pouvoir explicatif correct étant donné le nombre d'observations *(Colonne 2)*. La validité des instruments est encore une fois ici attestée.

Variable dépendante : Niveau de population entre 1876-1906	(1) First-stage			(2) Second-stage		
	Coef.	*Signif.*	*Ecart-type*	*Coef.*	*Signif.*	*Ecart-type*
Temps de parcours moyen	-	-	-	-0,634	***	0,092
Superposition aux réseaux historiques	-0,117	***	0,002			
Temps de parcours moyen sur les réseaux historiques	0,663	***	0,004			
Observations	34624			34624		
R^2	**0,8388**					
R^2 Second-stage				**0,1972**		
F-Test	**162,498**					
Over-id test p-value				**0,9483**		
Effets temporels	Oui			Oui		
Effets individuels	Oui			Oui		
Effets croisés	Non			Non		

Source : Base de données FRANcE

© Christophe Mimeur, Laboratoire ThéMA, CNRS, 2016

Instrumentation des temps de parcours moyens par les réseaux historiques entre 1876 et 1906

Annexe 2. Code R pour la simulation d'un réseau de transport

1. Chargement des fonctions du modèle

```
### FLOW BETWEENNESS COMPUTATION FUNCTIONS

#################################################################
## Return the amount of flow between nodes u and v
## according to the gravity model with exponent 2
#################################################################
flow_amount<-function(p1,p2,w1,w2){
  euclidian_dist<-as.numeric(dist(rbind(p1,p2)));
  if(euclidian_dist>0){return(w1*w2/(euclidian_dist*euclidian_dist))}
  else{return(0);}
}

#################################################################
## Return the total amount of flow going through each edges
## !!! WARNING : does not take into account multiple paths
## between a node pair !!!
#################################################################


flow_betweeness<-function(g,cost,layout,weight,is_starting_node){
  nodes<-V(g);
  if(!is.null(is_starting_node)){nodes<-V(g)[is_starting_node];}
  res<-numeric(length(E(g)));
  for(i in 1:(length(nodes)-1)){
    u<-nodes[i];
    sp<-get.shortest.paths(g,from=u,to=as.numeric(nodes)[(i+1):length(nodes)],mode="all",weights = cost,output = "epath" );

    for(j in 1:(length(nodes)-i)){
      w<-nodes[i+j];
      path<-sp$epath[[j]];
      res[path]<-res[path]+flow_amount(layout[i,],layout[i+j,],weight[i],weight[i+j]);
      #res[path]<-res[path]+1.
    }

  }
  return(res*cost);
}
```

```
#### ROUTING EXPANSION FUNCTIONS


#################################################################
## Normalization of vect for the selected candidates
#################################################################
normalize_minmax<-function(vect,candidates){
  vect[candidates]<-(vect[candidates] - min(vect[candidates]))/(max(vect[candidates])- min(vect[candidates]));
  return(vect);
}
#################################################################
## Cost update of candidates edges according to the flow
## for the selected candidates
## alpha : between 0 and 1
#################################################################
length_modification<-function(cost,flow,candidates,alpha){
  flow<-normalize_minmax(flow);
  cost[candidates]<-cost[candidates]*(1.-alpha*flow[candidates]);
  cost<-cost/sum(cost);
  return(cost);
}
```

2. Initialisation du modèle

```
################################
## Chargement du graphe vide ###
################################

GRAPHE <- read.dbf("TRI_PLM_MORVAN.dbf")

g <- graph.data.frame(GRAPHE, directed = FALSE)
summary(g)

####################################
## Attributs à l'initialisation ####
####################################

position = cbind(VERT$coord_x, VERT$coord_y)
cost = GRAPHE[,3]
weight = V(g)$pop

######################################
#### Affichage du graphe initial ####
######################################

plot(g,layout = position,edge.width=.1, vertex.size = weight, vertex.label = NA)
```

3. Exécution du modèle pour α = 0,1 entre 2 et 20 itérations

```
################################
####      Calcul ROUTING      ####
################################

flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,2)
GRAPHE$A1IT2<- flow
flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,4)
GRAPHE$A1IT4<- flow
flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,6)
GRAPHE$A1IT6<- flow
flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,8)
GRAPHE$A1IT8<- flow
flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,10)
GRAPHE$A1IT10<- flow
flow<-iteration_flow_routing(g,cost,weight,position,!logical(ecount(g)),0.1,20)
GRAPHE$A1IT20<- flow
```

RESUME

Les interactions entre transport et territoire sont l'objet d'une littérature scientifique permanente, questionnant les impacts économiques et démographiques d'une nouvelle infrastructure, souvent évoqués à l'échelle d'un projet. L'objectif de la thèse est de réinvestir les composantes de l'interaction par les larges échelles spatiales et temporelles, en posant l'hypothèse que la profondeur temporelle et l'échelle du territoire national sont porteuses de nouvelles explications. Ce travail s'appuie sur la collecte, l'exploitation et l'analyse de la large base de données FRANcE (French RAilway NEtwork), qui recense chaque section du réseau ferroviaire français depuis le début du XIXème siècle et les recensements démographiques. Cette base renferme également les *traces de la vitesse*, qui constituent une information inédite sur l'ensemble du réseau et qui permet de faire de l'accessibilité une variable décisive dans les explications.

Plutôt que de se concentrer sur l'acquisition de nouvelles données au prix d'une lourde collecte, nous misons sur la construction d'un appareil méthodologique pour étudier les deux sens de l'interaction entre réseau et territoire, qui requiert toutefois une adaptation des dispositifs de structuration des données et d'analyse. La démarche de la thèse consiste en une modélisation croissante du phénomène, de la compréhension et la formalisation des objets jusqu'à la formalisation des données et des analyses, ce qui nécessite le recours à d'autres disciplines. Ce travail utilise le formalisme des graphes pour investiguer les deux sens de la relation. Il aide à étudier l'effet du réseau à partir d'une diversification de la donnée et de sa modélisation pour rendre compte de portées spatiales et temporelles. Il aide à étudier l'impact d'une structure préexistante dans la morphogénèse du réseau ferroviaire français à partir d'un modèle d'évolution endogène, entre diffusion du rail et hiérarchisation des infrastructures. Ce travail vise à mieux comprendre les liens qui unissent réseau et territoire, dont les outils méthodologiques peuvent être appliquées à d'autres réseaux, d'autres temporalités, jusqu'à des problématiques actuelles.

Mots-clés : Réseau ferroviaire ; Territoire ; Graphe ; Modélisation ; Géohistoire

SUMMARY

The interaction between space and network are frequently questioned in the academic literature, by asking the economical and demographical impacts of a new infrastructure, often studied at the scale of a project. This work aims to investigate the components of the interaction in both large spatial and temporal scales. The hypothesis is that the temporal depth and the national scale could bring new explanations. This work is based on the collect, the exploitation and the analysis of the large spatio-temporal database FRANcE (French Railway Network). It identifies all sections of the network since the 19th century and the population census. This database also contains the traces of the speed, which are novel information for network, and allows the accessibility to become a decisive variable in the explanations.

Rather than acquisition new data with an intensive phase of collect, we aim to build a methodological chain to study the two senses of interaction between space and network. It requires the adaptation of data structuration and analysis. The approach of this thesis consists on the growing modelling of the phenomenon, from the comprehension to formalization of data to the analysis, which requires the use of other disciplines. This work uses the graph theory to investigate the two senses of the relationship. It permits to study the network effect in the long run by diversifying the data to identify spatial and temporal ranges. It permits to study the impact of a pre-existing structure in the morphogenesis of the network, by using a dynamic model of network evolution, between diffusion and hierarchical organization. This work aims to understand the link between space and network, where the methodological tools can be adapted to other networks, other times and actual questioning.

Keywords: Railway network; Space; Graph; Modelling; *Spatial Humanities*