THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Infectiologie, Physipoathologie, Toxicologie, Génétique et Nutrition

---

Présentée et soutenue par :

M. GIANG-NAM LE

le mercredi 2 novembre 2022

Titre :

Nouvelle mesure de la robustesse des animaux d'élevage par utilisation des données de phénotypage haut-débit.

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Génétique, Physiologie et Systèmes d'Elevage (GENPHYSE)

Directeur(s) de Thèse :

MME INGRID DAVID

M. TOM ROHMER

Rapporteurs :

MME SANDRINE GRASTEAU, INRA TOURS

M. RAFAEL MUNOZ-TAMAYO, INRA PARIS

Membre(s) du jury :

MME CATHERINE LARZUL, INRA TOULOUSE, Président

MME FLORENCE YTOURNEL, , Membre

MME INGRID DAVID, INRA TOULOUSE, Membre

M. TRISTAN MARY-HUARD, AGROPARISTECH, Membre

Remerciements

Cette thèse a été financée par l'Institut National de la Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE) et l'Alliance R&D (association regroupant les organismes de sélection porcine (Axiom, Choice Genetics et Nucleus) et l'Institut Technique du Porc - IFIP). Cette thèse a aussi été labelisée par l'Institut Convergences Agriculture Numérique (DigitAg).

Tout d'abord, je tiens à remercier mes directeurs de thèse : Ingrid David et Tom Rohmer. Merci beaucoup pour votre encadrement extraordinaire, pour vos précieux conseils et pour votre disponibilité. Votre suivi régulier tout au long de cette thèse m'a permis de finaliser cette étude de thèse, sans vous je n'aurais jamais pu atteindre cet objectif.

J'aimerais ensuite remercier tous les membres de mon jury de thèse : Rafael Munoz-Tamayo, Sandrine Mignon-Grasteau, Catherine Larzul, Tristan Mary-Huard et Florence Ytournel pour avoir accepté d'évaluer mes travaux de thèse.

Je voudrais aussi remercier les membres de mon comité de thèse : Nicolas Friggens, Alban Bouquet, Loïc Flatres-Grall, Bruno Ligonesche, Boris Hejblum et David Picard pour votre contribution au cours de cette étude. Merci pour tous les échanges, les conseils et les commentaires visant à améliorer la qualité de cette thèse.

Je remercie également toutes les personnes d'INRAE qui m'ont aidé avant et pendant la préparation de ma soutenance de thèse : Anne, Christel, Dominique, Charlotte, Céline, Davi, Jerome, Hervé, Pierre, Mathieu, Guillermo, Andres, Sylvia et Carolina. Merci pour votre temps et pour votre aide précieuse.

J'aimerais enfin remercier toute la bande des non-permanents que j'ai rencontré grâce à cette thèse : Guillermo, Simona, Marine, Riccardo, Vincent, Coralie, Marie, Océane, Quentin, Timothé, Martin, David, Fernando, Vanille, Audrey. Merci pour votre gentillesse, pour votre bonne humeur et votre soutien durant ces trois années de thèse. Vous avez rendu mon séjour à Toulouse inoubliable.

Table des matières

Introduction 1

Chapitre 1 - État de l'art 4

1.1 Élevage porcin 4

1.1.1 Présentation générale 4

1.1.2 Sélection génétique 8

1.2 Sélection pour la robustesse 10

1.2.1 Contexte 10

1.2.2 Définitions de la robustesse 13

1.2.3 Quantification de la robustesse 18

1.2.4 Organisation de l'étude de thèse 30

Chapitre 2 - Données d'étude 33

2.1 Origine 33

2.2 Données individuelles 35

2.2.1 Données issues des distributeurs automatiques de concentrés 35

2.2.2 Informations complémentaires sur les animaux 37

2.3 Données d'environnement 38

2.4 Mise en forme et nettoyage des données 38

2.4.1 Données de santé et météorologiques 38

2.4.2 Données issues des DAC 40

2.4.3 Statistiques descriptives des données longitudinales 45

Chapitre 3 - Impact des perturbations environnementales sur l'estimation des paramètres et valeurs génétiques 50

3.1 Introduction 50

3.2 Article 51

3.3 Discussion complémentaire de l'article 'Impact des perturbations environnementales sur l'estimation des paramètres et valeurs génétiques' 78
Chapitre 4 - Détection des perturbations et quantification de la robustesse.....82
4.1 Introduction 82
4.2 Matériels et méthodes 82
4.2.1 Partie Up 83
4.2.2 Partie Down 86
4.2.3 Caractérisation des perturbations 89
4.3 Application de la méthode Up & Down 91
4.3.1 Données simulées 91
4.3.2 Données réelles.....94
4.4 Résultats 94
4.4.1 Données simulées 94
4.4.2 Données réelles.....101
4.5 Discussion 102
4.6 Conclusion.....112
4.7 Analyses et résultats complémentaires 113
4.7.1 Identification du début des perturbations 113
4.7.2 Méthodes de lissages 114
4.7.3 Statistiques pour la partie Up 115
Conclusion générale et perspectives.....117
Références 121
Formations 132
Valorisations.....134
Publications scientifiques.....134
Article en cours d'écriture 134

Communications à des congrès.....134

Résumé 2

Abstract.....4

Listes des tables

Table 1 - Différentes définitions de la robustesse globale, robustesse spécifique, résilience et résistance. 15
Table 2 - Répartition des porcs de la station de phénotypage au Rheu par race de 2017 à 2019 34
Table 3 – Pourcentages dans la base de données des différentes combinaisons des caractères de comportements alimentaires selon leur validité 42
Table 4 - Statistiques descriptives de la consommation alimentaire par race de porcs.....43
Table 5 - Résultats sur l'évolution des corrélations partielles entre la VG prédite de la production et la VG simulée de la robustesse (résistance et résilience) selon différentes situations.....80
Table 6 - Détail des classes d'intensité, durée et moment du début de perturbation. 92
Table 7 - La sensibilité (Se), spécificité (Sp), valeur prédictive positive (VPP) et valeur prédictive négative (VPN) à l'échelle globale, bande, case et individu pour la méthode Up & Down utilisant la fonction de lissage Nadaraya-Watson (noyau gaussien et taille de fenêtre égale à 10) sur 1000 simulations.....95
Table 8 - Statistiques pour l'évaluation des caractéristiques des perturbations à l'échelle globale, bande, case et individu sur 1000 simulations.....95
Table 9 - Héritabilité estimée et précision de la VG prédite pour la résistance et résilience en fonction du pourcentage de perturbations (résultats sur 1000 simulations) 100
Table 10 – Nombre d'éléments détectés selon les phénotypes utilisés (poids vif ou CFI) pour la méthode Up & Down. 101
Table 11 – La médiane des écarts en valeur absolue entre les débuts (edp) (ou fins (efp)) des perturbations détectées avec le poids et CFI pour toutes les échelles (bande, case, individu). 102
Table 12 – La médiane des écarts entre le début simulé et le moment du maximum local du phénotype corrigé et lissé ($edp2$) et le moment du taux de décroissance maximale du phénotype corrigé et lissé ($edp1$) à l'échelle de la bande, case et individu. 114
Table 13 - La sensibilité et spécificité en pourcentage pour toutes les méthodes de lissage testées à l'échelle globale, bande, case et individu. 114

Table 14 – La sensibilité et spécificité en pourcentage à l'échelle de la bande, de la case et de l'individu pour les statistiques complémentaires concernant les résidus et le coefficient de variation. 115

Listes des figures

Figure 1 - Chaine de valeur du porc 5
Figure 2 - Pyramide d'élevage du porc 6
Figure 3 - Evolution historique des objectifs de sélection animale pour la production de viande (Phocas et al., 2016). 12
Figure 4 - Réponse dynamique d'un système à une perturbation.....16
Figure 5 - Modalités des réponses adaptatives aux perturbations 17
Figure 6 - Une représentation schématique des fonctions biologiques opérant aux niveaux sous-jacents (n-1) qui se combinent pour construire la robustesse au niveau (n) de l'animal (Friggens et al., 2017). 19
Figure 7 - Schéma d'organisation de l'étude de thèse. 31
Figure 8 - Représentation de la station de phénotypage de l'Alliance R&D et INRAE au Rheu 34
Figure 9 - Répartition des porcs dans des bandes et cases-intra-bande équipées d'un distributeur automatique de concentrés. 36
Figure 10 - Histogramme de la consommation alimentaire (g), du logarithme du temps passé au DAC et du logarithme de la vitesse d'ingestion des porcs étudiés à l'échelle de la visite. Les lignes rouges correspondent aux seuils pour déterminer les observations aberrantes.....41
Figure 11 - Illustration de la correction du poids pour un porc donné 44
Figure 12 - Exemple de poids mal corrigé 45
Figure 13 - Evolution moyenne du poids (à gauche) et de son coefficient de variation (à droite) par jour et par race.....46
Figure 14 - Evolution moyenne de la CFI (à gauche) et de son coefficient de variation (à droite) par jour et par race.....47
Figure 15 - Evolution moyenne de la VI (à gauche) et de son coefficient de variation (à droite) par jour et par race.....48
Figure 16 - Evolution du poids vif brut (graphique gauche), corrigé (par les effets fixes et la médiane de la population par âge) et lissé avec la fonction de lissage non-paramétrique de Nadaraya-Watson (graphique de droite). 84

Figure 17 - Modèle de mélange appliqué sur les minimums de pente pour classer la population d'animaux en deux groupes : perturbé (sous la courbe verte) et non-perturbé (sous la courbe rouge) ; la ligne bleue pointillée correspond au seuil τ qui sépare les deux populations. 85
Figure 18 - Identification du (des) moment(s) de référence (MR) sur la dérivée première du phénotype corrigé et lissé 87
Figure 19 - Identification du (des) moment(s) de référence (MR) sur l'évolution du phénotype corrigé et lissé 87
Figure 20 - Arbre de prise de décision dichotomique pour la méthode Up & Down 89
Figure 21 - Identification de la fin de la perturbation sur le phénotype corrigé et lissé. 90
Figure 22 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de bande (résultats sur 1000 simulations). 96
Figure 23 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de case (résultats sur 1000 simulations) 97
Figure 24 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de l'individu (résultats sur 1000 simulations) 97
Figure 25 - L'effet de la durée et du moment du début de la perturbation sur la corrélation entre l'intensité estimée et théorique par niveau de perturbation (bande, case, individu). ..98
Figure 26 - L'effet de l'intensité de perturbation sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations) 99
Figure 27 - L'effet de la durée de perturbation sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations) 99
Figure 28 - L'effet du début perturbé sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations).....100
Figure 29 - Effet de la durée de la perturbation sur la caractérisation de la perturbation sur le poids vif corrigé et lissé. 106
Figure 30 - Effet du moment perturbé sur la caractérisation de la perturbation sur le poids vif corrigé et lissé. 107
Figure 31 - Exemple d'un animal détecté comme perturbé avec son poids (en haut) vers mi-aout mais pas avec sa CFI (en bas). 109

Figure 32 - Exemple d'un animal détecté comme perturbé avec sa CFI (en bas) vers début Mars mais pas avec son poids (en haut). 110
Figure 33 - Identification du moment des maximums locaux du phénotype corrigé et lissé. 113

Listes des abréviations

ABC : L'indice de l'aire entre les courbes

BDD : Base de données

CFI : Consommation alimentaire cumulée

CV : Coefficient de variation

DAC : Distributeurs automatiques de concentrés

DS : Données de santé

edp : Ecart entre le vrai début de perturbation et le début estimé

$edur$: Ecart entre la vraie durée de perturbation et la durée de perturbation estimée

efp : Ecart entre la vraie fin de perturbation et la fin de perturbation estimée

FI : Consommation alimentaire

GxE : Interaction Génotype x Environnement

GMQ : Gain moyen quotidien

IFIP : Institut Technique du Porc

INRAE : Institut National de la Recherche pour l'agriculture, l'alimentation et l'environnement

MR : Moments de référence

OSP : Organismes de Sélection Porcine

PAS : Populations animales sélectionnées

PR_j : Poids de référence du jour *j*

RFID : Puce radio-identification

Se : Sensibilité

Sp : Spécificité

THI : Indice de température et d'humidité

TI : Temps passé au DAC

VI : Vitesse d'ingestion

VG : Valeur génétique

VPN : Valeur prédictive négative

VPP : Valeur prédictive positive

$W0$: Poids d'entrée en salle d'engraissement

Introduction

Dans le futur, les animaux vont être élevés dans des conditions plus diversifiées et perturbées qu'à l'heure actuelle en lien avec les demandes sociétales pour une amélioration du bien-être des animaux (abolition des cages individuelles, élevage en plein air, …) et les demandes des éleveurs pour une réduction des coûts alimentaires et énergétiques. Les changements environnementaux sont également le résultat d'une transition agroécologique des systèmes d'élevage visant à développer une agriculture durable et écologique. De plus, les conditions environnementales sont aussi sujettes à une plus grande variabilité du fait du changement climatique entrainant l'augmentation des phénomènes météorologiques extrêmes (canicule).

Un animal robuste sera peu sensible aux variations de milieu et aura la capacité de maintenir sa production quel que soit l'environnement dans lequel il est élevé. Cela assurera le maintien de la profitabilité de l'élevage, la réduction des problèmes sanitaires, l'amélioration du bien-être des animaux et la réduction de l'utilisation d'intrants (en particulier médicamenteux). Par conséquent, l'amélioration de la robustesse des animaux par la sélection génétique est donc devenue un levier majeur du renforcement de la durabilité des élevages répondant à toutes ces attentes importantes. Comme tous les caractères d'intérêt, mesurer la robustesse des animaux est un prérequis indispensable à sa sélection. Cependant, la robustesse est un caractère complexe et difficile à phénotyper. Actuellement, les rares stratégies de sélection pour une meilleure robustesse passent par la sélection de caractères fonctionnels (par exemple : les boiteries, la longévité, la résistance aux maladies, etc…). Il ne s'agit là que des mesures indirectes pour le maintien d'une robustesse globale et la réponse adaptative au cours du temps des animaux (robustesse spécifique) n'est pas prise en compte. Avec le développement d'outils d'élevages de précision, il est maintenant possible de mesurer en continu plusieurs phénotypes individuels. Cela offre de nouvelles possibilités de mesure de la robustesse des animaux en étudiant les dynamiques de réponse des phénotypes longitudinaux face aux perturbations environnementales.

L'objectif principal de cette thèse est d'extraire de l'analyse des données longitudinales, un nouveau critère de robustesse des animaux d'élevage. L'idée est de développer une méthode qui consiste, en étudiant la dynamique d'évolution d'un phénotype

au cours du temps à des échelles différentes : individuelle ou groupe d'animaux (case et bande), à identifier les éléments (animal, case, bande) soumis à des perturbations, caractériser les perturbations identifiées puis quantifier la robustesse des animaux en mesurant les réponses individuelles face à ces perturbations. Le travail de thèse repose sur l'analyse de données simulées, puis sur les données réelles fournies par l'Alliance R&D (association regroupant les organismes de sélection porcine (Axiom, Choice Genetics et Nucleus) et l'Institut Technique du Porc). Dans le premier chapitre de ce manuscrit, une analyse bibliographique a été menée sur la sélection porcine actuelle et les différents concepts et méthodes de quantification de la robustesse. Puis, les données d'étude fournies par l'Alliance R&D sont présentées dans le deuxième chapitre. Dans le chapitre 3, l'objectif était de répondre à la question scientifique suivante : est-ce que la sélection à partir de phénotype de production enregistrés sur des animaux élevés dans un environnement fortement contrôlé permet une sélection conjointe sur la production et la robustesse ? L'utilisation de données simulées permettait de comprendre l'impact des perturbations apparaissant dans un environnement bien contrôlé, tel que rencontré en élevage de sélection porcine, sur l'estimation des paramètres et valeurs génétiques de la production. Ce chapitre intègre l'article publié dans le Journal Animal « Impact of environmental disturbances on estimated genetic parameters and breeding values for growth traits in pigs ». Enfin, le chapitre 4 présente une méthode à deux étapes (Up & Down) que nous avons développée pour détecter et caractériser des perturbations environnementales inconnues affectant les animaux d'élevage, et quantifier la robustesse des animaux perturbés. Son application sur des données simulées et réelles est mise en avant au sein de ce chapitre.

Chapitre 1 - État de l'art

1.1 Élevage porcin

1.1.1 Présentation générale

Le porc est essentiellement élevé pour sa viande et plus anecdotiquement comme animal de compagnie ou d'expérimentation (Köhn et al., 2007). La production mondiale de viande porcine en 2020 est estimée à 109,2 millions de tonnes. La viande de porc est la première viande consommée en France, en Europe et dans le monde. En France, la production de porcs se concentre dans l'Ouest de la France (notamment en Bretagne qui compte 58% des élevages porcins de la métropole en 2020). La France est le troisième producteur de porc en Union Européenne, après l'Allemagne et l'Espagne, et devant le Danemark, les Pays-Bas et la Pologne (IFIP, 2013).

1.1.1.1 Chaîne de valeur du porc

La production de viande de porc s'effectue par une chaîne d'événements allant de la production de reproducteurs à la production de produits de porc vendus au détail. Cette chaine est appelée chaîne de valeur du porc (Dekkers et al., 2011). Elle comprend tous les partenaires impliqués dans la production de porc : depuis les fournisseurs de génétique et les producteurs de porcs jusqu'aux abattoirs et les magasins de vente au détail qui apportent les produits finis aux consommateurs (Figure 1). Ces partenaires sont très dépendants les uns des autres. La qualité de la viande vendue aux consommateurs dépend fortement de la qualité des porcs sélectionnés. S'il y a un souci concernant la qualité du produit final vendu aux consommateurs, il est nécessaire de remonter dans la chaine jusqu'à ce que la source du problème soit identifiée. Dans de nombreux pays, les partenaires de la chaîne de valeur du porc travaillent ensemble pour assurer la compétitivité de leur industrie porcine sur le marché mondial (un forum d'interaction entre les partenaires peut être créé), tandis que dans d'autres pays comme la France, ce sont des entreprises individuelles qui assument ce rôle. Enfin, en plus de la qualité de viande, il est important pour la chaine de valeur de développer leurs produits selon les demandes sociétales. En effet, il y a une augmentation de pression sociétale

et de désir de plusieurs producteurs de porcs d'inclure dans les objectifs de sélection des caractères tels que la santé et le bien-être des porcs, et l'impact environnemental de la production porcine (Verbeke et Viaene, 2000; Kanis et al., 2005) (plus de détail dans la suite du document).

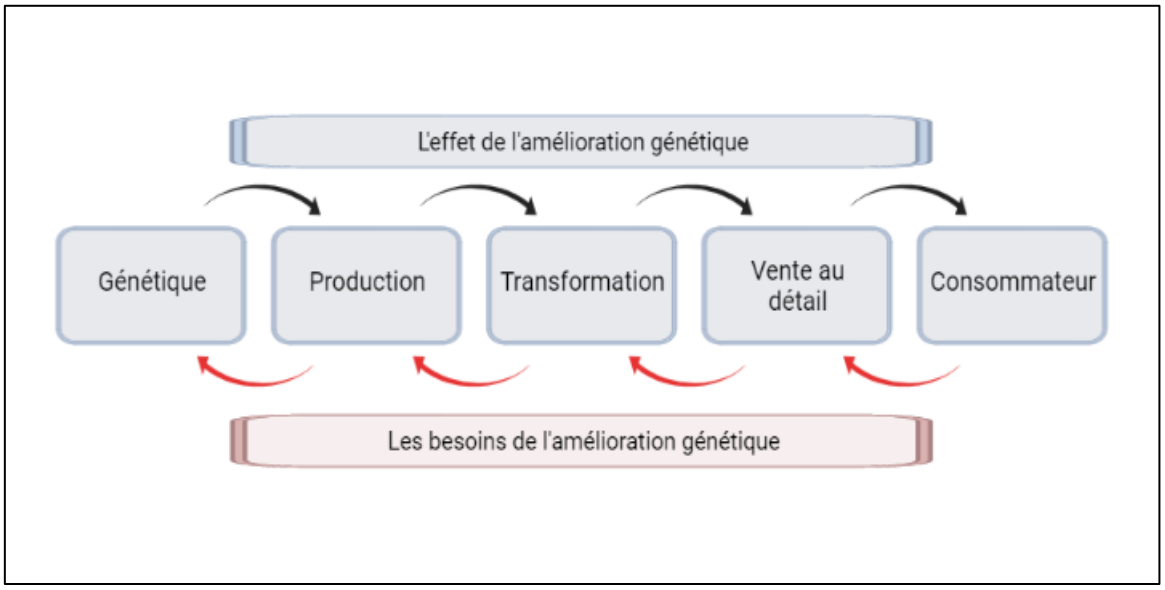


Figure 1 - Chaine de valeur du porc

1.1.1.2 Organisation pyramidale de l'élevage porcin

Dans la plupart des pays du monde, dont la France, l'organisation de la filière porcine est pyramidale à trois niveaux (Figure 2). Au sommet de la pyramide se trouvent les élevages de sélection. Cet étage comprend peu d'animaux et il s'agit ici de populations de races pures. A ce niveau, l'environnement dans lequel sont élevés les porcs est extrêmement contrôlé. Ces élevages sont « fermés » et un strict protocole d'entrée des personnes et des animaux y est appliqué. Le deuxième étage de la pyramide comprend les élevages de multiplication qui effectuent des accouplements ou des croisements spécifiques pour la production d'un grand nombre de femelles. Ces truies hybrides sont vendues à des producteurs commerciaux (troisième et dernier étage de la pyramide) où elles seront inséminées par des verrats de l'étage de sélection pour la production de porcelets qui seront envoyés aux abattoirs pour produire la viande de porc. Cet étage de production comprend beaucoup plus d'élevages par rapport aux deux autres étages (Dekkers et al., 2011).

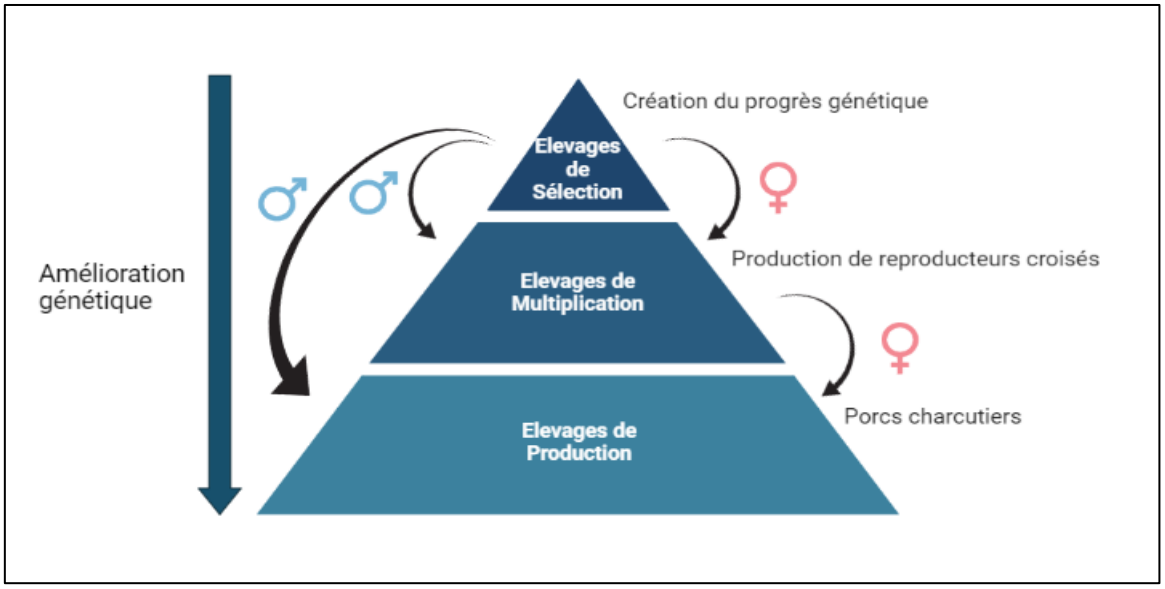


Figure 2 - Pyramide d'élevage du porc

Il existe une grande différence de 'dureté' de l'environnement entre les élevages de sélection et les élevages de production. En effet, comme indiqué précédemment, l'environnement des élevages de sélection est très contrôlé et standardisé. Les animaux élevés à cet étage subissent donc peu de perturbations environnementales de forte intensité et de longue durée. Ces animaux sont également suivis quotidiennement par les animaliers pour s'assurer de leur bien-être et de leur santé. A l'étage de production, les systèmes d'élevage sont plus variés (il existe par exemple des élevages en bâtiment sur caillebotis, des élevages en bâtiment sur paille ou des élevages en plein air) et l'environnement y est moins contrôlé (moins de règles sanitaires). Les animaux élevés dans ces élevages doivent faire face à plus de perturbations et doivent être capable de maintenir une production élevée dans des conditions environnementales différentes et plus perturbées que celle des élevages de sélection.

1.1.1.3 Organisation structurelle

Les **Organismes de Sélection Porcine (OSP)** ont un rôle important dans le maintien des livres généalogiques et des caractéristiques des races. Ces sociétés sont agréées par le Ministère de l'Agriculture pour la tenue des informations des races qu'elles sélectionnent,

aussi appelée **populations animales sélectionnées (PAS)**. Les OSP définissent et mettent en œuvre les programmes de sélection pour leurs PAS (soit populations collectives codétenues par plusieurs OSP soit populations autonomes détenues par une seule OSP) (IFIP, 2013). Elles fournissent aussi des services d'enregistrement des performances et d'évaluation génétique. Les OSP existent dans de nombreux pays et jouent un rôle important dans l'amélioration génétique par le biais de programmes nationaux d'amélioration dans leurs pays respectifs (Dekkers et al., 2011). En France, **l'Institut Technique du Porc (IFIP)** accompagne le développement, la compétitivité et la pérennité de la filière porcine et a un rôle important dans la conduite des programmes des OSP. Cet institut de Recherche-Innovation collabore étroitement avec un ensemble de partenaires, dont **l'Institut National de la Recherche pour l'agriculture, l'alimentation et l'environnement (INRAE).**

1.1.1.4 Races des porcs sélectionnées en France

Depuis 1950, peu de races porcines ont été sélectionnées dans le monde parmi un grand nombre de races répertoriées (environ 350 races) (IFIP, 2013). En France, les 4 principales races sélectionnées sont les Large White, Landrace, Duroc et Piétrain. Ces populations sont scindées en 2 lignées : une lignée femelle dont les objectifs de sélection sont partagés entre les caractères de reproduction et de production, et une lignée mâle uniquement sélectionnée sur les caractères de production. Plus précisément :

- La race Large White (aussi appelée Yorkshire) est sélectionnée pour ses performances de reproduction et de production. Elle est largement utilisée pour la lignée femelle dans les programmes de croisement (l'utilisation de la lignée mâle est plus réduite). Elle est présente dans tous les OSP des pays européens.
- La race Landrace est réputée pour ses qualités maternelles. Ses performances de production sont aussi très bonnes (mais moins bonnes que celles des Large White). Cette race est principalement utilisée pour la lignée femelle.
- La race Piétrain est utilisée comme verrat terminal en race pure pour produire des porcelets dans l'étage de production (Figure 2). Le Piétrain est connu pour ses performances en production (excellent rendement de carcasse et teneur en muscle).

- La race Duroc est originaire des Etats-Unis et est connue pour ses bonnes performances en production. Selon les pays ou les OSP, elle est utilisée dans les plans de croisements, soit en lignée mâle, soit en lignée femelle.

1.1.2 Sélection génétique

L'amélioration génétique des porcs en croissance est un enjeu important sur le plan économique, environnemental et sociétal. Les programmes de sélection animale ont pour but d'améliorer les caractéristiques génétiques des populations sur un ensemble de caractères héritables et d'intérêt pour les générations à venir (Phocas et al., 2013). En sélectionnant les animaux les plus performants comme reproducteurs, cette amélioration génétique peut être atteinte. Les reproducteurs sélectionnés produiront de meilleurs animaux que ceux de la génération précédente et au cours des générations, une amélioration des caractères sous sélection sera observée. Généralement, les principaux critères de sélection pour la reproduction sont la fertilité, le nombre de porcelets vivants par portée, le nombre de tétines fonctionnelles ; et pour les caractères de production ce sont la vitesse de croissance, l'efficacité alimentaire, la composition corporelle, la qualité technologique de la viande.

Le cœur de la sélection animale est la génétique quantitative qui consiste à utiliser et développer des modèles statistiques afin d'estimer les paramètres génétiques et prédire les valeurs génétiques des animaux d'élevage. Les modèles mixtes utilisés à cette fin diffèrent en fonction des caractéristiques des phénotypes analysés (variable continue ou non, répétée ou non…). La valeur phénotypique (P) d'un animal peut être décrite par la somme des effets de son génotype (G) et de son environnement (E) : P = G + E. La composante (G) peut être ensuite décomposée en composante génétique additive (A), de dominance (D) et une composante d'épistasie (I). La sélection génétique s'appuie sur la valeur génétique additive (A). Les modèles statistiques de génétique sont des modèles mixtes plus ou moins complexes en fonction du phénotype analysé. Lorsqu'une mesure par animal et effectuée, le modèle linéaire mixte d'étude de ce type de caractère peut être écrit par :

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{a} + \boldsymbol{\varepsilon},$$

où $\boldsymbol{y}$ représente le vecteur des phénotypes, $\boldsymbol{\beta}$ est le vecteur des effets fixes et $\mathbf{X}$ est la matrice d'incidence correspondante, $\boldsymbol{a}$ est le vecteur des effets génétiques additifs (aléatoire) et $\mathbf{Z}$ est la matrice d'incidence correspondante, et $\boldsymbol{\varepsilon}$ est le vecteur des résidus.

On suppose que les résidus sont indépendants des effets génétiques et qu'ils sont distribués selon une loi normale $N(0, \mathbf{R})$ où $\mathbf{R} = \mathbf{I}\sigma_e^2$ avec σ_e^2 la variance des résidus et $\mathbf{I}$ la matrice identité ; $\boldsymbol{a}$ suit une loi normale $N(0, \mathbf{G})$ où $\mathbf{G} = \mathbf{A}\sigma_a^2$ avec σ_a^2 la variance des effets génétiques et $\mathbf{A}$ la matrice de parenté de animaux du pédigrée.

L'objectif est d'estimer les composantes de variance de ce modèle, les effets fixes ($\widehat{\boldsymbol{\beta}}$) et prédire les effets aléatoires ($\widehat{\boldsymbol{a}}$). Les composantes de la variance sont généralement estimées par la méthode du maximum de vraisemblance restreint (REML), les prédictions des effets génétiques (BLUP) sont obtenues par le système d'équations d'Henderson (1975) :

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\boldsymbol{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\boldsymbol{y} \end{bmatrix}.$$

L'héritabilité du caractère, mesurant la part de variabilité du phénotype due aux effets génétiques, est alors calculée par : $h^2 = \frac{\sigma_a^2}{\sigma_p^2}$ avec $\sigma_p^2 = \sigma_a^2 + \sigma_e^2$. L'héritabilité décrit la transmission héréditaire du caractère et est comprise entre 0 et 1. Plus celle-ci est élevée, plus la variabilité au niveau du phénotype est due à une différence de composition génétique et donc plus grande sera la réponse à la sélection.

Pour l'analyse des données longitudinales, le modèle le plus couramment utilisé en génétique est le modèle de régression aléatoire (Schaeffer, 2004) qui est une extension du modèle précédent. Il a été introduit pour la première fois en 1982 par Henderson (1982) et peut être écrit par l'équation suivante :

$$\boldsymbol{y_t} = \mathbf{X}\boldsymbol{\beta} + \sum_{k=0}^{m_1} \boldsymbol{\alpha_k} f_k(t) + \sum_{k=0}^{m_2} \boldsymbol{p_k} f_k(t) + \boldsymbol{\varepsilon_t},$$

où $\boldsymbol{y_t}$ est le vecteur du phénotype mesuré au temps t, les $\boldsymbol{\alpha_k}$ et $\boldsymbol{p_k}$ représentent respectivement les vecteurs des coefficients de régression pour les effets génétiques et l'environnement permanent, f_k représente un polynôme d'ordre k (polynôme de Legendre par exemple), $\boldsymbol{\varepsilon_t}$ est le vecteur des résidus supposés indépendants des effets génétiques et des effets d'environnement permanent.

Le vecteur des coefficients de régression génétiques $\boldsymbol{\alpha} = (\boldsymbol{\alpha_k})_{k=0,\dots m_1}$ suit une loi normale $N(0, \mathbf{G} \otimes \mathbf{A})$ où $\otimes$ est le produit Kronecker et $\mathbf{G}$ est la matrice de variance-covariance des coefficients de régression aléatoire génétique additive d'ordre $m_1 + 1$, le vecteur des coefficients de régression de l'environnement permanent $\boldsymbol{p} = (\boldsymbol{p_k})_{k=0,\dots m_2}$ suit une loi normale $N(0, \mathbf{P} \otimes \mathbf{I})$ avec $\mathbf{P}$ est la matrice de variance-covariance des coefficients de régression d'ordre $m_2 + 1$.

1.2 Sélection pour la robustesse

1.2.1 Contexte

L'objectif de sélection correspond à un ensemble de caractères, avec leurs pondérations relatives, que l'on souhaite améliorer dans une population afin d'améliorer le bien-être et la santé des animaux et d'accroitre le profit de l'éleveur, de la filière ou de tout autre niveau d'organisation (Phocas et al., 2013). Les objectifs de sélection ont beaucoup évolué au cours du temps en fonction des attentes des éleveurs et de la société. L'évolution simplifiée des objectifs de sélection pour la production de viande en France est présentée dans la Figure 3 (Phocas et al., 2016). La prise en compte de nouveaux caractères se fait de manière cumulative : les nouveaux caractères viennent en complément des caractères déjà sélectionnés, en faisant évoluer le poids accordé à chacun d'eux dans l'objectif de sélection global (Roy et al., 2019).

À l'origine, dans les années 50, les programmes de sélection dans le secteur de la production de viande étaient principalement axés sur la production et la locomotion des animaux. L'objectif principal était d'atteindre une productivité maximale. Par conséquent, seulement un faible nombre de caractères étaient choisis : croissance, composition de carcasse (Bidanel et al., 2020).

Ensuite, dans les années 70, face à l'accroissement de la population mondiale, les objectifs de sélection devaient répondre aux nouvelles demandes des éleveurs : réduction de la quantité alimentaire pour les animaux d'élevage et changements de régime vers une nourriture moins chère. Une des solutions a été de sélectionner les animaux sur leur efficacité alimentaire (les caractères d'intérêt étant l'indice de consommation, la consommation résiduelle…).

Durant les années 90, les programmes d'amélioration génétique ont été progressivement réorientés vers la prise en compte des caractères concernant la qualité des produits et la reproduction. La mise en évidence d'effets défavorables de la sélection sur la qualité de la viande a conduit à inclure un indice de qualité technologique de la viande (combinaison d'une mesure de pH ultime et de critères d'évaluation de la couleur et du pouvoir de rétention d'eau de la viande) dans les objectifs de sélection. En parallèle, la sélection sur la reproduction est consécutive à une forte diminution de la fertilité suite à la sélection pour la production (Lawrence et Wall, 2014). Par conséquent, la prolificité (nombre de porcelets nés au totaux par portées) a été formellement intégrée dans les objectifs de sélection (Bidanel et al., 2020).

A partir des années 2000, la sélection d'individus robustes (les différentes voies de sélection sont développées dans la section suivante) est devenue un enjeu majeur pour différentes raisons : i) le changement climatique entraine de plus en plus d'évènements météorologiques extrêmes. En effet, la température moyenne globale à la surface de la Terre a augmenté de 1,3°C depuis le début de l'industrialisation. Par conséquent, cela entraine une augmentation de la durée, fréquence et intensité des vagues de chaleurs partout dans le monde. Depuis le milieu des années 2000, des canicules extrêmes (du type de celles qui n'affectaient qu'environ 1% de la surface terrestre dans les années 1900) ont touché près de 10% de la surface du globe. De plus, d'après différents scénarios concernant la projection au cours du XXI^{e} siècle, les vagues de chaleur inhabituelles (du type de celles qui ne se reproduisaient qu'une fois tous les 740 ans) deviendraient de plus en plus fréquentes et couvriraient jusqu'à 60% des terres émergées du globe à l'horizon 2100 (Van Gameren et al., 2014) ; ii) les systèmes d'élevage varient en lien avec les attentes de la société (abolition des cages individuelles, élevage en plein air, réduction des intrants médicamenteux) et des éleveurs (réduction des couts énergétiques, changement vers une alimentation moins chère). Les futurs animaux sélectionnés doivent donc être robustes et capables de s'adapter à des conditions environnementales variées. Pour répondre à ce besoin, il est essentiel d'enregistrer un phénotype sur la robustesse (ou autres caractères associés comme la résilience ou résistance) des animaux d'élevage.

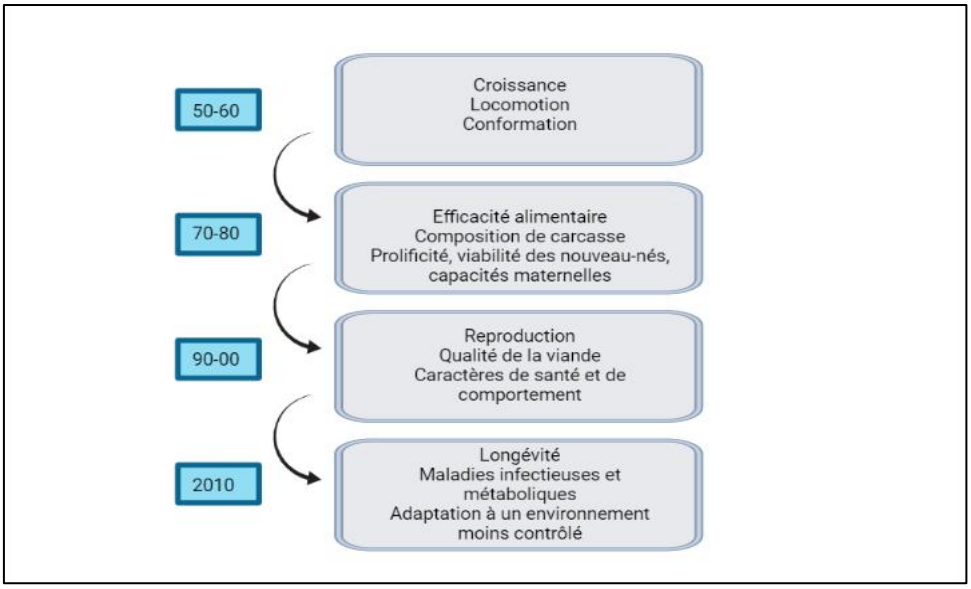


Figure 3 - Evolution historique des objectifs de sélection animale pour la production de viande (Phocas et al., 2016).

De plus, la sélection des animaux robustes s'explique également par le besoin d'une transition agroécologique des systèmes d'élevage. En effet, l'intensification de l'élevage a eu des conséquences négatives importantes sur l'environnement (la demande des ressources alimentaires a conduit à la déforestation massive, le recours aux intrants a entrainé des pollutions des eaux et des sols et augmenté l'émission de gaz à effets de serre) et sur le bien-être et la santé des animaux (accroissement de la densité d'animaux dans les bâtiments, mutilations comme coupe de queues ou la castration à vif des porcelets, élimination systématique de l'un des deux sexes dans certaines filières…) (Ducos et al., 2021). Par conséquent, les systèmes d'élevage doivent évoluer et une transition agroécologique de ces systèmes est une solution pertinente pour résoudre ces problèmes. Dumont *et al.* (2013) ont proposé cinq principes agroécologiques pour les systèmes d'élevage : i) réaliser une gestion intégrée de la santé et du bien-être des animaux ; ii) diminuer les intrants nécessaires à la production ; iii) optimiser le fonctionnement des systèmes et réduire les pollutions ; iv) gérer la diversité pour renforcer la résilience des systèmes d'élevage et v) préserver la biodiversité en adaptant les pratiques. Les actions ciblant l'amélioration de la robustesse des animaux sont une des solutions importantes pour le premier et le deuxième principe.

1.2.2 Définitions de la robustesse

L'origine du terme "robustesse" vient du mot "robuste", du latin *robustus* qui exprime le caractère "solide comme un chêne". Il exprime la capacité à résister (force, vigueur) et la capacité à " endurer l'effort, la fatigue, le mal " (Clément et al., 2018). Dans la suite de ce document, une revue de la littérature sur les définitions de la robustesse est présentée.

1.2.2.1 Synthèse sur les différents concepts de la robustesse

La robustesse est un caractère complexe qui n'a pas de définition standard unique et qui renvoie à différents concepts dans différentes disciplines : en statistique, la robustesse d'un estimateur désigne sa capacité à rester inchangé lorsqu'il y a un changement dans les paramètres du modèle choisi pour son estimation ou dans les données d'étude (Bodin et al., 2010) ; en ingénierie, la robustesse correspond à la capacité de l'organisation à survivre et à rester sous contrôle par l'émergence de nouveaux modèles organisationnels (Wybo, 2008); en science environnementale, un système est considéré comme robuste lorsque, en présence de perturbations externes, il continue à fonctionner sans changement fondamental de la configuration initiale (Dragicevic, 2016). Dans le monde du vivant, la robustesse d'une population sauvage est sa capacité de survie ou de maintien de l'espèce (Bodin et al., 2010). En production animale, initialement, de nombreuses définitions de la robustesse et d'autres termes liés à la robustesse (résistance ou résilience) ont été proposées (Table 1).

Les principales différences entre les définitions proposées reposent tout d'abord sur le système concerné : soit l'animal dans son intégralité (robustesse globale), soit une partie du système comme la réponse d'un phénotype à une variation de milieu (robustesse spécifique ou résilience). De manière générale, la robustesse globale fait référence à la capacité de survie de l'animal et au maintien d'un potentiel global de production ; cela comprend toutes les capacités biologiques des animaux et elles sont très difficiles à quantifier. En revanche, la robustesse spécifique est plus facile à quantifier. Mais elle ne correspond pas dans ce cas à l'intégralité du phénomène biologique sous-jacent (Friggens et al., 2010). La temporalité de la réponse aux perturbations est également une différence entre les définitions proposées. En effet, certains concepts de robustesse se basent uniquement sur la réponse de l'animal à un moment donné (Waddington, 1960; Kolmodin et al., 2002), tandis que d'autres concepts prennent en compte la réponse de l'individu sur toute sa vie (Colditz et al., 2016; Berghof et

al., 2018). En outre, pour certains auteurs la robustesse est la réponse d'un animal face à une perturbation qui se décompose en résistance (capacité de minimiser l'impact direct de la perturbation) et résilience (capacité de vite récupérer) (Blanc et al., 2013; Nguyen-Ba et al., 2020). Enfin, nous arrivons de plus en plus à un consensus : la robustesse globale fait plutôt référence à un maintien global des performances de l'animal à tous les niveaux et sur du long terme, alors que la robustesse spécifique ou la résilience fait référence à la réponse à court terme de l'animal face aux perturbations environnementales et concerne un nombre limité de phénotypes (Berghof et al., 2018).

Table 1 - Différentes définitions de la robustesse globale, robustesse spécifique, résilience et résistance.

Robustesse globale	Robustesse spécifique
Des animaux robustes sont des animaux qui combinent un potentiel de production élevé avec une résistance aux facteurs de stress externes. (Knap, 2005)	La robustesse est le maintien d'un potentiel de production spécifique dans des environnements variés. (Waddington, 1960)
La robustesse est la capacité de survie de l'animal et le maintien d'un potentiel global de production dans des environnements stressants. (Bodin et al., 2010)	La robustesse peut être considérée à l'échelle de l'expression d'un seul caractère. (Kolmodin et al., 2002)
La robustesse est définie comme la capacité, face aux contraintes environnementales, de continuer à faire les différentes choses que l'animal doit faire pour favoriser sa capacité future à se reproduire. (Friggens et al., 2017)	La robustesse, étant la capacité à faire face à une perturbation, est évaluée en fonction de la déviation générée sur un caractère de production et de la trajectoire suivie par ce caractère pour récupérer suite à l'arrêt de la perturbation. (Blanc et al., 2013)
La robustesse est la capacité d'adaptation individuelle en considérant divers critères de sélection, dont la production, la reproduction, la santé et le bien-être des animaux, et l'adaptation à des conditions environnementales fluctuantes. (Phocas et al., 2016)	La robustesse peut être définie comme la capacité d'un animal à minimiser les impacts des facteurs perturbateurs et à revenir rapidement à l'état antérieur à la perturbation. (Nguyen-Ba et al., 2020)

Résilience	Résistance
La résilience est la capacité d'un animal à être peu affecté par les perturbations ou à revenir rapidement à l'état qui prévalait avant l'explosion de la perturbation. (Colditz et al., 2016; Berghof et al., 2018)	La résistance est la capacité de l'animal hôte à exercer un contrôle sur la perturbation. (Bishop, 2012)
La capacité d'un animal à faire face à des perturbations environnementales, est appelée résilience (peu affecté par une perturbation, ou qui retrouve rapidement son niveau normal de fonctionnement). (Poppe et al., 2020)	La résistance est la capacité d'un animal à minimiser les impacts des facteurs perturbateurs. (de Goede et al., 2013; Nguyen-Ba et al., 2020)
La résilience fait référence à la productivité d'un animal face à une perturbation (Bishop, 2012)	
La résilience est le taux de récupération en cas de perturbation (Scheffer et al., 2015)	
La résilience aux maladies est définie comme la capacité d'un hôte à maintenir un niveau raisonnable de productivité lorsqu'il est confronté à une infection. (Hermesch et al., 2015; Putz et al., 2019)	

Dans tous les cas, la robustesse spécifique (ou résilience) traduit comment l'animal va répondre face à des contraintes environnementales. Cette réponse est dynamique et peut se décomposer en 3 grandes phases (Sauvant et Martin, 2010), qui considèrent la résilience comme une composante de la robustesse (Figure 4). Au début, le système est sur sa trajectoire d'équilibre. Ensuite, une perturbation apparait et influence l'évolution du système après une phase d'incubation (un délai), le système change d'état avec une amplitude dépendant de l'intensité de la perturbation et de la résistance de l'animal. Dans la phase suivante, le système tend à récupérer et revient au niveau avant l'exposition de la perturbation ou sur une nouvelle trajectoire. La vitesse de récupération dépend de la résilience du système. Enfin, à la fin de réaction de l'animal, la courbe évolue de manière stable sur un niveau soit inférieur (système flexible), soit équivalent (système élastique), soit supérieur au niveau du départ (système plastique) (Figure 5). Il existe enfin un cas extrême où aucune modification de l'état n'est observée, il s'agit du système rigide.

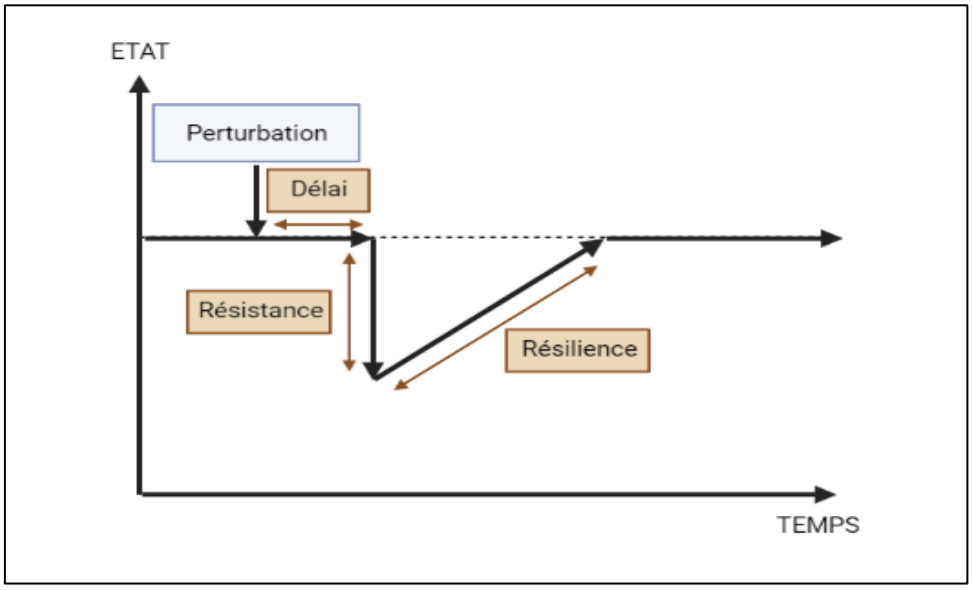


Figure 4 - Réponse dynamique d'un système à une perturbation

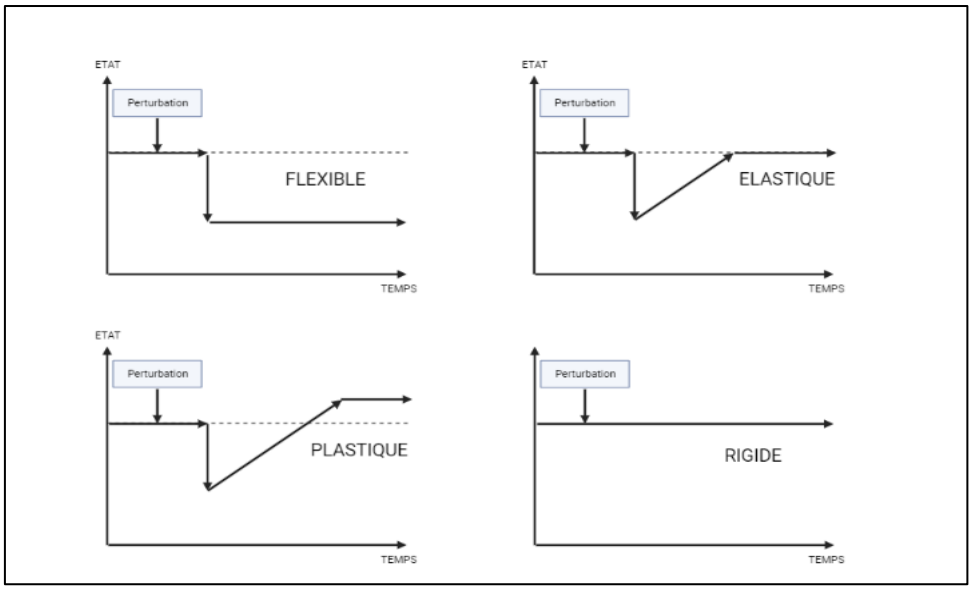


Figure 5 - Modalités des réponses adaptatives aux perturbations

Il n'est pas évident de choisir quelle réponse est la plus adaptée en fonction des systèmes ou des phénotypes. En effet, un animal peut avoir une bonne robustesse spécifique (peu sensible aux variations environnementales de court terme) et une mauvaise robustesse globale (par épuisement). L'analogie populaire de la robustesse globale est le roseau, une plante qui se plie facilement sous l'effet d'une rafale de vent (faible robustesse spécifique) et qui se redresse une fois la rafale passée et ne casse jamais. L'analogie populaire de la robustesse spécifique est le chêne qui ne plie pas (bonne robustesse spécifique) mais casse sur le long terme (Friggens et al., 2022). Selon cette théorie (chêne versus roseau), la corrélation génétique entre robustesse spécifique et globale devrait être négative. Cependant, d'autres avis existent : Poppe *et al.* (2020) ont montré que la corrélation entre les nouveaux indicateurs de la résilience (ou robustesse spécifique) et la longévité (qui représente une robustesse globale) était positive (0,28 à 0,34).

1.2.2.2 Hypothèses génétiques sous-jacente à la robustesse

Des différences de robustesse d'origine génétique s'interprètent comme de l'interaction génotype x environnement (GxE, l'effet de l'environnement est différent selon le génotype). Trois hypothèses génétiques ont été proposées pour expliquer la capacité des animaux à s'adapter aux variations environnementales (Bodin et al., 2010; Roy et al., 2019) :

le modèle de « surdominance », le modèle de « pléitropique » et le modèle « épistatique ». Le modèle de « surdominance » suppose qu'un animal sera plus robuste et s'adaptera mieux aux variations environnementales s'il est plus hétérozygote car cela lui offre un plus grand choix de voies métaboliques pour répondre aux divers environnements qu'un animal moins hétérozygote. Le modèle « pléitropique », quant à lui, repose sur l'hypothèse d'une expression différente des gènes selon l'environnement. Enfin, le modèle « épistatique » suppose qu'il existe un ou plusieurs gènes spécifiques qui contrôlent l'expression phénotypique des génomes dans les environnements différents.

1.2.3 Quantification de la robustesse

Afin d'introduire la robustesse dans les programmes de sélection, il est essentiel de pouvoir identifier et mesurer les caractères qui lui sont associés. Selon la définition utilisée pour la robustesse (globale ou spécifique), les méthodes de quantification peuvent être différentes. Dans la suite de ce document, nous présentons plusieurs stratégies qui ont été proposés dans la littérature pour caractériser et mesurer la robustesse des animaux d'élevage.

1.2.3.1 Robustesse globale

Combinaison des caractères fonctionnels

Concernant la catégorie « robustesse globale », la robustesse peut être considérée comme une combinaison de plusieurs fonctions sous-jacentes permettant « *face aux contraintes environnementales, de continuer à faire les différentes choses que l'animal doit faire pour favoriser sa capacité future à se reproduire* » (Friggens et al., 2017). En effet, les animaux d'élevage doivent de plus en plus faire « *face aux contraintes environnementales* » à cause des changements d'environnements dans lesquels ils sont élevés (en lien avec les demandes de la société et des éleveurs, et avec la transition agroécologique des systèmes d'élevage) et à cause du changement climatique entrainant plus souvent des phénomènes météorologiques extrêmes tels que les vagues de chaleur. Les animaux d'élevage doivent être capable de s'adapter à ces conditions environnementales variées et donc avoir une bonne robustesse.

D’après la définition présentée ci-dessus par Friggens *et al.* (2017), sélectionner pour une meilleure robustesse revient alors à sélectionner conjointement sur les différents caractères permettant de favoriser la capacité future à se reproduire (Figure 6). A minima, pour que l’animal se reproduise, il faut que ce dernier ait une croissance correcte (nécessaire pour atteindre la maturité sexuelle), qu’il survive jusqu'aux futures opportunités de reproduction (qu'il s'agisse de la mort naturelle ou de réforme), et qu’il soit fertile (réussite de la reproduction). Les critères de réforme au sein des élevages étant spécifiques à chaque espèce, au contexte de production et aux contraintes environnementales, différents autres caractères peuvent être pris en compte pour limiter les risques de réformes.

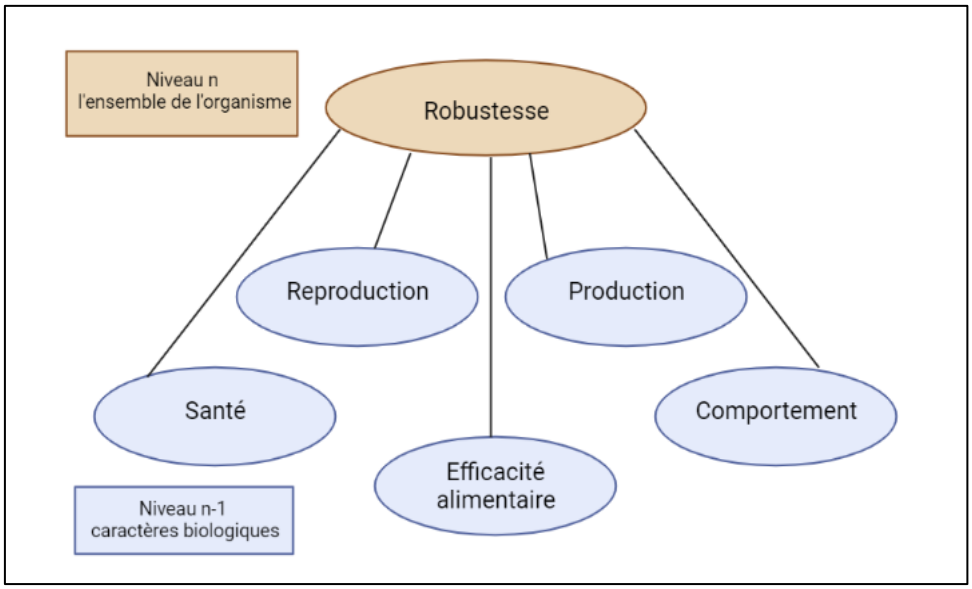


Figure 6 - Une représentation schématique des fonctions biologiques opérant aux niveaux sous-jacents (n-1) qui se combinent pour construire la robustesse au niveau (n) de l'animal (Friggens et al., 2017).

Des combinaisons précises de composants sous-jacents ont été proposées dans la littérature pour expliquer la robustesse. Par exemple, d’après Star *et al.* (2008), le concept de la robustesse des animaux d’élevage peut faire référence aux concepts de la santé (absence de maladie (Nordenfelt, 1995)), du bien-être (l'animal doit se sentir bien en étant libéré de toute peur prolongée et intense, de toute douleur et de tout autre état négatif, et en éprouvant des plaisirs normaux (Fraser et al., 1997)) et d’intégrité (intégralité et préservation

de l'animal (Rutgers et Heeger, 1999)) des animaux d'élevage. Plusieurs caractères de robustesse liée à ce concept sont considérés et ils couvrent les caractères comportementaux, immunologiques et physiologiques. Par exemple chez les volailles, les caractères comportementaux liés à la robustesse peuvent être la peur, le stress social, le picage de plumes et le cannibalisme ; les caractères immunologiques correspondent à ceux qui réduisent l'apparition de maladies comme les anticorps naturels déterminés dans le sang (l'animal a besoin d'un système immunitaire bien développé qui répond de manière adéquate aux agents pathogènes envahissants) ; enfin, un des caractères physiologiques liés à la robustesse peut être la solidité des os des animaux. D'autre part, Lenoir *et al.* (2022), quant à eux, ont proposé d'autres mesures de la robustesse des porcs en phase d'engraissement en combinant la santé et la production. Ces nouveaux phénotypes ont été définis à partir des mesures et des traitements médicaux enregistrés pendant la période de test (phase d'engraissement).

Il est ensuite important de connaitre la corrélation génétique entre les différentes fonctions pour mieux manipuler la robustesse par l'effet des multiples composants. Une explication possible est de supposer que l'interaction entre les fonctions et les caractères associés est considérée en termes d'allocation de ressources. Selon la théorie de l'allocation des ressources qui considère que les ressources sont réparties entre les différents caractères sous-jacents de la robustesse, la ressource utilisée pour une activité ne peut être utilisée pour une autre activité ou un autre processus. Ainsi, un animal ne peut pas maximiser simultanément l'expression de chaque fonction biologique. En conséquence on ne doit donc pas s'attendre à ce que les animaux soient robustes à tous les types de perturbations environnementales (sélectionner uniquement pour l'amélioration de la robustesse au stress thermique n'impliquera pas nécessairement l'amélioration de la robustesse à la charge pathogène par exemple).

Axe corticotrope

La robustesse a été considérée comme un ensemble de fonctions biologiques différentes dans la section précédente, cependant elle peut être quantifiée 'théoriquement' à un niveau plus élevé dans ce qui contrôle la répartition entre différentes fonctions, l'axe corticotrope. En effet, l'axe corticotrope est l'élément essentiel et privilégié du contrôle de la

réponse au stress (Bodin et al., 2010). Les principaux éléments de sortie de l'axe corticotrope sont le cortisol et l'hormone adrénocorticotrope (ACTH) qui exercent un large éventail d'effets sur le métabolisme, le système immunitaire, les processus inflammatoires et les fonctions cérébrales (Mormède et al., 2011). De nombreuses expériences ont montré une relation significative entre l'activité de l'axe corticotrope et les caractères liés à la production ou la robustesse. En effet, selon plusieurs auteurs, les taux de cortisol (ou corticostérone) plus élevés ont des effets négatifs sur les caractères liés à la production. Hennessy *et al.* (1988) ont montré que la vitesse de croissance et l'efficacité alimentaire sont négativement liées au cortisol. Des résultats similaires ont été obtenus chez le mouton, la consommation résiduelle étant directement proportionnelle au taux de cortisol (Knott et al., 2008). Chez le porc, plusieurs exemples montrent que le taux d'engraissement est influencé par le taux de production de cortisol (Foury et al., 2005, 2007). Au contraire, plusieurs auteurs ont prouvé que le cortisol avait des effets positifs sur les caractères liés à la robustesse et à l'adaptation. Michel *et al.* (2007) ont montré que les rats présentant la plus forte réponse de l'axe corticotrope s'adaptaient mieux au stimulus thermique. Leenhouwers *et al.* (2002), quant à eux, ont montré qu'une des caractéristiques biologiques corrélées positivement avec la survie des porcelets était la concentration de cortisol dans le sang prélevé à la naissance. Enfin, chez les volailles, des preuves expérimentales ont aussi montré que la sélection génétique pour l'intensité de la réponse au stress de l'axe corticotrope avait une influence complexe sur les réponses immunitaires et la résistance aux maladies (Gross, 1976). Par conséquent, une stratégie 'théorique' pour améliorer la robustesse globale consisterait donc à sélectionner des animaux dont l'activité de l'axe corticotrope est plus élevée. Cependant, cette stratégie est très difficile à mettre en œuvre puisque cela nécessiterait des prises de sang à répétition des animaux (une seule prise de sang ne suffit pas à caractériser un animal sur le fonctionnement de son axe corticotrope, il est nécessaire d'analyser l'évolution du taux de cortisol au cours du temps) afin de mesurer le cortisol, ce qui est impossible en routine. De plus, une augmentation de l'activité de l'axe corticotrope s'accompagne d'une diminution de la production des animaux (Phocas et al., 2014).

1.2.3.2 Robustesse spécifique

Contrairement à l'approche précédemment décrite qui concevait la robustesse comme le maintien d'une production globale, cette seconde approche est attachée à un phénotype particulier. Pour la génétique quantitative, cela correspond au maintien de l'expression d'un caractère donné (croissance chez les porcs, production laitière chez les bovins laitiers, production d'œufs chez les poules…) de façon stable face aux macro- ou micro-variations environnementales (modifications marginales des températures, de l'alimentation, voire des sources de stress) (Bodin et al., 2010). Plusieurs questions scientifiques ont été posées concernant la réponse au cours du temps des animaux face à une perturbation environnementale :

- Doit-on évaluer la réponse à un type de perturbation ou pour tout type de perturbation ?
- La quantification de la robustesse doit-elle se faire à un instant t ou à plusieurs moments différents tout au long de la vie de l'animal ?
- L'étude de la réponse dynamique face aux aléas devrait-elle être menée sur un ou plusieurs phénotypes différents ?

Mesurer la robustesse spécifique par quantification des interactions GxE

Le but est de comprendre les changements dans l'expression phénotypique des génotypes dans des milieux différents et quantifier la sensibilité environnementale des animaux d'élevage en utilisant des modèles génétiques et mathématiques.

Dans le cas où le nombre de milieux est faible (en général inférieur à 5), il est possible de quantifier les interactions GxE par un modèle multicaractère. Dans le cas où le nombre d'environnements considérés est plus élevé, le nombre de paramètres à estimer devient trop important et difficile à estimer avec le modèle multicaractère. Il existe alors deux façons pour réduire le nombre de paramètres à estimer : utiliser un modèle de norme de réaction lorsque les environnements sont connus (ou supposés connus après détection des perturbations) et mesurés selon un gradient, ou utiliser des modèles structuraux lorsque les environnements sont inconnus.

Les normes de réaction servent à expliquer la variation de l'expression phénotypique de chaque génotype dans les différents environnements par une fonction de quelques paramètres (par exemple avec une fonction polynomiale). Les valeurs génétiques associées aux paramètres de cette fonction permettent de caractériser la valeur moyenne et la sensibilité environnementale de chaque individu étudié. Cette méthode est efficace dans le cas des variations de milieux continues telles que la quantité de protéines dans la ration, la charge en agents pathogènes ou la température. Par exemple, en étudiant la résilience des animaux d'élevage à la variabilité climatique, Sánchez-Molano *et al.* (2019) ont quantifié la résilience des animaux par un modèle de norme de réaction sur **l'indice de température et d'humidité (THI)**.

En l'absence d'information sur le milieu, les modèles structuraux décomposent non seulement la moyenne mais aussi la variance des performances en effets génétiques, environnementaux et résiduels (Veerkamp et al., 2009). Dans ces modèles, les interactions GxE sont alors résumées sous la forme d'une valeur génétique sur la variance traduisant la variabilité des performances d'un génotype dans différents environnements. Plusieurs auteurs ont appliqué ces modèles : chez les porcs (Sorensen et Waagepetersen, 2003), chez les lapins (Garreau et al., 2008) ou encore chez les escargots (Ros et al., 2004).

Mesurer la robustesse spécifique par utilisation des données longitudinales

Une autre stratégie pour quantifier la robustesse spécifique est d'analyser la dynamique d'évolution d'un phénotype de l'animal au cours du temps face aux variations environnementales. Pour pouvoir analyser les réponses dynamiques au cours du temps de l'animal, il est indispensable de pouvoir collecter des données longitudinales (données de phénotypage haut-débit). Heureusement, depuis quelques années, l'automatisation et l'« électronisation » des élevages explosent dans de nombreux pays, donnant naissance à ce que l'on appelle l'élevage de précision. Ce dernier se développe dans les différentes filières animales pour faciliter la surveillance de troupeaux dans un contexte d'accroissement de la taille des élevages et de diminution de la disponibilité en main-d'œuvre et pour améliorer l'efficacité de la production par un meilleur ajustement de l'alimentation aux besoins individuels réduisant ainsi l'impact environnemental de l'élevage.

Les nouvelles technologies sont très diversifiées et se retrouvent dans plusieurs domaines pour des espèces différentes. Par exemple, avec le système de positionnement global (GPS) ou des capteurs de mouvements, il est désormais possible d'obtenir des informations concernant l'état de santé des animaux (Bahlo et al., 2019). Au Vietnam, la boiterie et le cycle d'œstrus des vaches peuvent être détectés en utilisant un accéléromètre de trois dimensions (Khanh et al., 2016). Chez les volailles, les maladies des pattes des poulets de chair peuvent aussi être détectées avec des caméras et un système numérique de détection de mouvement (Kristensen et Cornou, 2011). Chez les porcs, les technologies comme des microphones peuvent aider les animaliers à identifier et soigner des cochons ayant des maladies respiratoires en analysant les toux de ces animaux (Berckmans, 2014), et les distributeurs automatiques de concentrés (DAC) permettent d'alimenter individuellement chaque porc et ainsi améliorer le suivi des animaux. Ce développement important des outils au sein des élevages de précision a permis aux sélectionneurs de collecter un grand nombre de données mesurées de manière répétée et précise au cours du temps. Cela ouvre en même temps de nouvelles possibilités pour étudier et identifier de nouveaux critères plus pertinents de la robustesse des animaux d'élevage.

Perturbations

Avant d'aller plus loin dans l'analyse des réponses dynamiques des animaux d'élevage face aux perturbations afin de quantifier la robustesse, il semble intéressant de connaitre la caractérisation du contexte environnemental (c.-à-d. les principaux types de perturbations). En effet, certains environnements sont rudes mais relativement stables et certains environnements sont en moyenne bons mais avec des perturbations fréquentes. Colditz *et al.* (2016) pensent qu'il existe deux catégories de perturbations environnementales possibles. La première catégorie est fournie par des caractéristiques cycliques ou persistantes de l'environnement (facteurs stables tels que les conditions météorologiques moyennes du lieu, les conditions nutritionnelles, les contraintes imposées par le système agricole, la disponibilité des pâturages …). Cela fait référence à la dureté générale de l'environnement comme indiqué par Friggens *et al.* (2010). Dans la deuxième catégorie, les perturbations environnementales ont tendance à être sporadiques, épisodiques ou spécifiques à une situation. Elles peuvent être de nature différente, étant soit physiques (par exemple : maladie, stress thermique), soit

psychologiques (par exemple : nouvel environnement, facteur de stress social, interaction humaine). Dans la suite du document, nous examinons la robustesse spécifique des animaux aux perturbations de la deuxième catégorie.

Utilisation des données longitudinales

Dans les sections précédentes (1.2.3.1 et au début de cette section), les aspects temporels de la robustesse ont été ignorés. En effet, ces réponses sont considérées en tant qu'ensemble d'états stables et elles correspondent à la conséquence « finale » d'une adaptation à une contrainte (Friggens et al., 2010). En utilisant les données longitudinales, il est possible de mieux comprendre la réponse à travers le temps des animaux d'élevage. Plusieurs méthodes différentes peuvent être trouvées dans la littérature pour mesurer ce changement dynamique du phénotype et chacune des méthodes a une nature particulière.

Tout d'abord, de nombreux caractères ont été proposés dans la littérature pour quantifier directement la robustesse spécifique : Berghof *et al.* (2018) ont élaboré plusieurs indicateurs de la robustesse (le terme résilience utilisé dans leur article correspond à une robustesse spécifique). Selon ces auteurs, les indicateurs de robustesse appropriés basés sur un seul caractère en production animale pourraient être obtenus à partir des résidus obtenus en ajustant un modèle de régression aux données. Les phénotypes longitudinaux utilisés peuvent être de type productif ou non-productif (Laghouaouta et al., 2021). L'idée est de supposer qu'un animal peu robuste sera plus susceptible d'avoir une déviation importante par rapport à la tendance ajustée aux données lors d'une perturbation. Ces indicateurs dérivés des résidus sont :

- La variance des résidus qui donne une indication de l'impact des perturbations. Une variance faible (ou plus faible) traduit des animaux sans perturbations ou non influencés par les perturbations ; et une variance élevée (ou plus élevée) des animaux influencés par les perturbations.
- L'autocorrélation des résidus donne une indication de la durée de l'impact des perturbations (c'est-à-dire le taux de récupération). Une autocorrélation autour de 0 (c'est-à-dire que les déviations ne sont pas liées) fait référence aux animaux sans perturbations, non influencés par les perturbations, ou avec une récupération rapide

après les perturbations. Une autocorrélation vers +1 (c'est-à-dire que les déviations sont plus semblables) désigne les animaux influencés par les perturbations et avec une récupération lente après perturbations. Enfin, une autocorrélation vers -1 (c'est-à-dire que les déviations sont opposées) représente les animaux influencés par les perturbations et une réponse rapide et surcompensatrice aux perturbations, par exemple une croissance compensatoire.

- L'asymétrie des résidus rend compte de la gravité des perturbations environnementales subies par un individu : une asymétrie autour de 0 pour les animaux sans perturbations ou non influencés par les perturbations ; une asymétrie positive pour les animaux ayant des déviations positives dues à des réponses positives aux améliorations environnementales ; et une asymétrie négative pour les animaux ayant des déviations négatives dues aux perturbations.

Pour résumer, concernant la déviation par rapport à la tendance ajustée aux données longitudinales, les animaux moins résilients devraient avoir une plus grande variance, une autocorrélation positive et une asymétrie négative. En revanche, les animaux résilients devraient avoir une variance plus faible, une autocorrélation et une asymétrie autour de 0. Dans la littérature, ces indicateurs de robustesse ont été utilisés par plusieurs auteurs pour différentes espèces : par exemple chez les poissons (Mengistu et al., 2022), les porcs (Homma et al., 2021) ou les bovins laitiers (Poppe et al., 2020).

Une autre stratégie pour quantifier la robustesse spécifique est de tout d'abord déterminer la courbe théorique du phénotype longitudinal d'intérêt de l'animal (trajectoire théorique de l'animal s'il n'avait subi aucune perturbation) puis quantifier les déviations entre cette courbe et son phénotype observé (modifié par les perturbations). Plusieurs propositions ont été faites dans la littérature afin de reconstruire la courbe de production théorique idéale. La construction de la courbe théorique peut être facilitée lorsque les moments des perturbations potentielles sont connus. Par exemple, Revilla *et al.* (2019) ont construit la courbe théorique sur des porcelets en considérant la phase de sevrage comme un moment où les animaux subissent une perturbation. Un modèle dynamique a été construit sur la base de la loi de Gompertz-Makeham pour décrire les trajectoires de poids vif. L'utilité du modèle est qu'il fournit des paramètres biologiquement significatifs qui renseignent sur l'amplitude et la

durée de la perturbation, ainsi que sur le taux de récupération de l'animal. La construction devient plus difficile lorsqu'on ne dispose pas d'information sur les perturbations. Par exemple, Poppe *et al.* (2020) ont proposé différentes fonctions afin d'ajuster au mieux les courbes de lactation individuelles des vaches. La courbe de lactation théorique ajustée devrait être aussi proche que possible de la courbe qu'une vache aurait réalisée en l'absence de perturbations, car les déviations de la courbe construite représenteraient alors le plus d'informations sur les réponses aux perturbations. Par conséquent, différentes méthodes ont été explorées pour ajuster les courbes de lactation individuelles en utilisant les enregistrements quotidiens du rendement laitier : soit une modélisation paramétrique (la courbe de lactation de Wilmink (Wilmink, 1987) et la régression polynomiale quantile) soit une modélisation non-paramétrique basée uniquement sur la tendance du phénotype (la moyenne mobile et la médiane mobile). Toujours dans l'analyse des courbes de lactation des vaches laitières, Codrea *et al.* (2011), quant à eux, ont utilisé une fonction B-spline et ont introduit une pénalité sur la dérivée seconde de cette fonction pour reconstruire la courbe théorique (plus la pénalité est grande, plus on élimine les bruits sur les données). Le problème de cette technique est qu'elle corrige adéquatement les fortes perturbations mais pas les perturbations de faible intensité ou de courte durée. Nguyen Ba *et al.* (2020), quant à eux, en travaillant sur la robustesse des porcs en croissance, ont développé un modèle mécanistique pour estimer la trajectoire cible de la consommation d'aliments (quantité d'aliments qu'un porc souhaite consommer lorsqu'il n'est confronté à aucune perturbation).

Une fois la reconstruction de la courbe théorique effectuée, des indicateurs pertinents décrivant au mieux les déviations entre cette courbe et la courbe observée perturbée sont utilisés. Ces derniers peuvent une fois de plus être la variance, l'asymétrie, l'autocorrélation des résidus (Berghof et al., 2018) ou d'autre indicateur comme **l'indice de l'aire entre les courbes (ABC)** qui reflète la différence entre la courbe non-perturbée et perturbé (Revilla et al., 2021).

Détection des perturbations

Pour quantifier la réponse dynamique des animaux d'élevage face aux perturbations environnementales, certains auteurs ont eu besoin de détecter les moments où l'individu (ou le groupe d'animal) a subi une (ou plusieurs) perturbation(s). En étudiant la consommation

alimentaire cumulée, Nguyen-Ba *et al.* (2020) ont appliqué une fonction B-spline d'ordre 6 à la différence entre la courbe observée et à la courbe théorique. Toutes les périodes pendant lesquelles les données observées s'écartent de la courbe théorique pendant plus de 5 jours et pour plus de 5% ont été considérées comme le résultat d'une perturbation. D'après Garcia-Baccino *et al.* (2021), des variations phénotypiques accrues d'un groupe d'animaux peuvent être la cause de réponses individuelles extrêmes à un stress et peuvent ainsi constituer un indicateur d'une perturbation environnementale car les perturbations environnementales produisent des réactions différentes chez des animaux différents. En se basant sur cette logique, un modèle de mélange gaussien à deux composantes a été appliqué sur le coefficient de variation (CV) journalier des données pour séparer les jours « normaux » (sans perturbation, CV faible) des jours « stressants » (CV plus élevé).

1.2.3.3 Paramètres génétiques

Selon le phénotype utilisé pour quantifier la robustesse, une large « plage » d'héritabilité a été rapportée dans la littérature. Concernant la robustesse globale, pour les caractères liés à la survie ou la longévité fonctionnelle, les héritabilités sont généralement faibles. Par exemple, les trois caractères de la robustesse proposés par Knap (2005) (taux de survie avant sevrage, taux de survie des porcs en croissance et nombre de portées qu'une truie aura au cours de sa vie) ont une très faible héritabilité (<0,1) ; Engblom *et al.* (2010), quant à eux, ont analysé la longévité de reproduction chez les truies Landrace, Yorkshire et croisées et les héritabilités ont été estimées de 0,00 à 0,16 pour ces trois races de porcs. Pour les caractères liés à la santé, les héritabilités sont aussi généralement faibles. En travaillant sur la résistance aux maladies non spécifiques des lapins, Gunia *et al.* (2018) ont proposé trois caractères de maladie (les signes de maladie respiratoire, les signes de maladie digestive et un caractère composite prenant en compte les signes de maladie digestive, respiratoire et diverses maladies infectieuses) et trouvé que les héritabilités de ces caractères variaient de 0,04 ± 0,01 à 0,11 ± 0,03. Lenoir *et al.* (2022), quant à eux, ont trouvé des héritabilités inférieures à 0,1 pour tous les caractères proposés de la robustesse combinant la santé et la production chez les porcs en engraissement. Les héritabilités des différents caractères fonctionnels de robustesse des volailles présentés par Star *et al.* (2008) varient beaucoup. Les héritabilités estimées pour le comportement craintif et le comportement de picage des

plumes se situent entre 0,35 et 0,60, et 0,10 et 0,24, respectivement (Rodenburg et al., 2004). L'héritabilité estimée pour les anticorps naturels déterminés dans le sang est comprise entre 0,11 et 0,42 (Siwek et al., 2006). Les caractères décrivant la force des os sont modérément à fortement héritables (0,30-0,45) (Bishop et al., 2000).

Concernant les héritabilités liées à l'étude de l'interaction GxE, elles ne sont généralement pas élevées. Sánchez-Molano *et al.* (2019) ont trouvé que les héritabilités des nouveaux phénotypes pour la résilience des animaux d'élevage à la variabilité climatique variaient entre 0,09 et 0,11. Lung *et al.* (2020) dans une revue sur la génétique et génomique de la résilience chez les animaux d'élevage et d'aquaculture ont montré que les héritabilités des caractères proposés étaient faibles (médiane = 0,01 ; intervalle 0-0,10). Enfin, Garreau *et al.* (2008), en travaillant sur les lapins, ont trouvé que l'héritabilité estimée pour la variabilité environnementale du poids individuel à la naissance était 0,014.

Concernant les phénotypes proposés issus de l'analyse des données longitudinales, les héritabilités varient plus selon les auteurs. En analysant la résilience des porcs aux maladies naturelles, Putz *et al.* (2019) ont montré que l'héritabilité de la variance résiduelle issue des régressions linéaires de la consommation alimentaire et de la durée passée au distributeur alimentaire était modérée : 0,21 (±0,07) et 0,26 (±0,07), respectivement. Homma *et al.* (2021) ont aussi trouvé que l'héritabilité de la résilience était modérée et variait entre 0,31 et 0,40. En revanche, Mengistu *et al.* (2022), quant à eux, ont trouvé que l'héritabilité du nouveau phénotype de résilience (proposé à partir des résidus) était faible : de 0,10 à 0,12. Enfin, en travaillant sur les courbes de lactation, Poppe *et al.* (2020) ont montré que l'héritabilité de la variance, de l'autocorrélation et de la symétrie des résidus variaient entre 0,20 et 0,24, 0,08 et 0,10, et 0,01 et 0,02, respectivement selon la courbe théorique appliquée.

1.2.3.4 Corrélation entre robustesse et production

Selon la théorie de l'allocation des ressources, la corrélation entre la robustesse et les autres caractères productifs devrait être négative (Friggens et al., 2017). Une autre hypothèse pour expliquer une corrélation négative entre la robustesse et la production est que les systèmes qui évoluent pour accomplir des tâches spécifiques dans des conditions environnementales données présentent une sensibilité plus forte aux changements environnementaux (de Goede et al., 2013). En effet, les résultats de Bytyqi *et al.* (2007)

indiquent une sensibilité accrue à l'environnement des races bovines ayant un potentiel génétique plus élevé pour la production de lait.

Certaines études sont cohérentes avec ces théories. En étudiant la résistance des chèvres face aux nématodes gastro-intestinaux, Heckendorm *et al.* (2017) ont trouvé une corrélation génétique positive entre le nombre d'œufs de parasites dans les féces et le rendement laitier des chèvres (ce qui implique une corrélation négative entre la robustesse et la production). De même, en analysant la capacité adaptative des porcs face aux pathogènes, Clapperton *et al.* (2009) ont conclu que la corrélation génétique entre le **gain moyen quotidien (GMQ)** et les caractères immunitaires était négative chez le porc.

A l'inverse, d'autres études ont montré que la corrélation génétique entre la robustesse et la production était positive. En travaillant sur la réponse des porcs au syndrome reproducteur et respiratoire porcin, Hess *et al.* (2016) et Dekkers *et al.* (2017) ont démontré une corrélation génétique négative entre la croissance et le niveau de virémie ou de la charge virale (ce qui implique une relation favorable entre production et robustesse). Enfin, Garcia-Baccino *et al.* (2021) ont rapporté une corrélation négative (-0,46±0,21) entre la sensibilité aux facteurs perturbés et la production.

1.2.4 Organisation de l'étude de thèse

Le but de la thèse est de proposer un critère de la robustesse des animaux d'élevage qui serait plus pertinent que ceux déjà proposés dans la littérature. Ce nouveau phénotype devrait permettre de mieux expliquer la réponse dynamique des animaux face aux facteurs perturbés et rendre comparable les animaux n'ayant pas subi les mêmes perturbations environnementales. Ainsi, la méthode proposée devrait être applicable sur plusieurs cas différents selon les espèces ou les phénotypes longitudinaux choisis. Pour atteindre cet objectif, nous avons travaillé sur les données issues des porcs en phase d'engraissement (plus de détail dans la suite du document). Un schéma d'organisation de l'étude de thèse est présenté dans la Figure 7.

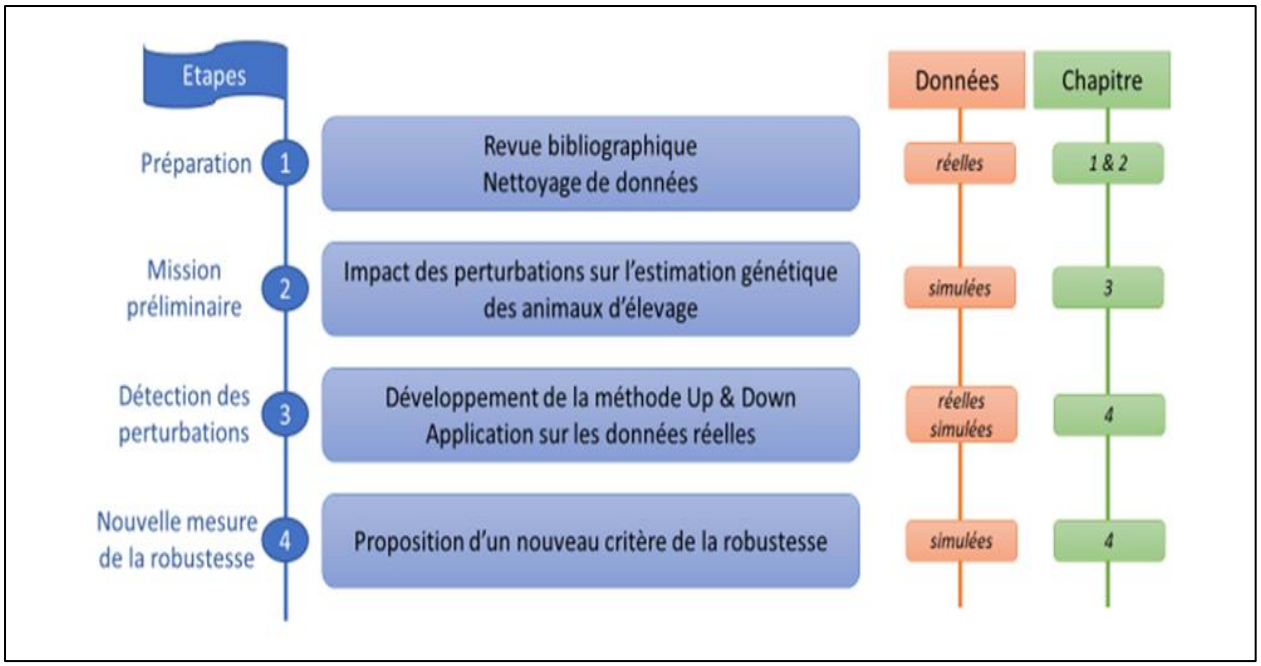


Figure 7 - Schéma d'organisation de l'étude de thèse.

Chapitre 2 - Données d'étude

2.1 Origine

Les données de la station de phénotypage du Rheu de l'année 2017 à 2019 ont été utilisées au cours de cette thèse. Dans cette station, les porcs de races pures, dont la plupart sont de race Large White, sont élevés dans différents bâtiments selon leur stade physiologique. De nouveaux porcelets (non castrés, queue réduite) provenant de plusieurs sources (élevages de sélection des OSP : Axiom, Nucleus ou Choice Genetics) arrivent toutes les 3 semaines (à environ 3 semaines d'âge). Les porcelets restent entre 5 et 7 semaines en 'post sevrage' selon les âges et poids de livraison. Ensuite, tous les porcs d'une même case passent au même moment en engraissement. Ils y resteront jusqu'à atteindre le poids vif cible de 120 kg (environ 100 jours, cette période est considérée comme la période d'observation) avant de partir à l'abattoir. Ces animaux sont pesés manuellement en cage de pesée 1 fois au début et 1 fois à la fin de la phase d'engraissement. Ils sont nourris avec un aliment conventionnel composé en majorité de blé (38,29% pour l'aliment croissance, et 42,57% pour l'aliment finition), de maïs (25% pour l'aliment croissance et finition) et d'orge (15% pour l'aliment croissance et finition).

Une représentation de la station de phénotypage du Rheu est présentée dans la Figure 8. Afin de garantir une bonne gestion des risques sanitaires, cette station est un élevage fermé : obligation de prendre une douche avant d'entrer dans la station ; changement de bottes entre salles ; protocoles stricts de nettoyage, désinfection du matériel et salles… D'autre part, les systèmes mis en place dans la station permettent de réduire l'impact environnemental de l'élevage : le raclage en V sépare les déjections directement sous les animaux et permet de réduire les émissions d'ammoniac et les nuisances olfactives pour le voisinage ; le laveur d'air dirige l'air extrait de la station vers un caisson de lavage pour y être traité et permet de réduire les émissions d'ammoniac, d'odeurs et de poussières en sortie du bâtiment.

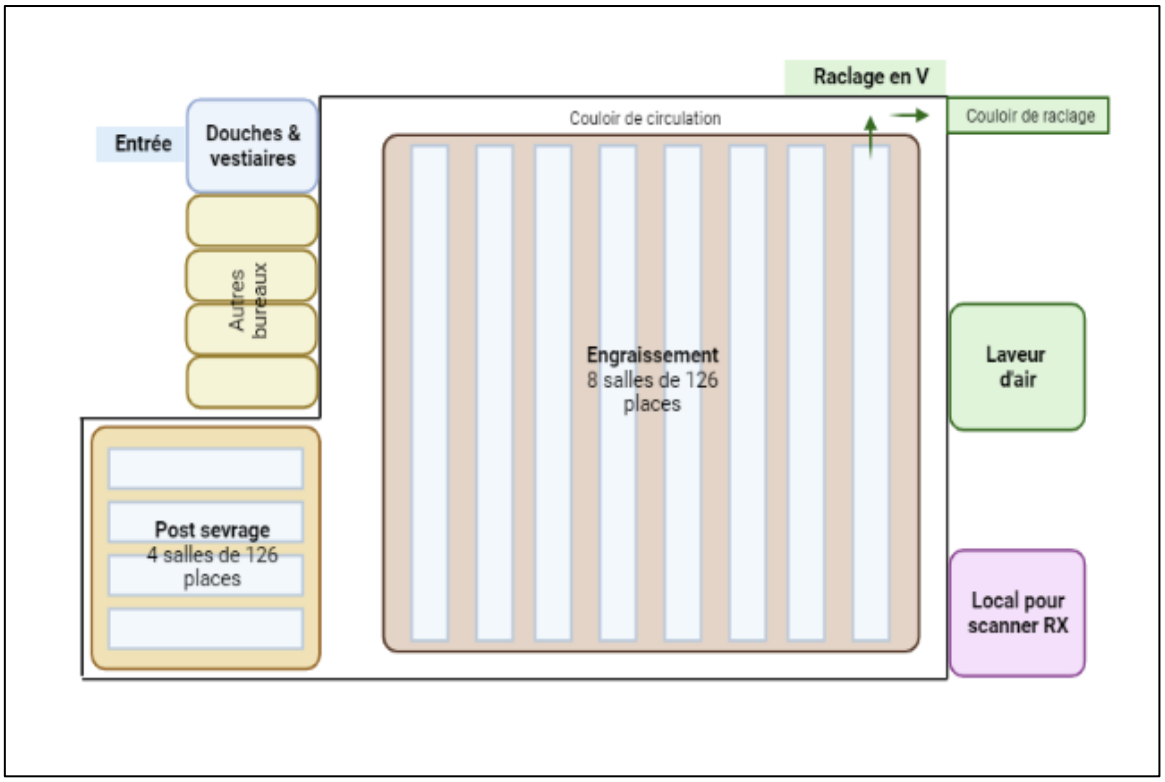


Figure 8 - Représentation de la station de phénotypage de l'Alliance R&D et INRAE au Rheu

Au total, 6298 porcs ont été étudiés de 2017 à 2019 (cela correspond aux données validées après un nettoyage de données, plus de détail concernant ce travail dans la suite du document). Ces animaux sont répartis en différentes races (Table 2). La race la plus nombreuse est la race LargeWhite (représentant environ 50% de la population totale), tandis que les races les moins nombreuses sont les races TaiZumu et Duroc (représentant respectivement moins de 1% et 3% de la population totale).

Table 2 - Répartition des porcs de la station de phénotypage au Rheu par race de 2017 à 2019

Race	Effectif
Landrace	1471
Pietrain	1434
LargeWhite	3180
Duroc	185
TaiZumu	28
TOTAL	**6298**

2.2 Données individuelles

2.2.1 Données issues des distributeurs automatiques de concentrés

Les données de comportement alimentaire et de croissance des porcs en croissance (phase d'engraissement) distribués dans des bandes (chaque bande correspond à une salle) et dans des cases-intra-bande différentes à l'intérieur du bâtiment ont été utilisées au cours de cette étude. Ces données sont issues des **distributeurs automatiques de concentrés (DAC)** de marque Acemo®. Un schéma de la répartition des porcs dans les cases-intra-bande équipées d'un DAC est présenté dans la Figure 9. Pour chaque visite au DAC, plusieurs phénotypes individuels sont automatiquement enregistrés par lecture de la puce **RFID (puce radio-identification)** de l'animal : la date et l'heure d'entrée au DAC, la date et l'heure de sortie du DAC (ces dernières informations permettant de calculer le temps passé individuellement au DAC), la consommation alimentaire en grammes, et le poids vif de l'animal en grammes (avec une précision de 1g). Le DAC peut contenir au maximum 1200g de granulés. Ce dispositif se recharge lorsque la nourriture restante est inférieure à 400g (s'il reste moins de 400g dans le DAC et que l'animal veut manger davantage, il doit ressortir et revenir dans le DAC). L'entrée du DAC est aménagée avec une porte évitant que deux porcs entrent simultanément sauf en début de période d'engraissement pour laquelle la porte du DAC est laissée ouverte pour que les porcelets apprennent à utiliser le dispositif, dans ce cas il y a possibilité de chevauchement des animaux. Enfin, un dispositif anti-confort empêche les porcs de s'endormir à l'intérieur du DAC.

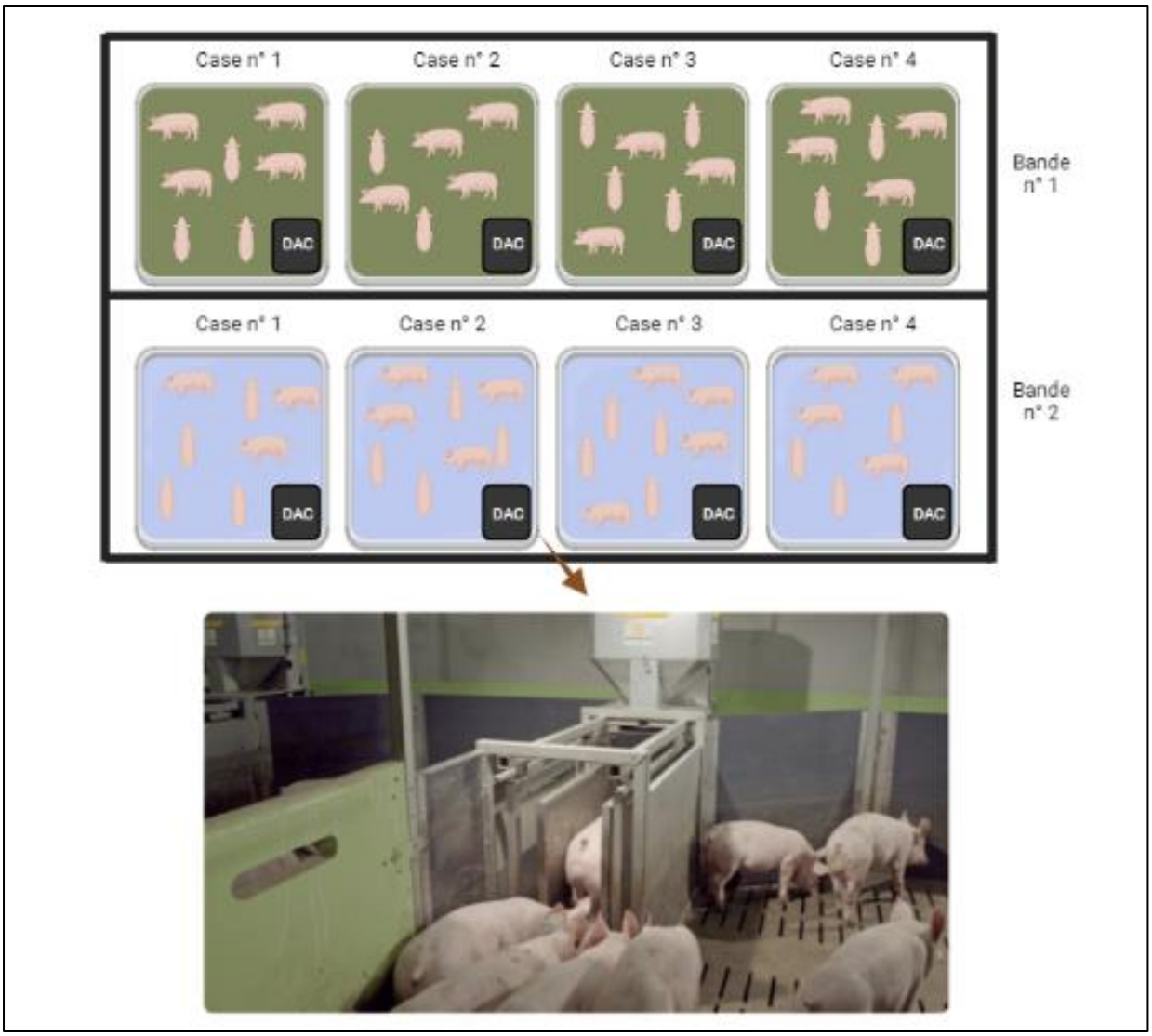


Figure 9 - Répartition des porcs dans des bandes et cases-intra-bande équipées d'un distributeur automatique de concentrés.

2.2.2 Informations complémentaires sur les animaux

Pour chaque porc, les animaliers enregistrent également les informations suivantes :

- La date de naissance
- L'âge d'entrée en salle d'engraissement (en moyenne 62 jours)
- Le poids d'entrée en salle d'engraissement : pesée réalisée sur une balance indépendante des plateaux DAC et qui sert de mesure de référence (utile ensuite pour le nettoyage de données)
- Le poids de fin de la phase d'engraissement : en théorie 120 kg (pesée réalisée manuellement sauf pour les porcs enlevés avant la fin de période d'engraissement (malade, mort ou autre raison))
- La date de départ à l'abattoir : en théorie le lendemain de la fin de la phase de croissance il y a un départ à l'abattoir toutes les semaines et les animaliers doivent s'organiser pour gérer tous les départs le plus rapidement possible (il est souvent difficile de gérer toutes les sorties en une seule journée).
- La cause de réforme : un code de validation qui permet d'indiquer comment le porc est réformé avant, pendant ou après la période d'observation. Avant le début et pendant la période d'observation, il existe plusieurs causes de réforme : mort (maladie respiratoire, maladie du système nerveux, diarrhée, …), croissance insuffisante, bande annulée ou d'autres diverses raisons. Après la période d'observation, l'animal est indiqué s'il est mort à l'abattoir ou malheureusement pendant le transport.
- Les données de santé de l'animal : Les **données santé (DS)** fournissent les informations concernant les traitements reçus par les porcs. L'identifiant, la bande, la case, la date de traitement, la famille du traitement, la maladie de l'animal et d'autres éléments sont collectés. En total, 4919 porcs différents ont reçu au moins un traitement.

Toutes ces données sont enfin fusionnées avec les données DAC pour compléter les informations des porcs étudiés.

2.3 Données d'environnement

Les données météorologiques du Rheu par jour de l'année 2017 à 2019 ont été récupérées depuis le site Internet météo-France. Les données concernant la température et le taux d'humidité par jour sont disponibles dans cette base de données. L'indice de température et d'humidité à l'échelle du jour (potentiellement utile pour détecter des perturbations environnementales) a été ensuite calculé par la formule suivante (Zumbach et al., 2008):

$$\mathrm{THI}\ (°\mathrm{C}) = \mathrm{T}\ (°\mathrm{C}) - (0{,}55 - 0{,}0055 \times \mathrm{HUM}) \times (\mathrm{T}\ (°\mathrm{C}) - 14{,}5),$$

où T est la température en °C et HUM est l'humidité relative.

2.4 Mise en forme et nettoyage des données

2.4.1 Données de santé et météorologiques

Dans l'élevage porcin, les perturbations environnementales ayant des répercussions sur les animaux peuvent apparaitre à plusieurs niveaux : à l'échelle de la bande, de la case ou de l'animal. Une perturbation de bande est une perturbation qui peut impacter la totalité des animaux élevés à l'intérieur d'une bande. Cela peut être dû à un changement de température à l'intérieur du bâtiment d'élevage (canicule) ou un changement d'alimentation à l'échelle de la bande. Les animaux concernés vont subir une perturbation de bande en commun et devoir répondre à cette perturbation selon leur robustesse. Ensuite, une perturbation de case est une perturbation qui impacte l'ensemble des animaux d'une case. Il existe plusieurs exemples concernant ce type de perturbation comme le cannibalisme, une panne de DAC, un virus peu contagieux… Enfin, une perturbation individuelle impacte seulement un animal, elle peut être une maladie non infectieuse, une patte cassée par exemple.

Les informations météorologiques et celles sur la santé des animaux peuvent être utilisées pour obtenir des informations sur les éventuelles perturbations auxquelles les animaux ont dû faire face qui pourront être confrontées aux résultats que nous obtiendrons en analysant les données de phénotypage haut-débit pour quantifier la robustesse des animaux. Les conditions perturbées sont déterminées en identifiant les traitements curatifs des porcs et en isolant les jours avec un THI important (supérieur à 25 °C).

2.4.1.1 Données de santé

L'objectif est de différencier les traitements préventifs (pas considérés comme traitements pour des perturbations) des traitements curatifs et dans ce dernier cas identifier si ces derniers étaient des traitements à l'échelle de la bande (traitements pour des perturbations de bande), des traitements à l'échelle de la case (traitements pour des perturbations de case) ou des traitements à l'échelle de l'individu (traitements pour des perturbations individuelles). Huit familles de traitements ont été identifiées : antibiotique, vermifuge, vitamine, vaccin, anti-inflammatoire, analgésique, anti-infectieux et antihémorragique. Tous les traitements de la famille vermifuge, vitamine et vaccin sont considérés comme des traitements préventifs et ne sont pas pris en compte parmi les traitements pour des perturbations potentielles.

Traitement à l'échelle de la bande

L'identification des traitements de bande se fait sur les familles restantes (considérés comme des traitements curatifs). Une bande est considérée comme perturbée si, au moins à un jour donné, toutes les cases de cette bande ont reçu un traitement de la même famille et au moins 50% des animaux de cette bande ont reçu aussi un traitement de la même famille.

Traitement à l'échelle de la case

Une case est considérée comme perturbée si, au moins à un jour donné, au moins 50% des animaux de cette case ont reçu un traitement de la même famille à l'exclusion des traitements de bande préalablement identifiés.

Traitement à l'échelle de l'individu

Le reste des traitements curatifs sont considérés comme des traitements pour des perturbations individuelles. Les traitements qui correspondent aux traitements de bande et de case ne sont pas considérés comme des traitements individuels.

2.4.1.2 Données météorologiques

La moyenne glissante sur 3 jours des valeurs de THI a été calculée afin de pouvoir identifier les jours correspondants a des fortes valeurs de THI sur une durée suffisamment longue pour avoir un impact sur les animaux. Les jours ayant un THI moyen sur 3 jours supérieur ou égal à 25 sont considérés comme des jours perturbés (ce choix de seuil a été validé par les professionnels). Par conséquent, les bandes qui sont observées pendant ces jours-là vont être considérées comme des bandes perturbées.

Au total, 26 perturbations potentielles à l'échelle des bandes (15 avec la base de données santé et 11 avec la base de données météorologiques), 16 perturbations potentielles à l'échelle des cases et 1786 perturbations potentielles individuelles ont été identifiées. Ces dernières serviront à valider celles qui seront mises en évidence par l'analyse des données de phénotypage haut-débit.

2.4.2 Données issues des DAC

2.4.2.1 Données de comportement alimentaire

Même si un pré-nettoyage est effectué par l'IFIP (exclusion et correction de certains poids et quantités d'ingéré, identification des problèmes concernant les DAC), les données des DAC ont été validées à l'échelle de la visite avant d'être converties à l'échelle de la journée.

Avant le nettoyage de données, la **base de données (BDD)** initiale du DAC contient des informations sur 6304 porcs ; le nombre total de visites au DAC est de 6 773 462. La **consommation alimentaire (FI)**, **le temps passé au DAC (TI)** et **la vitesse d'ingestion (VI)** obtenue par VI=FI/TI, sont les caractères concernant l'alimentation qui nous intéressent. Le tout premier nettoyage consiste à supprimer toutes les visites ayant un FI nul. Après ce nettoyage initial, il reste 94% des données brutes.

Ensuite, pour le reste des visites, **FI**, **TI** et **VI** ont été nettoyés simultanément à l'échelle de la visite. Plus précisément :

- Les visites d'un jour donné ayant un FI inférieur à 10g ont été regroupées ensemble pour former une seule visite (d'après les experts, les porcs ne devraient pas manger trop peu à

chaque visite et il est probable qu'un porc soit interrompu à cause d'un autre porc au milieu de sa visite au DAC).

- Pour ces 3 phénotypes d'intérêt, des seuils indiquant si les données observées étaient aberrantes ont été déterminés (en analysant la distribution des données c.f. Figure 10) et validé par les professionnels de l'IFIP. Les données étaient alors considérées comme aberrantes si FI > 1200g ; si TI < 20s ou TI > 30min (c.-à-d. si log(TI) ∉ [3 ;7,5]); ou si VI < 0,1g/s) ou VI > 3,3g/s (c.-à-d. log(VI) ∉ [-2,3 ;1,2])).

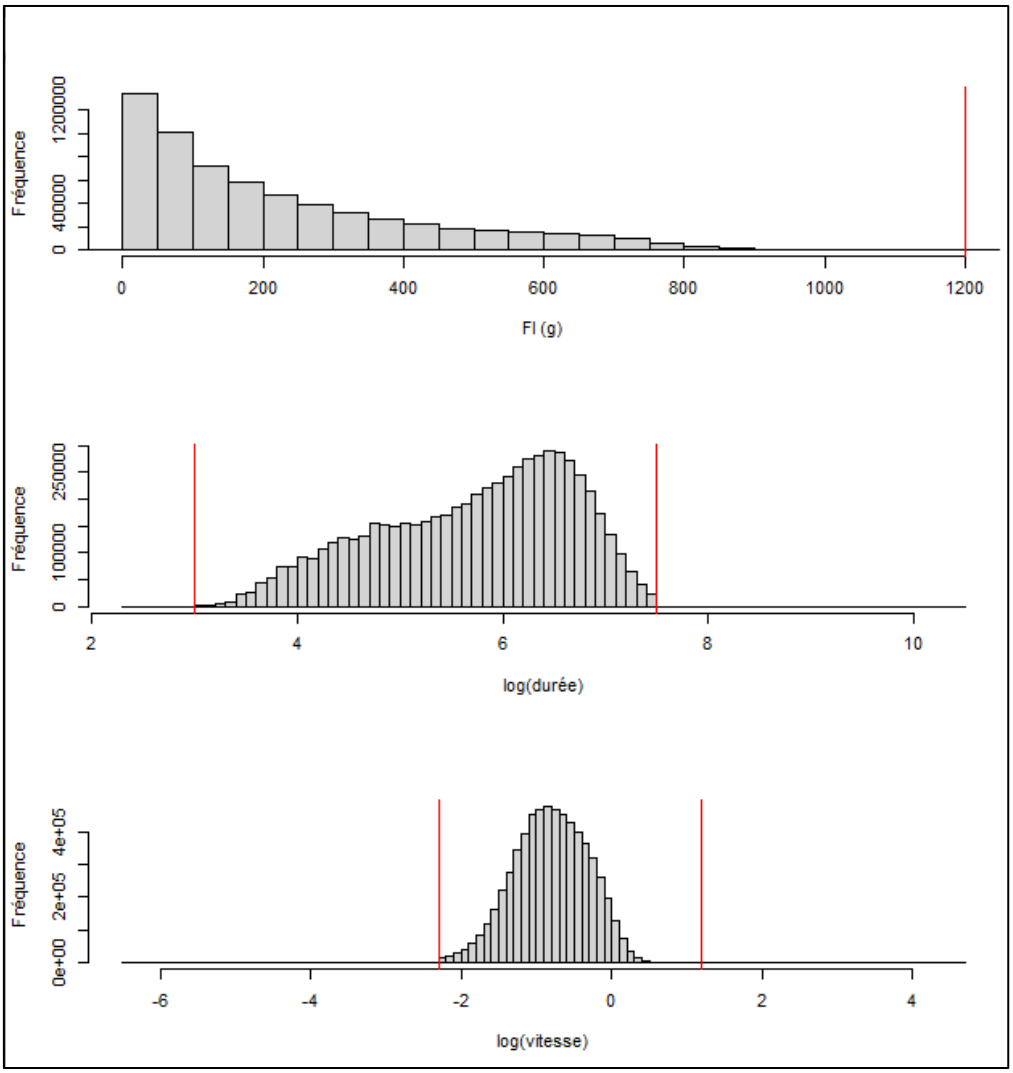


Figure 10 - Histogramme de la consommation alimentaire (g), du logarithme du temps passé au DAC et du logarithme de la vitesse d'ingestion des porcs étudiés à l'échelle de la visite. Les lignes rouges correspondent aux seuils pour déterminer les observations aberrantes.

- Huit différents cas possibles concernant la validation des données sont considérés, comme présentés dans la Table 3.

Table 3 – Pourcentages dans la base de données des différentes combinaisons des caractères de comportements alimentaires selon leur validité

Cas	log(VI) ∉ [-2,3 ; 1,2]	FI>1200g	log(TI) ∉ [3 ; 7,5]	Pourcentage
1	Non	Non	Non	99,14%
2	Non	Oui	Non	<0,01%
3	Non	Non	Oui	0,33%
4	Non	Oui	Oui	<0,01%
5	Oui	Non	Non	0,51%
6	Oui	Oui	Non	<0,01%
7	Oui	Non	Oui	0,02%
8	Oui	Oui	Oui	<0,01%

Si log(VI) ∉ [-2,3 ; 1,2], la vitesse d'ingestion est aberrante. Si FI>1200g, l'ingéré mesuré est aberrant. Si log(TI) ∉ [3 ; 7,5], le temps passé au DAC est aberrant.

La plupart des données sont considérées comme valides : le cas 1 où aucun des trois phénotypes concernés n'est aberrant correspond à 99,14% des enregistrements de la BDD totale (BDD brute après le nettoyage initial). Elles ont été conservées pour la suite des analyses. La base de données finale représente donc 93% des données de la BDD initiale. Elle concerne les informations de 6298 porcs (dont 3180 porcs Large White) répartis dans 52 bandes et 467 cases différentes entre l'année 2017 et 2019 ; le nombre total de visites au DAC est de 6 306 079. Il y avait en moyenne 9 cases par bande et 13 individus par case et chacun des individus était observé 100 jours en moyenne.

Les données sont ensuite converties à l'échelle du jour : la moyenne de VI et TI par visite est calculée pour chaque individu du premier au dernier jour de sa période de suivi ; la consommation alimentaire journalière (la somme de FI par visite) est calculée pour chaque individu pour tous les jours de sa période de suivi et la **consommation alimentaire cumulée (CFI)** est ensuite calculée en fonction de ces sommes comme $CFI_1 = FI_1 \; et \; CFI_t = CFI_{t-1} + FI_t$. A cette étape, il est possible d'avoir des 'trous' dans les données (jours où il n'y a pas de données). Pour résoudre ce problème, nous avons effectué une interpolation

linéaire : pour tous les porcs qui possèdent au moins un jour sans observation, les valeurs manquantes ont été imputées par une fonction linéaire appliquée au phénotype longitudinal concerné. Les statistiques descriptives pour les caractères concernant la consommation alimentaire par race de porcs disponibles sont présentées dans la Table 4.

Table 4 - Statistiques descriptives de la consommation alimentaire par race de porcs

Race	Temps (s) moyen passé au DAC par visite	Consommation alimentaire (g) par visite	Vitesse d'ingestion moyenne (g/s) par visite	Nombre de visites par jour	Nombre de jours d'observation
Duroc	597 ± 229	302 ± 139	0,50 ± 0,20	9 ± 4	96 ± 11
Landrace	500 ± 202	274 ± 142	0,55 ± 0,23	10 ± 4	103 ± 11
LargeWhite	495 ± 217	246 ± 128	0,50 ± 0,21	11 ± 5	102 ± 11
Pietrain	561 ± 207	298 ± 130	0,53 ± 0,20	9 ± 3	98 ± 12
Taizumu	537 ± 248	267 ± 127	0,50 ± 0,20	9 ± 5	114 ± 10

Moyenne ± écart-type des différentes variables

2.4.2.2 Poids vif

Pour chaque porc étudié**, un poids d'entrée en salle d'engraissement ($W0$)** est utilisé comme repère pour valider les données de poids vif de l'animal à l'échelle de la visite. La valeur de $W0$ sert à valider les poids pesés automatiquement avec le DAC au premier jour de la période d'observation : tous les poids par visite qui sont en dehors de l'intervalle [0,85x$W0$; 1,25x$W0$] sont considérés comme aberrants (ces seuils ont été validés par les professionnels). Ensuite, la moyenne des poids considérés comme valides au jour $j > 0$ est calculée et est utilisée comme **le poids de référence du jour *j* (PR_j)** pour valider les poids du jour $j + 1$. Tous les poids du jour $j + 1$ qui sont en dehors de l'intervalle [0,85xPR_j ; 1,25xPR_j] sont considérés comme non-valides et ne sont pas utilisés pour calculer PR_{j+1}. Le processus se répète jusqu'au dernier jour de la période de contrôle. Si au jour j, il n'y a aucun poids qui est valide, alors $\mathrm{PR}_j = \mathrm{PR}_{j-1}$. Par exemple, si au jour 15, tous les poids sont considérés comme aberrants en utilisant le poids de référence PR_{14}, le poids de référence PR_{15} sera supposé égal à PR_{14} et sera utilisé pour valider les poids au jour 16. Une représentation graphique illustrant la transformation des données de poids brutes (non-corrigés) par visite aux données de poids corrigés par jour (les PR_j) est présentée dans la Figure 11 pour un animal donné.

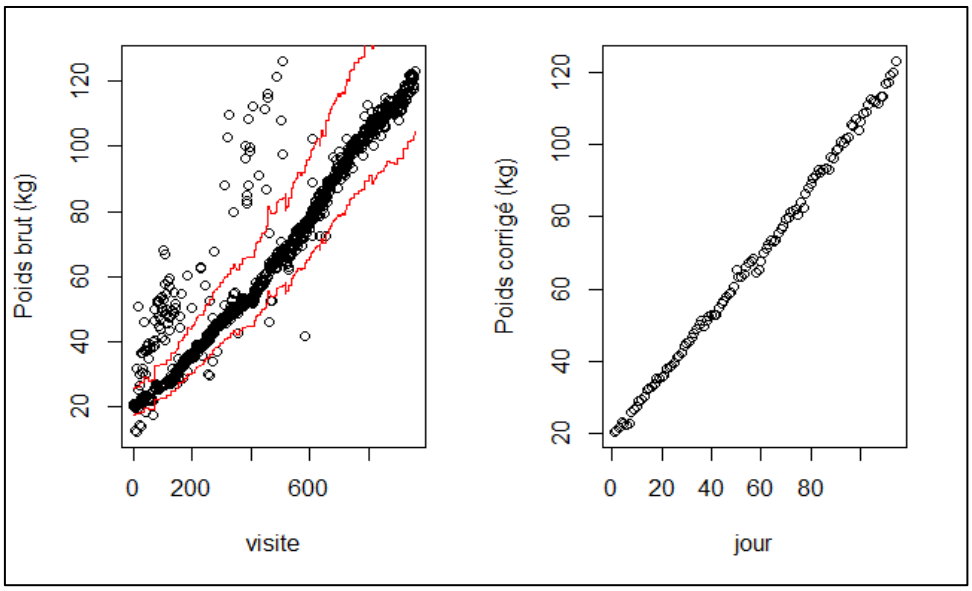


Figure 11 - Illustration de la correction du poids pour un porc donné

Figure de gauche : les poids bruts par visite (les lignes rouges représentent l'intervalle [0,85xPR_j ; 1,25xPR_j], Figure de droite : les poids corrigés par jour.

Cependant, il existe des cas spéciaux que nous n'arrivons pas à corriger complètement, par exemple le cas où il y a un biais systématique pour un jour ou une période donnée qui reste dans les bornes (Figure 12). En effet, il est possible que plusieurs porcs aillent au même DAC en même temps, ce qui fausse les données concernant le poids des porcs. Dans ce cas, l'évolution corrigée du poids vif de l'animal va malheureusement suivre les poids 'biaisés' car le PR_j est calculé par la moyenne de l'ensemble des poids valides. Les individus pour lesquels le poids n'était pas correctement corrigé ont été détectés en analysant les pentes du poids vif, pour chaque animal i, entre le jour j et $j + 10$ avec $j = 1, \ldots n_i - 10$ et n_i est le nombre d'observations de l'animal i. Les animaux ayant une pente maximale fortement supérieure à la moyenne de la population (analyse de la distribution de l'ensemble des données) ont été considérés comme 'mal corrigés'. Au final, il y a 64 sur 6298 animaux 'mal corrigés' (c.-à-d. cas où il existe des poids biaisés), soit 1% des animaux. Ces animaux ont été exclus des analyses.

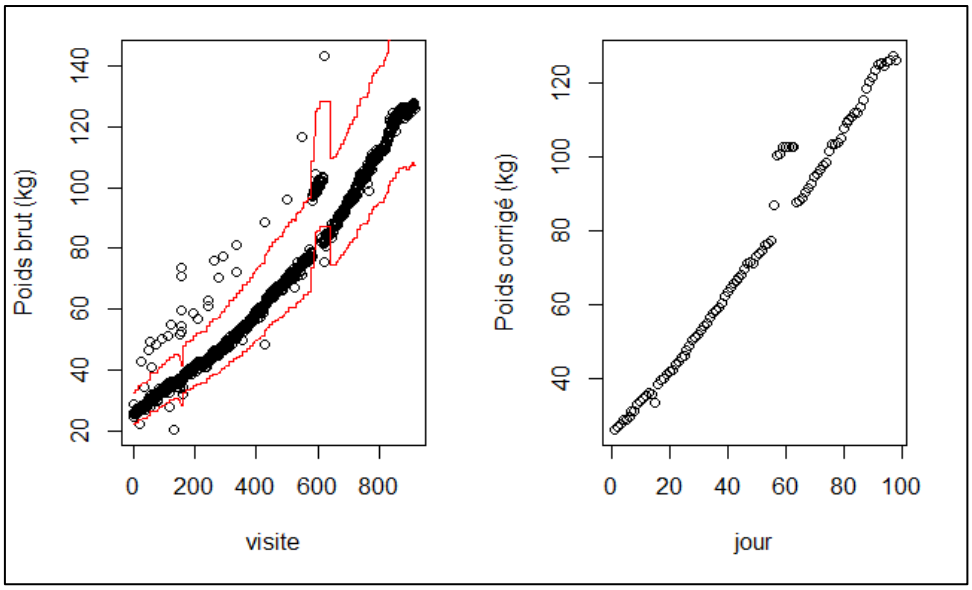


Figure 12 - Exemple de poids mal corrigé

Figure de gauche les poids bruts par visite (les lignes rouges représentent l'intervalle [0,85xPR_j ; 1,25xPR_j] pour valider les poids), Figure de droite les poids corrigés par jour.

2.4.3 Statistiques descriptives des données longitudinales

2.4.3.1 Poids

Concernant les différentes races de porcs disponibles dans notre base de données d'études, une nette différence peut être observée entre les plus lourdes et les moins lourdes. La race la plus lourde est la race Duroc (poids moyen de 26kg au jour 1 et de 110kg au jour 85) et la race la moins lourde est la Taizumu (poids moyen de 21kg au jour 1 et de 93kg au jour 85) ; le jour 85 étant le dernier jour où il y a au moins 90% de la population phénotypée.

D'autre part, il existe une diminution de la variabilité du poids entre les individus au cours du temps pour toutes les races. Pour le poids, la baisse du coefficient de variation entre les individus entre le jour 1 et le jour 85 varie entre 38% et 51% selon les races. La Figure 13 représente l'évolution moyenne du poids par jour et de son coefficient de variation par jour par race.

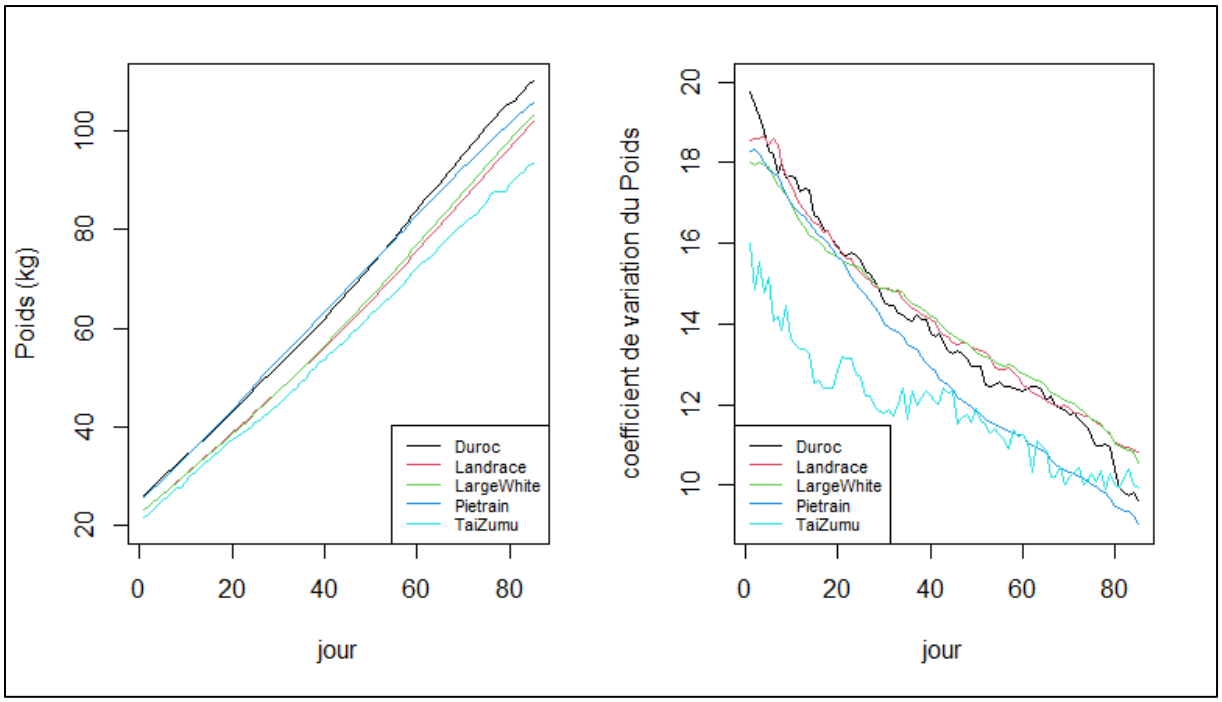


Figure 13 - Evolution moyenne du poids (à gauche) et de son coefficient de variation (à droite) par jour et par race

2.4.3.2 CFI

Une nette différence de consommation est également observée entre races. Les races qui consomment le plus sont les races Duroc et Piétrain (CFI moyenne de 0,61kg pour ces deux races au jour 1, et de 187kg et 186kg au jour 85, respectivement pour la race Duroc et Piétrain). Les races qui mangent le moins sont les races Taizumu et Landrace (CFI moyenne de 0,54kg et 0,53kg au jour 1, et de 169kg et 178kg au jour 85, respectivement).

Une diminution de la variabilité de la consommation alimentaire entre individus au cours du temps peut être observée pour toutes les races. La baisse du coefficient de variation entre les individus entre le jour 1 et 85 varie entre 72% et 79% selon les races. La Figure 14 représente respectivement l'évolution moyenne de la CFI et de son CV par jour et par race.

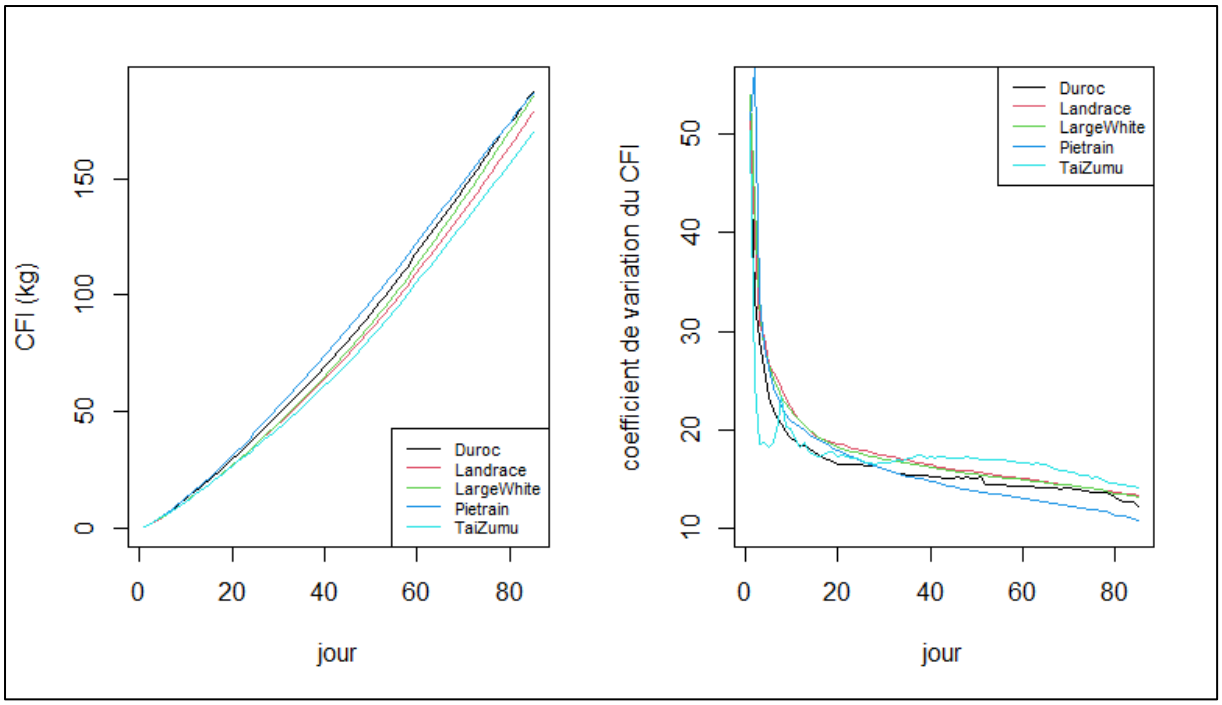


Figure 14 - Evolution moyenne de la CFI (à gauche) et de son coefficient de variation (à droite) par jour et par race

2.4.3.3 Vitesse d'ingestion

Les courbes de VI sont plus bruitées graphiquement que celles du poids vif et de la CFI. Aucune tendance concernant le coefficient de variation par jour pour toutes les races de porcs disponibles n'a été mise en évidence. La Figure 15 représente respectivement l'évolution moyenne de la vitesse d'ingestion par jour des races de porcs disponibles (graphique à gauche) et l'évolution du coefficient de variation concernant la vitesse d'ingestion de la population totale (graphique à droite) par jour.

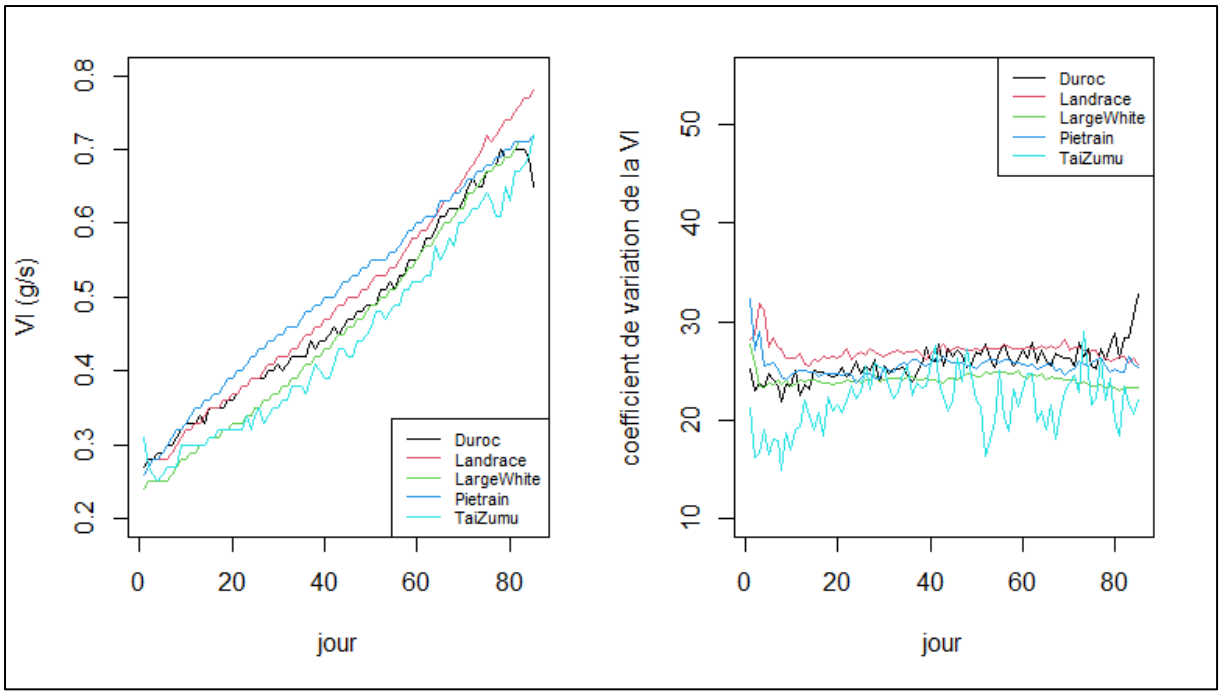


Figure 15 - Evolution moyenne de la VI (à gauche) et de son coefficient de variation (à droite) par jour et par race

Chapitre 3 - Impact des perturbations environnementales sur l'estimation des paramètres et valeurs génétiques

3.1 Introduction

Puisque le phénotype de production (croissance, production de lait…) qu'un animal exprime en présence de perturbations est plus ou moins modifié en fonction de sa robustesse, il est communément admis que les **valeurs génétiques (VG)** prédites pour un phénotype mesuré en milieu perturbé sont le reflet d'une combinaison des potentiels génétiques de production et de robustesse des animaux. Leur utilisation en sélection permettrait alors d'effectuer une sélection conjointe sur les deux caractères. Ceci est peut-être vrai dans le cas où l'ensemble des animaux phénotypés est soumis aux mêmes perturbations mais est moins trivial dans le cas où, comme en sélection porcine, l'environnement de sélection est fortement contrôlé avec des aléas sporadiques contrairement à celui de production. Dans une telle situation, la question scientifique qui se pose est la suivante : le potentiel génétique estimé, dans le cas d'un phénotypage dans un environnement bien contrôlé avec peu de perturbations, permet-il de combiner la robustesse et la production et donc d'assurer une sélection permettant une bonne production dans un environnement plus soumis aux aléas ?

Le but de cette première étude était donc d'évaluer l'impact des faibles perturbations aléatoires sur l'évaluation génétique à partir d'un caractère de production et de comprendre la nature des VG prédites dans cette situation. Pour atteindre ce but, nous avons réalisé une étude par simulation reflétant un élevage de sélection porcine avec une structuration de la population en case-intra-bande. Le phénotype longitudinal de production simulé reflète le poids des porcs en engraissement que nous avons considéré dans 2 situations différentes : en présence ou absence de perturbations. Afin de coller au mieux à la réalité, nous nous sommes inspirés des données de la station de phénotypage du Rheu pour calibrer notre modèle de simulation. Les phénotypes d'intérêt analysés ont été le gain moyen quotidien (GMQ) entre l'entrée et la fin d'engraissement et le poids à 100j d'âge et non les données longitudinales de poids afin de correspondre une fois de plus à ce qui est réalisé actuellement en élevage de sélection. L'impact des perturbations sur la sélection a été évaluée en analysant la différence entre les VG prédites en situations perturbées et non perturbées.

3.2 Article

Animal 16 (2022) 100496

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences



Impact of environmental disturbances on estimated genetic parameters and breeding values for growth traits in pigs

Vincent Le [a,b,*], Tom Rohmer [a], Ingrid David [a]

[a] GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326 Castanet Tolosan, France
[b] Alliance R&D, 35650 Le Rheu, France

Abstract

Due to the diversification of farming systems and climate change, farm animals are exposed to environmental disturbances to which they respond differently depending on their robustness. Disturbances such as heat stress or sanitary challenges (not always recorded, especially when they are of short duration and low intensity) have a transitory impact on animals, resulting in changes in phenotypes of production (feed intake, body weight…). The aim of this study was to evaluate the impact of such unknown disturbances on the estimated genetic parameters and breeding values for production traits. A population of 6 120 individuals over 5 generations divided into 8 batches of 10 pens was generated, each individual underwent an ≃100-day test period. A longitudinal phenotype mimicking piglet weight during the fattening period was simulated for each individual in two situations: disturbed and non-disturbed. The disturbed phenotype was modified according to the robustness of the animal and the intensity and duration of the disturbance that the animal was subjected to. Various sets of simulations (1000 replicates per set) were considered depending on the type of disturbance (at the level of the batch, pen, or individual), the genetic correlation (negative, neutral, or positive) between the two components of the robustness (resistance and resilience), the genetic correlation (negative, neutral, or positive) between growth and the components of robustness, and the heritability of the components of robustness (weak or moderate). An animal model was used to estimate the genetic parameters and **breeding values (BV)** for two production traits: the **BW at 100 days of age (BW_{100})** and **average daily gain (ADG)**. The estimated heritability of the production traits was lower in the disturbed situation compared to the non-disturbed one (reduction of 0,08 and 0,05 points respectively for BW_{100} and ADG). The correlations between **estimated breeding values of the observed phenotypes (EBV)** and BV for production traits in absence of disturbance were lower in the disturbed situation (reduction of 0,04 and 0,06 points for BW_{100} and ADG respectively) while the partial correlation between EBV and BV for robustness was not significantly different from 0 in the two situations. These results suggest that selection in a well-controlled environment with random disturbances of low intensities does not allow to improve animal robustness while it is less effective for improving production traits than selection under no environmental disturbances.

Keywords: Simulation, Longitudinal performance, Robustness, Resistance, Resilience.

Implications

The impact of low-intensity and random disturbances, encountered in nucleus farms, on estimated genetic parameters and breeding values for production traits is unknow. This study showed that, for production traits recorded at a given time, this impact is noticeable but is not sufficient to concomitantly select for animal robustness and production. To reach this last objective, it would be interesting to study the dynamics of the phenotype over time to decipher production from robustness, which is made possible by the development of automatic devices in farms.

Introduction

Conventional animal breeding conditions are becoming more diversified, in line with the reduction of pharmaceutical inputs, diversification of feed resources, and reduction of energy costs of livestock farms (Rauw and Gomez-Raya, 2015). All these measures are due to societal and farmer demands (aims to improve animal welfare and minimize costs). In addition, environmental conditions are also subject to greater variability due to climate change (heat waves) and new regulations regarding rearing conditions (abolition of individual cages) (Star et al., 2008). These changes generate a range of suboptimal production environments, to which animals respond differently depending on their robustness (Knap, 2005). This robustness can be described in terms of resistance and resilience (Nguyen-Ba et al., 2020), which are respectively defined as the ability of an animal to minimize the direct impact of the disturbance (de Goede et al., 2013) and its ability to quickly return to the state before the disturbance (Colditz et al., 2016). When the disturbances that the animals are subject to are known and recorded (temperature, disease, feed, farm management changes, etc.), it is possible to take them into account in the genetic evaluation model in order to obtain unbiased and accurate estimates of genetic parameters and **breeding values (BV)** for the trait of interest according to the environment. Conventional models that account for genotype by environment interaction used in such cases are the multiple-trait (Meyer, 2009) and the reaction norm models (Kolmodin et al., 2002). When the disturbances are unknown, a reaction norm model with unknown covariates has been proposed, although it can suffer from identifiability issues (i.e. no unique solution) (Shariati et al., 2009). The effect of unknown disturbances is generally considered in the evaluation model by the inclusion of a contemporary group effect (most often the herd-year-season effect in ruminants (Van Vleck, 1987) and the pen-within-batch effect in monogastric animals (Frey et al., 1997)). This model assumes that all animals in the same contemporary group are exposed to the same disturbances and that there is no genotype by environment interaction (i.e., response to change in the environment is the same for all genotypes, meaning that all animals have the same robustness) (Phocas et al., 2016). In nucleus farms where animals are raised in a highly controlled environment, disturbances are expected to be of low magnitude. It is however unclear what is the impact of these disturbances on the estimates of the genetic parameters and BV in this situation especially when the phenotypes of interest are measured at a given

time point and thus not necessarily recorded concomitantly with the exposure of disturbances. The objective of this study was to evaluate the impact of random low-intensity disturbances that animals can face during a given period of time on genetic parameters' estimates and **estimated breeding values of the observed phenotypes (EBV)**. Simulations of the BW of fattening pigs was used to meet this objective. Impact of disturbances on genetic evaluation for **BW at 100 days of age (BW_{100})** and for **average daily gain (ADG)** over the fattening period was evaluated.

Material and Methods

The aim of the simulations was to reproduce the effects of disturbances on pig's growth that may occur in a pig production system. They comprise the simulation of i) a simplified population of non-overlapping generations reared in different pens and batches ii) disturbances of different duration, intensity and starting date during the fattening period that occur at the individual, pen or batch level iii) the resistance and resilience (unobserved traits) of each individual iv) the observed trait which is the daily BW of each animal during the fattening period ignoring the effect of the disturbance (non-disturbed phenotype) and including the effect of the disturbance (disturbed phenotype). In the latter case, the dynamic of the phenotype over time is modified according to the disturbance's intensity and duration; as well as the animal's resistance and resilience.

2.1 - Population simulation

The simulated oversimplified pig population consisted of 5 non-overlapping generations without selection. The founder generation comprised 12 sires and 150 dams that were randomly mated to give birth to 1 200 phenotyped offspring (8 offspring per dam, 4 males and 4 females). Among the progeny, 12 males and 150 females were sampled randomly to be the parents of the next generation. The same process was repeated from the **first generation (G1)** to the **last generation (G5)**. The final population without the founders comprised $N = 6\,000$ individuals. To mimic a pig production system whereby animals are raised in small groups during the fattening period; animals of G1 to G5 were, within each generation, randomly distributed across batches and across pens within a batch. Eight batches and 10 pens per batch were considered, leading to 15 individuals in each pen for each generation.

2.2 - Disturbance simulation

Three different types of disturbances were simulated: batch disturbances (all animals in the same batch are subjected to the same disturbance), pen (all animals in the same pen are subjected to the same disturbance), and individual (the disturbance acts on a single animal). All batches were subject to a batch disturbance, while pen disturbances and individual disturbances occurred in a certain proportion of pens and animals. The pens and animals affected by these disturbances were randomly sampled. Two proportions of pen and individual disturbances were considered (20 and 40%). They were chosen based on input from experts in the field (Canario, personal communication). The intensity, starting point, duration of a given disturbance were randomly sampled according to the distributions provided in Table 1. The intensities were considered to be constant over time. A given batch could suffer from one batch disturbance only (same for pen and individual level: a given pen could suffer from one pen disturbance only…), but disturbances at different levels were not exclusive. Thus, an animal could be exposed throughout its life to one batch disturbance, one pen disturbance, and one individual disturbance that can occur at the same time.

2.3 - Resistance and resilience simulation

The resistance ($\boldsymbol{resis_i}$) and **resilience** ($\boldsymbol{resil_i}$) of each animal were considered on a [0,1] scale. The resistance corresponds to the ability of an animal to minimize the direct impact of the disturbance on growth performance. A resistance value of 0 corresponds to an absence of resistance, the direct impact of the disturbance will be maximal. Conversely, a resistance value of 1 indicates an animal insensible to disturbance, the perceived intensity will be null and the disturbance will have no impact on growth. The resilience corresponds to the ability to quickly return to the state before the disturbance. A resilience value of 0 indicates absence of resilience, the animal will remain in the state it was at the end of the disturbance (i.e. constant difference with the value of the phenotype it should have had if there had been no disturbance), while a resilience value of 1 indicates strong resilience (the growth of the animal will be maximal in order to quickly return to the BW it should have had if there had been no disturbance). Resistance and resilience were simulated, for each animal, $i = 1, \dots, N$, according to the following model:

$$\begin{aligned} logit(resis_i) &= \boldsymbol{x_{r_i}}\boldsymbol{\beta_r} + u_{r_i} + e_{r_i} \\ logit(resil_i) &= \boldsymbol{x_{R_i}}\boldsymbol{\beta_R} + u_{R_i} + e_{R_i}, \end{aligned} \quad [1]$$

where $\boldsymbol{\beta_r}$ and $\boldsymbol{\beta_R}$ are the vectors of fixed effects (chosen arbitrarily, see Table 1) affecting resistance and resilience (two levels for each), $\boldsymbol{x_{r_i}}\ and\ \boldsymbol{x}_{R_i}$ the corresponding incidence row vectors, u_{r_i} and u_{R_i} are the resistance and resilience additive genetic effects (distribution given later in the text), and e_{r_i} and e_{R_i} are the residuals for the resistance and resilience of animal i, with joint distribution $\begin{bmatrix}\boldsymbol{e_r}\\ \boldsymbol{e_R}\end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \begin{bmatrix}\sigma^2_{e_r} & 0\\ 0 & \sigma^2_{e_R}\end{bmatrix} \otimes \boldsymbol{I_N})$, with $\boldsymbol{I_N}$ the identity matrix of size N, $\boldsymbol{e_r} = (e_{r_i})_i$ and $\boldsymbol{e_R} = (e_{R_i})_i$, and $\otimes$ the Kronecker product. Note that the resistance and the resilience in Equation [1] are considered to be constant over time for each animal.

2.4 - Simulation of longitudinal phenotypes

We considered that each animal i enters the fattening period at an animal-specific age that ends at 100 days of age, thus leading to a test period of length n_i that varies from 93 to 100 days (the enter-age is unique to each animal regardless of animal group, see Table 1). During the fattening period, a longitudinal phenotype that mimics the daily BW was simulated for each individual of generation G1 to G5 in the following two situations: without disturbance versus disturbances possible, indexed with wd and d respectively. For individual $i = 1, \dots, N$ and day j ($j \in \{1, \dots, n_i\}$), the respective longitudinal phenotypes are denoted $\boldsymbol{y_{wd,i}} = (y_{wd,ij})_{j=1,\dots,n_i}$ and $\boldsymbol{y_{d,i}} = (y_{d,ij})_{j=1,\dots,n_i}$.

The phenotype in the non-disturbed situation was simulated by a random regression model (Schaeffer, 2004) using second-order orthogonal polynomials :

$$y_{wd,ij} = \mu_j + enter_age_i + \boldsymbol{x}_{ij}\boldsymbol{\beta} + \sum_{k=0}^{2} a_{k,i}\,\varphi_k\left(age_{ij}\right) + e_{ij}, \qquad [2]$$

where $\mu_j = 15 + j$ represents the population mean at day j ($j \in \{1, \dots, 100\}$) that increased linearly with time. The entity $enter_age_i$ is the enter-age of animal i. The vector $\boldsymbol{\beta} = [\beta_1\ \beta_2\ \beta_3]^T$, with incidence row vector $\boldsymbol{x}_{ij}$, is the vector of fixed effects arbitrarily chosen to reflect the effect of one cross-classified factor with three levels (the daily social status of the animal for instance: aggressor, victim and none). The level of the factor was randomly sampled for each individual and each day j from a discrete uniform distribution $U[\{1,2,3\}]$ (i.e, if 2 is sampled then $\boldsymbol{x}_{ij} = [0,1,0]$) (see Table 1). The entity age_{ij} denotes the age of animal i at day j, that is to say: $age_{ij} = enter_age_i + j$. Functions φ_k are the Legendre

orthogonal polynomials of degree k (Robson, 1959) and the coefficients $a_{k,i}$ are the random additive genetic coefficients for animal i.

The joint distribution of the random coefficients $\boldsymbol{a_k} = (a_{k,i})_i$, appearing in Eq [2] and of the random effects $\boldsymbol{u_r} = (u_{r_i})_i$ and $\boldsymbol{u_R} = (u_{R_i})_i$ of Eq [1] was the following:

$$\begin{bmatrix} \boldsymbol{a_0} \\ \boldsymbol{a_1} \\ \boldsymbol{a_2} \\ \boldsymbol{u_r} \\ \boldsymbol{u_R} \end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}), \text{with } \mathbf{G} = \begin{bmatrix} \sigma^2_{a_0} & \sigma_{a_0 a_1} & \sigma_{a_0 a_2} & \sigma_{a_0 u_r} & \sigma_{a_0 u_R} \\ & \sigma^2_{a_1} & \sigma_{a_1 a_2} & \sigma_{a_1 u_r} & \sigma_{a_1 u_R} \\ & & \sigma^2_{a_2} & \sigma_{a_2 u_r} & \sigma_{a_2 u_R} \\ & sym & & \sigma^2_{u_r} & \sigma_{u_r u_R} \\ & & & & \sigma^2_{u_R} \end{bmatrix},$$

where $\mathbf{A}$ is the genetic relationship matrix. Residuals e_{ij} were independent and had centered Gaussian distributions with variances at day j equal to $11\,\varphi_0(j)^2 + 8\varphi_1(j)^2 + 0{,}6\varphi_2(j)^2$. Variances and covariances of the symmetric $\mathbf{G}$ matrix and formula for the variance of the residuals were chosen in order to obtain moderate heritabilities of the undisturbed BW over time (from 0,20 to 0,35). A detailed description of the parameters/distributions used in the simulation program is provided in **Table 1**.

Table 1 - Distribution of the random variables and the values of the parameters used in the simulations of the BW of pigs.

Parameter / variable	Description	Distribution/value
$enter_age_i$	Age in day when animal i enters the test period	$U(\{0,1,\dots,7\})$
μ_j	Population mean for y_{wd} at day j	$\mu_j = 15 + j$
$\boldsymbol{\beta}$	Vector of fixed effects for y_{wd} and y_d, one fixed effect with three levels	$\boldsymbol{\beta} = [1,2,3]^T$
$\boldsymbol{x_{ij}}$	Vector linking the fixed effect to the observed phenotype for animal i at day j	[1,0,0], [0,1,0] or [0,0,1] randomly sampled
$\boldsymbol{\beta_r}$	Vector of fixed effects for resistance, one fixed effect with two levels	$\boldsymbol{\beta_r} = [1,-1]^T$
$\boldsymbol{x_{ri}}$	Vector linking the fixed effect to the resistance for animal i	[0,1] or [1,0] randomly sampled
$\boldsymbol{\beta_R}$	Vector of fixed effects for resilience, one fixed effect with two levels	$\boldsymbol{\beta_R} = [1,-1]^T$
$\boldsymbol{x_{Ri}}$	Vector linking the fixed effect to the resilience for animal i	[0,1] or [1,0] randomly sampled
int	Intensity of the disturbance at the batch level	$\Gamma(0,3,1)$
	Intensity of the disturbance at the pen or individual level	$U([0,1,2])$
$start$	Start time of a disturbance	$U(\{2,3,\dots,100\})$
dur	Duration of a disturbance	$U(\{1,2,\dots,25\})$
$\sigma^2_{a_0}$	Variance of the genetic regression coefficient a_0	4
$\sigma^2_{a_1}$	Variance of the genetic regression coefficient a_1	2,4
$\sigma^2_{a_2}$	Variance of the genetic regression coefficient a_2	0,8
$\sigma_{a_0 a_1}$	Covariance between a_0 and a_1	0,62
$\sigma_{a_0 a_2}$	Covariance between a_0 and a_2	-0,36
$\sigma_{a_1 a_2}$	Covariance between a_1 and a_2	-0,28
$\sigma^2_{e,j}$	Residual variance at day j	$11\varphi_0(j)^2 + 8\varphi_1(j)^2 + 0{,}6\varphi_2(j)^2$
$\sigma^2_{u_r} + \sigma^2_{er}$	Sum of genetic and residual variance of resistance	1
$\sigma^2_{u_R} + \sigma^2_{eR}$	Sum of genetic and residual variance of resilience	1
$h^2\ ADG$	Heritability of the average daily gain	0,29
$h^2\ BW_{100}$	Heritability of the BW at 100 days	0,26

$U()$: uniform distribution, $\Gamma()$: gamma distribution

Resistance and resilience were simulated, for each animal , $i = 1, \dots, N$, according to equation [1]:

$$logit(resis_i) = \boldsymbol{x_{r_i}}\boldsymbol{\beta_r} + u_{r_i} + e_{r_i}$$
$$logit(resil_i) = \boldsymbol{x_{R_i}}\boldsymbol{\beta_R} + u_{R_i} + e_{R_i}$$

The phenotype in the non-disturbed situation was simulated according to equation [2]:

$$y_{wd,ij} = \mu_j + enter_age_i + \boldsymbol{x_{ij}}\boldsymbol{\beta} + \sum_{k=0}^{2} a_{k,i}\,\varphi_k(age_{ij}) + e_{ij}$$

The BW in the disturbed situation was modeled by considering the dynamic nature of the response to a perturbation using a dynamic model. Let's consider the following notations: $y^*_{wd,ij}$ and $y^*_{d,ij}$ are the non-disturbed (wd = without disturbance) and the disturbed BW of animal i at day j corrected for fixed effects (i.e., $y^*_{wd,ij} = y_{wd,ij} - \boldsymbol{x_{ij}\beta}$ and $y^*_{d,ij} = y_{d,ij} - \boldsymbol{x_{ij}\beta}$), $\Delta_{d,ij} = y^*_{d,ij} - y^*_{d,i(j-1)}$ and $\Delta_{wd,i} = y^*_{wd,ij} - y^*_{wd,i(j-1)}$, K_{ij} is the number of disturbances that animal i was subjected to on day j ($K_{ij} \in [\![0, 3]\!]$), and int_k, $k = 0, .., K_{ij}$ are their intensities ($int_0 = 0$), $int_{ij} = \sum_{k=0}^{K_{ij}} int_k$ (i.e., the sum of the intensities of the disturbances that animal i was subjected to on day j). At the first time point, $y_{d,i1} = y_{wd,i1}$ for all animals. Then, for $j > 1$, the disturbed phenotype is recursively defined as:

$$\frac{\Delta_{d,ij}}{\Delta_{wd,ij}} = 1 + resil_i \frac{y^*_{wd,ij}}{y^*_{wd,i1}} \left(1 - \frac{y^*_{d,i(j-1)}}{y^*_{wd,i(j-1)}}\right) - (1 - resis_i) int_{ij}, \quad [3]$$

The right side of the Equation [3] can be split into two terms $A_{ij} = 1 + resil_i \frac{y^*_{wd,ij}}{y^*_{wd,i1}} \left(1 - \frac{y^*_{d,i(j-1)}}{y^*_{wd,i(j-1)}}\right)$ and $B_{ij} = (1 - resis_i)\ int_{ij}$ that depict the response of an animal that is subject to a disturbance. Specifically, A_{ij} takes into consideration the effect of the resilience of animal i and B_{ij} the effect of the resistance of animal i weighted by the total intensity of disturbances int_{ij} experienced by animal i at day j. Indeed, throughout the life of the animal, three situations can arise:

1/ Before the first disturbance occurs, $y_{d,i(j-1)} = y_{wd,i(j-1)}$ and $int_{ij} = 0$, thus $A_{ij} = 1$; and $B_{ij} = 0$ so $\Delta_{d,ij} = \Delta_{wd,ij}$ and consequently $y_{d,ij} = y_{wd,ij}$.

2/ Then, when the animal is subjected to disturbance(s); during the disturbed period $int_{ij} > 0$, thus B_{ij} is greater than 0. The disturbance is hence considered to have a negative impact on the phenotype proportional to the intensity of the disturbance(s) that is moderated by the resistance of the animal. At the same time, as soon as the disturbed phenotype is lower than the undisturbed one, a resilience mechanism is involved (A_{ij}>1) to also limit the effect of the perturbation.

3/ Once the disturbance(s) is/are over, only the resilience mechanism remains ($B_{ij} = 0$ and A_{ij}>1) until the disturbed phenotype reaches the non-disturbed one.

This model to account the effect of disturbances (Eq [3]) was inspired from the one proposed by Nguyen Ba *et al*. (2020): $\frac{\Delta_{d,ij}}{\Delta_{wd,ij}} = 1 + resil_{NGi}\left(1 - \frac{y^*_{d,i(j-1)}}{y^*_{wd,i(j-1)}}\right) - resis_{NGi}$, where $resil_{NGi}$ and $resis_{NGi}$ refer to the resistance and resilience, respectively, of animal i as defined by Nguyen Ba *et al*. (2020).

From the longitudinal non-disturbed and disturbed phenotypes, we extracted two synthetic phenotypes that are measurable in practice in all farms and that were considered of interest for selection: BW_{100} and ADG. The impacts of disturbances on genetic parameter estimates and EBV were evaluated for these two synthetic phenotypes of interest. In the non-disturbed situation, only genetic effects affecting growth are involved in the observed phenotype while in the disturbed situation, the observed phenotype may be the result of genetic effects for growth but also for the resistance and resilience.

2.5 - Sets of simulations

Various sets of simulations were considered. These different alternatives were obtained by considering:

- Different proportions of pen and individual disturbances (only pen disturbance, only individual disturbance, or both disturbance types),
- Different correlations (negative, neutral, or positive) between the production traits and the two components of the robustness (resistance and resilience),
- Different correlations (negative, neutral, or positive) between resistance and resilience,
- Different heritabilities for resistance and resilience (moderate or low).

The various situations listed above are summarized in **Table 2**, and they led to 54 different alternatives. For each alternative, 1 000 independent datasets were generated by Monte Carlo simulations. For simplicity, in the rest of the document, the term “robustness” will refer to resistance and resilience. Then, a positive correlation between production and robustness means a positive correlation between production and resistance and a positive correlation between production and resilience. A moderate heritability of robustness corresponds to

situation with moderate heritability for resistance and moderate heritability for resilience and so on.

Table 2 - Description of the various alternatives of the simulation of the BW of pigs.

Proportion of pen (P_{pen}) and individual (P_{ind}) disturbances	$P_{pen} = 40\%, P_{ind} = 0\%$ $P_{pen} = 0\%, P_{ind} = 40\%$ $P_{pen} = 20\%, P_{ind} = 20\%$
Correlation ρ between phenotype of interest and resistance or resilience[1]	
Neutral ($\rho = 0$)	$\sigma_{a_0 u_r} = 0,\quad \sigma_{a_0 u_R} = 0,$ $\sigma_{a_1 u_r} = 0,\quad \sigma_{a_1 u_R} = 0,$ $\sigma_{a_2 u_r} = 0,\quad \sigma_{a_2 u_R} = 0$
Positive ($\rho \simeq 0{,}4$)	$\rho_{a_0 u_r} = 0{,}4,\quad \rho_{a_0 u_R} = 0{,}4,$ $\rho_{a_1 u_r} = 0{,}6,\quad \rho_{a_1 u_R} = 0{,}6,$ $\rho_{a_2 u_r} = -0{,}4,\quad \rho_{a_2 u_R} = -0{,}4$
Negative ($\rho \simeq -0{,}4$)	$\rho_{a_0 u_r} = -0{,}4,\quad \rho_{a_0 u_R} = -0{,}4,$ $\rho_{a_1 u_r} = -0{,}6,\quad \rho_{a_1 u_R} = -0{,}6,$ $\rho_{a_2 u_r} = 0{,}4,\quad \rho_{a_2 u_R} = 0{,}4$
Correlation ($\rho = \frac{\sigma_{u_r u_R}}{\sigma_{u_r}\sigma_{u_R}}$) between resistance and resilience[1]	Neutral $\rho = 0$ Positive $\rho = 0{,}4$ Negative $\rho = -0{,}4$
Heritability of resistance and resilience[2]	Moderate: $h^2_{resis} = h^2_{resil} = 0{,}29$ Low: $h^2_{resis} = h^2_{resil} = 0{,}16$

Abbreviations: a_0, a_1, a_2 = genetic regression coefficients of the Legendre polynomials; u_r, u_R = breeding values for resistance and resilience; h^2_{resis}, h^2_{resil} = heritability of resistance (resis) and resilience (resil).

[1] When the correlation between resistance and resilience is negative, the sign of the correlation between the trait of interest and the resistance is the inverse of the sign of the correlation between the trait of interest and the resilience.

[2] Heritability on the underlying scale

Statistical analysis

An animal model was applied for each synthetic phenotype of interest in order to obtain the estimates of heritability and BV in the absence or presence of disturbances: $\mathbf{y_{obs}} = \mathbf{X}\boldsymbol{\beta}_{\boldsymbol{obs}} + \mathbf{Z}\mathbf{u_{obs}} + \mathbf{e_{obs}}$ where $\mathbf{y_{obs}}$ is the vector of ADG or BW_{100}, $\boldsymbol{\beta}_{\boldsymbol{obs}}$ the vector of fixed effects with incidence matrix $\mathbf{X}$, $\mathbf{u_{obs}}$ the genetic additive effects ($\mathbf{u_{obs}} \sim N(\mathbf{0}, \mathbf{A}\boldsymbol{\sigma}^2_{u_{obs}})$) with incidence matrix $\mathbf{Z}$, and e_{obs} the vector of independent residuals $\mathbf{e_{obs}} \sim N(\mathbf{0}, \mathbf{I}\boldsymbol{\sigma}^2_{e_{obs}})$. Fixed effects included in the model were the contemporary group effect (pen x batch interaction) for both traits and the cross-classified factor with three levels used for the simulation at the age of 100 days for BW_{100}. We then (*i*) compared the heritabilities obtained in the disturbed and the non-disturbed situations, (*ii*) calculated the Pearson correlations between **BV for the production trait (BV_p)** and EBV; the BV_p corresponds to the true genetic potential of the animal for growth in the absence of disturbances and it was computed, for individual $i = 1, \dots, N$, as $\sum_{k=0}^{2} a_{k,i}\, \varphi_k(100)$ for BW_{100} and $\sum_{k=0}^{2}(a_{k,i}\, \varphi_k(100) - a_{k,i}\varphi_k(1))/99$ for ADG, (iii) calculated the partial correlation between EBV and BV for resistance given BV_p and BV for resilience and the partial correlation between EBV and BV for resilience given BV_p and BV for resistance, (*iv*) compared **the percentage of animals in common among the best 10% of animals (PB)** based on their EBV or BV_p and the PB based on their EBV obtained in the disturbed or the non-disturbed situation. In addition, (*v*) considering that 10% of the population was used as the parent of the next generation, for the last generation of each simulated data set (G5) we compared the mean of the true BV_p in the disturbed and the non-disturbed case for the selected animals (i.e., the best 10% of animals based on their EBV). One-sided paired Z-tests were used at the α risk of 5% for the comparisons of criteria obtained in the disturbed and the non-disturbed situation, and t-tests were used to compare the average difference of the various criteria (without disturbance-with disturbance) between the various simulated alternatives.

Results

To illustrate the simulated response of animals to a disturbance, the median BW over time of groups of animals undergoing three different situations are depicted in **Fig.1**: the group of animals does not face disturbance, the group of animals undergoes a batch disturbance of 'moderate' intensity (1,81) from day 20 to 35 and the group of animals undergoes a batch disturbance of 'weak' intensity (0,52) from day 61 to 82. When there is a disturbance, the maximal difference between the disturbed and the non-disturbed phenotype was observed on the last day of the perturbation (it was 27,54% on day 35 for the 'moderate' case and 6,37% on day 82 for the 'weak' case). Once the perturbation was over, due to the resilience of the animal, ADG was greater for disturbed animals than ADG of undisturbed animals so that their BW could return to their non-disturbed curve. At the population level, with the parameters used in this study to simulate phenotypes, on average, the impact of the disturbance led to a reduction of the slope of the curve by 40% during the disturbed period. Consequently, at the end of the disturbed period, the value of the disturbed phenotype was, on average, 90% of the non-disturbed one (increasing with the age of the animal), which represents a moderate impact on the phenotype. In all of the simulations sets, the correlation between the disturbed and the non-disturbed phenotype was high for ADG and BW_{100} (0,89 ± 0,01 and 0,81 ± 0,02, respectively).

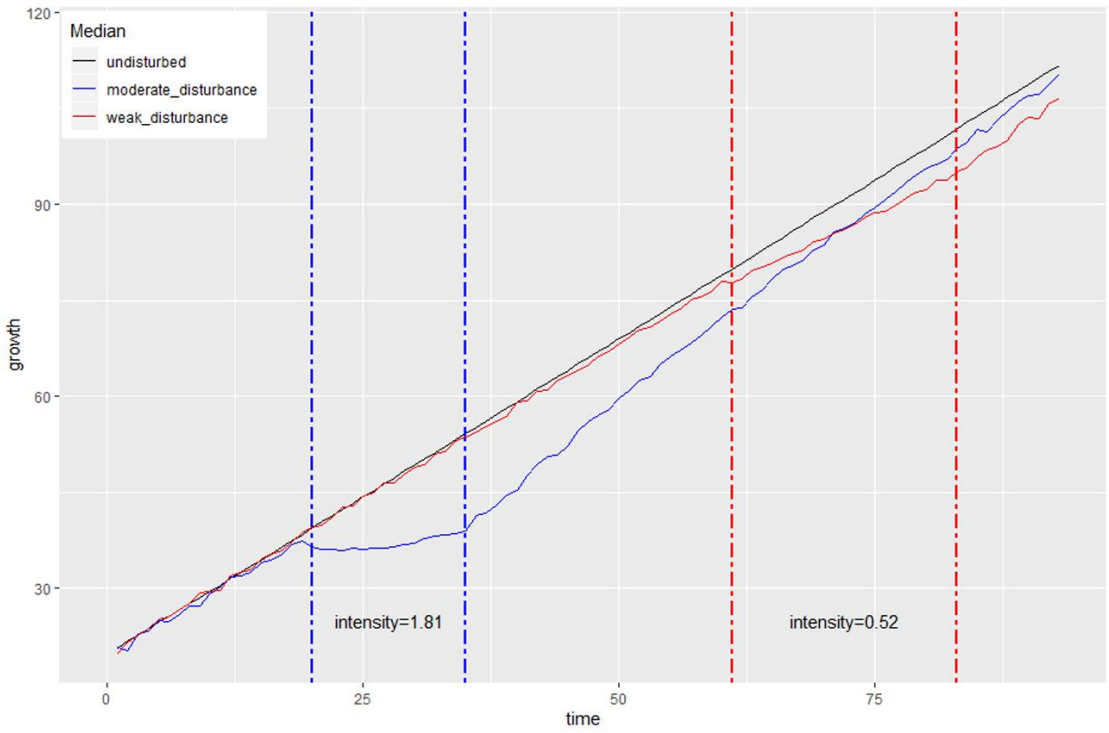


Fig.1 - Illustration of the effect of disturbances on the growth curve of pigs.

In black: median BW of the population that is not disturbed.

In blue: median BW of a group of animals undergoing a batch disturbance of "moderate" intensity (1,81) from day 20 to day 35 (blue dotted lines).

In red: median BW of a group of animals undergoing a batch disturbance of "low" intensity (0,52) from day 61 to day 82 (red dotted lines).

All statistic criteria are presented in **Table 3** for the disturbed and the non-disturbed situations in order to understand the impact of disturbances on the estimated genetic parameters and BV. The estimated heritability obtained in the disturbed situation was significantly lower than that obtained in the non-disturbed situation for both traits. The average difference (without disturbance-with disturbance) between heritabilities was $0{,}08 \pm 0{,}02$, and $0{,}05 \pm 0{,}02$ for the BW_{100} and ADG, respectively. The Pearson correlations between BV_p and EBV were significantly weaker for the disturbed phenotypes than for the non-disturbed ones for the two phenotypes of interest: the average differences being $0{,}04 \pm 0{,}01$, and $0{,}07 \pm 0{,}02$ for the BW_{100} and ADG, respectively. The partial correlation between EBV and BV for resistance or BV for resilience did not differ significantly between disturbed and non-disturbed situations for either traits. The PB based on their EBV or BV_p was significantly higher in the non-disturbed situation compared to the disturbed one for the BW_{100}, the difference being $0{,}04 \pm 0{,}02$. For

both traits, we did not observe any significant differences between the disturbed and the non-disturbed situation for the mean BV_p of the animals selected in the last generation based on their EBV. The PB based on their EBV between the non-disturbed and the disturbed situation were $0{,}73$, and $0{,}81$ for the BW_{100} and ADG, respectively, and were significantly lower than one.

Table 3 - Impact of disturbances on estimated genetic parameters and breeding values for production traits in pigs.

Phenotype	Criterion[1]	Without disturbance[2]	With disturbance[2]	p-value paired Z-test
BW_{100}	h^2	0,29 ± 0,03	0,21 ± 0,03	< 0,01
	ρ_{EBV,BV_p}	0,69 ± 0,02	0,65 ± 0,03	< 0,01
	PB_{EBV,BV_p}	0,47 ± 0,03	0,43 ± 0,03	0,02
	μBV_p	3,07 ± 1,09	2,87 ± 1,11	0,22
	$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	0,00 ± 0,05	0,04 ± 0,06	0,28
	$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	0,00 ± 0,05	0,04 ± 0,06	0,24
ADG	h^2	0,26 ± 0,03	0,21 ± 0,03	< 0,01
	ρ_{EBV,BV_p}	0,71 ± 0,03	0,65 ± 0,04	< 0,01
	PB_{EBV,BV_p}	0,46 ± 0,03	0,44 ± 0,04	0,15
	μBV_p	0,04 ± 0,01	0,04 ± 0,01	0,42
	$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	0,00 ± 0,06	0,03 ± 0,07	0,20
	$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	0,00 ± 0,06	0,03 ± 0,07	0,19

Abbreviations: BW_{100} = weight at 100 days of age; ADG = average daily gain.

[1] h^2, ρ_{EBV,BV_p}, PB_{EBV,BV_p}, μBV_p, $\rho_{EBV,\boldsymbol{u}_r|BV_p,\boldsymbol{u}_R}$ and $\rho_{EBV,\boldsymbol{u}_R|BV_p,\boldsymbol{u}_r}$ are the heritability, the Pearson correlations between the simulated breeding values (BV) for the production trait (BV_p) and the estimated breeding values of the observed phenotypes (EBV), the percentage of animals in common among the best 10% of animals based on their EBV or BV_p, the mean of the simulated BV_p for the best 10% of animals of the last generation of the phenotypes of interest, the partial correlation between EBV and the simulated BV for resistance given BV_p and the simulated BV for resilience, and the partial correlation between EBV and the simulated BV for resilience given BV_p and the simulated BV for resistance.

[2] μ ± sd calculated on all simulation sets

The impact of the disturbance on the estimated heritability and the EBV varied according to the correlation between robustness and production (**Table 4**). The decreases in heritability, correlations between EBV and BV_p, and the PB based on EBV or BV_p between the non-disturbed and the disturbed situation were significantly higher when the genetic correlation between robustness and production was negative compared to when it was positive for both traits. More specifically, the heritability in the disturbed case was lower in the case of a negative correlation between robustness and production (0,19 and 0,20, for BW_{100} and ADG respectively) compared to the positive case (0,22 and 0,23, for BW_{100} and ADG respectively). The same trend was observed for the correlation between the BV_p and EBV (0,65 and 0,64 in the negative case, and 0,67 and 0,66 in the positive case, for BW_{100} and ADG respectively), and the PB based on their BV_p or EBV (0,43 and 0,43 in the negative case, and 0,44 and 0,45 in the positive case, for BW_{100} and ADG respectively). The partial correlation between EBV and BV for resistance or BV for resilience tended to increase with the correlation between robustness and production in the disturbed situation for BW_{100} (by 4 points between the positive and negative correlation). No significant differences in the change in the mean BV_p of animals selected in the last generation between the disturbed and the non-disturbed situation were observed depending on the correlation between robustness and production for both traits.

Table 4 - Average difference between the non-disturbed and the disturbed situation for the various criteria according to the correlation between robustness and production of pigs.

		Correlation between robustness and production		
Criterion[1]	Phenotype	*Negative*	*Neutral*	*Positive*
h^2	BW_{100}	0,09 ± 0,01 [a]	0,08 ± 0,01 [ab]	0,07 ± 0,01 [b]
	ADG	0,06 ± 0,00 [a]	0,05 ± 0,00 [b]	0,03 ± 0,00 [c]
ρ_{EBV,BV_p}	BW_{100}	0,04 ± 0,00 [a]	0,03 ± 0,00 [b]	0,02 ± 0,00 [c]
	ADG	0,07 ± 0,00 [a]	0,06 ± 0,00 [ab]	0,05 ± 0,00 [b]
PB_{EBV,BV_p}	BW_{100}	0,05 ± 0,01 [a]	0,04 ± 0,00 [ab]	0,03 ± 0,00 [b]
	ADG	0,03 ± 0,00 [a]	0,02 ± 0,00 [ab]	0,01 ± 0,00 [b]
μBV_p	BW_{100}	0,20 ± 0,04 [a]	0,17 ± 0,04 [a]	0,14 ± 0,04 [a]
	ADG	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	BW_{100}	-0,01 ± 0,01 [a]	-0,04 ± 0,01 [b]	-0,05 ± 0,01 [b]
	ADG	-0,02 ± 0,01 [a]	-0,03 ± 0,01 [a]	-0,02 ± 0,01 [a]
$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	BW_{100}	-0,02 ± 0,01 [a]	-0,04 ± 0,01 [b]	-0,06 ± 0,01 [b]
	ADG	-0,02 ± 0,01 [a]	-0,03 ± 0,01 [a]	-0,02 ± 0,01 [a]

Abbreviations: BW_{100} = weight at 100 days of age; ADG = average daily gain.

[1] h^2, ρ_{EBV,BV_p}, PB_{EBV,BV_p}, μBV_p, $\rho_{EBV,\boldsymbol{u}_r|BV_p,\boldsymbol{u}_R}$ and $\rho_{EBV,\boldsymbol{u}_R|BV_p,\boldsymbol{u}_r}$ are the heritability, the Pearson correlations between the simulated breeding values (BV) for the production trait (BV_p) and the estimated breeding values of the observed phenotypes (EBV), the percentage of animals in common among the best 10% of animals based on their EBV or BV_p, the mean of the simulated BV_p for the best 10% of animals of the last generation of the phenotypes of interest, the partial correlation between EBV and the simulated BV for resistance given BV_p and the simulated BV for resilience, and the partial correlation between EBV and the simulated BV for resilience given BV_p and the simulated BV for resistance.

μ ± sd of the average of the sets concerned, sets with a negative correlation between resistance and resilience were not taken into account for these comparisons.

[a,b] Values within a row with different superscripts differ significantly at $P < 0,05$.

The changes in the different criteria depending on the correlation between resistance and resilience, the type of disturbances, and the heritability of robustness are presented in the Supplementary Materials (Table S1, S2, **and** S3). For both traits, we did not observe any

significant differences or trend in the impact of disturbance for the various criteria depending on these sets of parameters.

Supplementary Table S1 - Average difference between the non-disturbed and the disturbed situation for the various criteria according to the correlation between resistance and resilience of pigs.

		Correlation between resistance and resilience		
Criterion[1]	Phenotype	*Negative*	*Neutral*	*Positive*
h^2	BW_{100}	0,08 ± 0,00 [a]	0,08 ± 0,00 [a]	0,08 ± 0,00 [a]
	ADG	0,05 ± 0,00 [a]	0,05 ± 0,00 [a]	0,05 ± 0,00 [a]
ρ_{EBV,BV_p}	BW_{100}	0,03 ± 0,00 [a]	0,03 ± 0,00 [a]	0,03 ± 0,00 [a]
	ADG	0,06 ± 0,00 [a]	0,06 ± 0,00 [a]	0,06 ± 0,00 [a]
PB_{EBV,BV_p}	BW_{100}	0,04 ± 0,00 [a]	0,04 ± 0,00 [a]	0,04 ± 0,00 [a]
	ADG	0,02 ± 0,00 [a]	0,02 ± 0,00 [a]	0,02 ± 0,00 [a]
μBV_p	BW_{100}	0,17 ± 0,03 [a]	0,17 ± 0,03 [a]	0,17 ± 0,03 [a]
	ADG	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	BW_{100}	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]	-0,03 ± 0,01 [a]
	ADG	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	BW_{100}	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]
	ADG	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]

Abbreviations: BW_{100} = weight at 100 days of age; ADG = average daily gain.

[1] h^2, ρ_{EBV,BV_p}, PB_{EBV,BV_p}, μBV_p, $\rho_{EBV,\boldsymbol{u}_r|BV_p,\boldsymbol{u}_R}$ and $\rho_{EBV,\boldsymbol{u}_R|BV_p,\boldsymbol{u}_r}$ are the heritability, the Pearson correlations between the simulated breeding values (BV) for the production trait (BV_p) and the estimated breeding values of the observed phenotypes (EBV), the percentage of animals in common among the best 10% of animals based on their EBV or BV_p, the mean of the simulated BV_p for the best 10% of animals of the last generation of the phenotypes of interest, the partial correlation between EBV and the simulated BV for resistance given BV_p and the simulated BV for resilience, and the partial correlation between EBV and the simulated BV for resilience given BV_p and the simulated BV for resistance.

μ ± sd of the average of the sets concerned

[a,b] Values within a row with different superscripts differ significantly at $P < 0,05$.

Supplementary Table S2 - Average difference between the non-disturbed and the disturbed situation for the various criteria according to the type of disturbance in a pig production system.

		Type of disturbance[2]		
Criterion[1]	Phenotype	$P_{pen} = 40\%$, $P_{ind} = 0\%$	$P_{pen} = 0\%$, $P_{ind} = 40\%$	$P_{pen} = 20\%$, $P_{ind} = 20\%$
h^2	BW_{100}	0,07 ± 0,01 [a]	0,09 ± 0,00 [a]	0,08 ± 0,00 [a]
	ADG	0,04 ± 0,00 [a]	0,05 ± 0,00 [a]	0,05 ± 0,00 [a]
ρ_{EBV,BV_p}	BW_{100}	0,03 ± 0,00 [a]	0,03 ± 0,00 [a]	0,03 ± 0,00 [a]
	ADG	0,06 ± 0,00 [a]	0,06 ± 0,00 [a]	0,06 ± 0,00 [a]
PB_{EBV,BV_p}	BW_{100}	0,04 ± 0,00 [a]	0,04 ± 0,00 [a]	0,04 ± 0,00 [a]
	ADG	0,02 ± 0,00 [a]	0,02 ± 0,00 [a]	0,02 ± 0,00 [a]
μBV_p	BW_{100}	0,16 ± 0,03 [a]	0,17 ± 0,03 [a]	0,17 ± 0,03 [a]
	ADG	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	BW_{100}	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]	-0,03 ± 0,01 [a]
	ADG	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	BW_{100}	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]	-0,04 ± 0,01 [a]
	ADG	-0,03 ± 0,00	-0,03 ± 0,00 [a]	-0,03 ± 0,00 [a]

Abbreviations: BW_{100} = weight at 100 days of age; ADG = average daily gain.

[1] h^2, ρ_{EBV,BV_p}, PB_{EBV,BV_p}, μBV_p, $\rho_{EBV,\boldsymbol{u}_r|BV_p,\boldsymbol{u}_R}$ and $\rho_{EBV,\boldsymbol{u}_R|BV_p,\boldsymbol{u}_r}$ are the heritability, the Pearson correlations between the simulated breeding values (BV) for the production trait (BV_p) and the estimated breeding values of the observed phenotypes (EBV), the percentage of animals in common among the best 10% of animals based on their EBV or BV_p, the mean of the simulated BV_p for the best 10% of animals of the last generation of the phenotypes of interest, the partial correlation between EBV and the simulated BV for resistance given BV_p and the simulated BV for resilience, and the partial correlation between EBV and the simulated BV for resilience given BV_p and the simulated BV for resistance.

μ ± sd of the average of the sets concerned

[2] P_{pen} is the percentage of pens subjected to a disturbance, P_{ind} is the percentage of individuals subjected to a disturbance; three different alternatives were considered (1 - 40% of pens disturbed, 2 – 40% of individuals disturbed, 3 – 20% of pens disturbed and 20% of individuals perturbed; and 100% of batches disturbed for all sets).

[a,b] Values within a row with different superscripts differ significantly at $P < 0,05$.

Supplementary Table S3 - Average difference between the non-disturbed and the disturbed situation for the various criteria according to the heritability of robustness of pigs.

Criterion[1]	Phenotype	Heritability of robustness	
		Moderate	*Weak*
h^2	BW_{100}	0,08 ± 0,00 [a]	0,08 ± 0,00 [a]
	ADG	0,05 ± 0,00 [a]	0,05 ± 0,00 [a]
ρ_{EBV,BV_p}	BW_{100}	0,03 ± 0,00 [a]	0,03 ± 0,00 [a]
	ADG	0,06 ± 0,00 [a]	0,06 ± 0,00 [a]
PB_{EBV,BV_p}	BW_{100}	0,04 ± 0,00 [a]	0,04 ± 0,00 [a]
	ADG	0,02 ± 0,00 [a]	0,02 ± 0,00 [a]
μBV_p	BW_{100}	0,17 ± 0,03 [a]	0,17 ± 0,03 [a]
	ADG	0,00 ± 0,00 [a]	0,00 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_r\|BV_p,\boldsymbol{u}_R}$	BW_{100}	-0,04 ± 0,01 [a]	-0,03 ± 0,01 [a]
	ADG	-0,03 ± 0,00 [a]	-0,02 ± 0,00 [a]
$\rho_{EBV,\boldsymbol{u}_R\|BV_p,\boldsymbol{u}_r}$	BW_{100}	-0,04 ± 0,01 [a]	-0,03 ± 0,01 [a]
	ADG	-0,03 ± 0,00 [a]	-0,02 ± 0,00 [a]

Abbreviations: BW_{100} = weight at 100 days of age; ADG = average daily gain.

[1] h^2, ρ_{EBV,BV_p}, PB_{EBV,BV_p}, μBV_p, $\rho_{EBV,\boldsymbol{u}_r|BV_p,\boldsymbol{u}_R}$ and $\rho_{EBV,\boldsymbol{u}_R|BV_p,\boldsymbol{u}_r}$ are the heritability, the Pearson correlations between the simulated breeding values (BV) for the production trait (BV_p) and the estimated breeding values of the observed phenotypes (EBV), the percentage of animals in common among the best 10% of animals based on their EBV or BV_p, the mean of the simulated BV_p for the best 10% of animals of the last generation of the phenotypes of interest, the partial correlation between EBV and the simulated BV for resistance given BV_p and the simulated BV for resilience, and the partial correlation between EBV and the simulated BV for resilience given BV_p and the simulated BV for resistance.

μ ± sd of the average of the sets concerned

[a,b] Values within a row with different superscripts differ significantly at $P < 0,05$.

Discussion

To model the effect of a disturbance on a longitudinal phenotype, the dynamic model used (Eq [3]) was inspired from the one proposed by Nguyen Ba *et al*. (2020). Both models consider that when a disturbance occurs, it will trigger both the mechanism of resistance and resilience of the animal at the same time. Once there is no longer a disturbance, as long as the ratio between the observed and theoretical curve is smaller than 1, the resilience of the animal will always have an impact on the evolution of the phenotype. It should be noted that both models assume that the negative impact of the disturbance on the phenotype induces a decrease in the value of the phenotype. The model proposed by Nguyen-Ba (2020) is limited in its number of parameters to estimate to prevent identifiability issues. Since the purpose of our model is to simulate data, additional parameters were included, thereby leading to more flexibility. The resistance and resilience of an animal have been used in the present study on a 0 to 1 scale and included in the model such that the higher the resistance or resilience, the lower the negative response to the perturbation. In Eq [3], the resilience of the animal was multiplied by $\frac{y^*_{wd,ij}}{y^*_{wd,i1}}$ to account for the change in the mean value of the phenotype over time. Indeed, at constant resilience, the capacity of an animal to reach the state before the manifestation of the disturbance is not the same at the beginning as at the end of the test period. It is widely known for example that, it will be harder for a pig to gain 1 kg of weight at the beginning than at the end of the test period. Furthermore, our model considers that the resistance mechanism involved in the response to a perturbation can vary when a disturbance occurs, as it is a time-specific function of both the intrinsic resistance potential of the animal and the intensity of the perturbation, while the resistance mechanism is supposed to be constant during the time window in which the perturbation occurs in the model proposed by Nguyen Ba *et al*. (2020). In our model, the value of the resistance and resilience were considered as stable throughout the life of the animal since the duration of the test period of each animal is relatively short (~100 days). Nevertheless, the model can be expanded to account for varying resistance and resilience that may be dependent on the animal's age or experience. Indeed, the resistance (or resilience) of the animal may decrease over time due to exhaustion if it is subject to repeated disturbances, or conversely, in the case of disease responses, it may increase through the development of acquired immunity. We modeled the robustness by the combination of resistance and resilience in order to depict the response of the animal for a

given phenotype over a short period of time. However, robustness is a complex concept and several other definitions have been proposed in the literature. For example, Friggens *et al.* (2017) defined the robustness of an animal by '*the ability, in the face of environmental constraints, to carry on doing the various things that the animal needs to do to favour its future ability to reproduce'*. Using this framework, modeling the robustness should consider multiple phenotypes throughout the life of the animal.

In our study, the ADG and BW_{100} were extracted from longitudinal BW for each individual. This choice was guided by the main (non-longitudinal) phenotypes under selection encountered in breeding (Martin et al., 2014; Tortereau et al., 2020). Nonetheless, the results obtained are certainly applicable to any other phenotypes measured at a given time point, as long as, following a disturbance (whatever the impact of the disturbance on the trait, decrease or increase), the animal can eventually recover its original trajectory. Even if the results were obtained in an oversimplified pig system, it is unlikely that the results would have been different when considering a more complex system (overlapping breeding scheme) since the perturbations essentially add noises to the phenotype measurement.

We have simulated three different types of disturbances that mimic those that can occur in practice. Disturbances at the batch level comprise, for example, a change in temperature, feeding regime, or group prophylaxis. The problem of cannibalism and moderately contagious lameness can be considered to affect close conspecifics and thus occurring at the pen level. Finally, metabolic and non-contagious lameness should affect single individuals. In all three sets of disturbance types, 100% of the batches were subjected to batch disturbance but with an intensity that can be very low. Indeed, it is reasonable to consider that the overall environmental conditions of an animal cannot be optimal during the entire test period (i.e., the temperature can change, the caretakers intervene…).

Different correlations between resistance (resilience) and production were investigated in the present study to cover the various possibilities that can occur. Indeed, there has been no consensus to date regarding the genetic relationship between these traits. For instance, Friggens *et al*. (2017) have deduced from the resource allocation theory that the correlation between the components of robustness and production should be negative. However, several

experiments have shown that it is not always true. The relationship between production and response of the animal to disturbance has been reported in studies evaluating the response of animals to sanitary challenges. By working on the response of pigs to porcine reproductive and respiratory syndrome, Hess *et al.* (2016) have demonstrated a negative correlation between growth and viremia level (which implies a favorable relationship between production and robustness). Whereas Heckendorm *et al*. (2017) have demonstrated the opposite by working on the resistance of goats to gastrointestinal nematodes. There is also no consensus regarding the heritability of measurements related to the robustness: Hess *et al*. (2016) showed a moderately high heritability for the viral load (between 0,31 and 0,51) by working on pigs' response to porcine reproductive and respiratory syndrome, while Mazé-Guilmo *et al*. (2014) demonstrated that the heritability for resistance (or tolerance) should be low (between 0,18 and 0,19) by working on *Leuciscus burdigalensis* (a freshwater fish) parasitized by *Tracheliastes polycolpus*.

The simulated disturbances were of low intensities in order to evaluate the practical impact of disturbances on actual selection. The results obtained show that, between the disturbed and non-disturbed case, there is a significant difference in the heritability, EBV, and in the selected candidates. The impact of the disturbance on the selection of a production trait is exacerbated in the case of a negative correlation between robustness and production compared to the positive case due to a higher reranking of animals (production of the best animals is more affected by a disturbance than the worst animals). Inclusion of a contemporary group effect in the model that corrects for common environmental factors in a group is not sufficient to account for disturbances that occur at the group (batch or pen) level (i.e., there was no difference between the type of perturbation sets on the various criteria) because, since this effect only corrects for the deviation of the mean performance of the group, it does not take into account the variability in the response of the animals to a disturbance due to their variable robustness. Thus, even in a well-controlled environment as observed in conventional breeding program, the selection for production traits can be impacted by disturbances. Considering that the EBV in such a situation corresponds to a combination of the robustness and production, the breeders could then select for these two terms in the same time. However, this study showed that the EBV in the disturbed situation is not a satisfactory combination of production and robustness: the partial correlation between

the EBV and the simulated BV of resistance (and resilience) did not differ from the non-disturbed and disturbed case. In addition, these values were not significantly different from zero. Therefore, we do not select animals for production potential and robustness at the same time in that case. The reason for this is probably because the animals are subjected to random disturbances of small magnitude and that may appear long before the measurement of the phenotype. Additional simulations, to be presented elsewhere, have been carried out to confirm these hypotheses. Because animals raised in production herds are in less controlled and more various environments (lower-quality feed sources, higher sanitary pressure, older facilities, conventional and organic herds…) compared to nucleus herd, it is important to select animals that will be able to adapt to these conditions, which implies an improvement of robustness while maintaining a high level of production. Thus, production and robustness have to be included in the breeding goal. Given our results, to simultaneously improve production and robustness, one should create stronger disturbed environment common to all animals in nucleus herds. However, this may not be ethically acceptable and too expensive for the breeders (Loïc Flatres-Grall, AXIOM company, personal communication). Thus, a more desirable alternative is to obtain separate EBV for growth and robustness. In order to do so, dealing with longitudinal data now available thanks to the development of electronic phenotyping tools should be useful. Such approach has already been investigated by several authors. For instance, Revilla *et al*. (2021) proposed to quantify the robustness of the animals by the deviation from the observed and theoretical curve of growth of piglets. However, reconstruction of the theoretical production curve is difficult when the time of the perturbation is unknown and further research is needed.

Conclusion

The results of this study show that random weak disturbances, typically encountered in a breeding farm, have an impact on the estimates of the genetic parameters and BV for production traits and on selection. However, conversely to expected, the EBV obtained in such situation do not correspond to a combination of production and robustness abilities that would be interesting to select. Thus, to select accordingly to a predefined breeding goal for production and robustness, it would be interesting to separate these two components by studying the dynamic of the phenotype over time instead of measuring the trait at a given time point. This subject is currently undergoing further research.

3.3 Discussion complémentaire de l'article 'Impact des perturbations environnementales sur l'estimation des paramètres et valeurs génétiques'

Les résultats de cette étude montrent que les VG prédites obtenues à partir du phénotype de production dans un environnement faiblement perturbé ne correspondent pas à une combinaison de la capacité de production et de robustesse. Nous avons postulé que ce résultat venait du fait que les perturbations n'apportaient que du bruit au phénotype observé et ne donnaient pas d'indication sur la robustesse parce que les perturbations simulées étaient peu fréquentes, de faible intensité et aléatoires. De plus, les mesures des phénotypes d'intérêt (poids à 100j) ou de leurs composantes (poids initial et final pour le GMQ) pouvaient être éloignées dans le temps du moment où se produisait la (ou les) perturbation(s). En effet, si la perturbation arrive trop tôt par rapport à ces phénotypes qu'on mesure en regardant le dernier poids, le porc a potentiellement déjà récupéré au moment où on mesure le phénotype. Pour confirmer ces hypothèses, quatre nouveaux cas de simulations supplémentaires ont été réalisés. En utilisant les paramètres de simulation qui considèrent une corrélation neutre entre la robustesse et la production, une corrélation neutre entre la résistance et la résilience, et une héritabilité modérée pour la robustesse, nous avons modifié la façon dont les perturbations ont été simulées (1000 simulations ont été effectuées pour chacun des 4 nouveaux cas) :

1) Tous les animaux reçoivent la même perturbation (intensité tirée uniformément entre 0,5 et 2,5, début de la perturbation tiré uniformément entre le jour 2 et 99).
2) Tous les animaux reçoivent la même perturbation d'intensité en moyenne plus élevée (intensité tirée uniformément entre 2,0 et 2,5, début de la perturbation tiré uniformément entre le jour 2 et 99).
3) Tous les animaux reçoivent la même perturbation d'intensité en moyenne plus élevée et de début tardif (intensité tirée uniformément entre 2,0 et 2,5, début tiré uniformément après le 75ème jour).
4) L'intensité de la perturbation de l'ensemble simulé précédent (cas 3) est encore augmentée (intensité tirée uniformément entre 3,0 et 4,5, début tiré uniformément après le 75ème jour).

Nous avons ensuite évalué les changements dans les corrélations partielles entre les VG prédites de la production et les vraies VG de résistance et résilience, pour ces simulations complémentaires afin d'évaluer si la VG prédite de la production inclut le potentiel de robustesse des animaux. La même tendance a été observée pour les 2 caractères (BW100 et ADG) (Table 5). La corrélation partielle avec la résistance a tendance à augmenter du cas 1 au cas 4 : de 0,10 à 0,35 pour BW100 et de 0,07 à 0,28 pour ADG, alors que la corrélation partielle avec la résilience a augmenté du cas 1 au cas 2 : pour BW100 (0,09, 0,12, respectivement) et pour ADG (0,06, 0,09, respectivement) ; et a ensuite diminué dans les cas 3 et 4 : pour BW100 (0,05, 0,06, respectivement) et pour ADG (0,04, 0,05, respectivement). Ces résultats sont en accord avec notre hypothèse : lorsque tous les animaux sont soumis aux mêmes perturbations, le phénotype mesuré reflète davantage le potentiel génétique de robustesse (résistance et résilience) des animaux. Cette tendance augmente avec la dureté de l'environnement (c.-à-d. avec l'intensité de la perturbation), prouvant que la réponse concernant la robustesse est plus visible dans des conditions environnementales difficiles. En outre, plus le début de la perturbation est proche de l'instant de la mesure du phénotype (par exemple proche de la mesure de BW100), plus la relation avec la résistance est forte, tandis que la relation avec la résilience diminue car le mécanisme de résilience n'a pas eu suffisamment de temps pour se mettre en place avant la mesure du phénotype. Ainsi, pour sélectionner simultanément les animaux en fonction de la production et la robustesse à partir du phénotype mesuré à un moment donné, il serait nécessaire de créer un environnement plus fortement perturbé (challenge sanitaire ou alimentaire) où les perturbations affectent tous les animaux et qui se produisent au même moment que le phénotypage. Cependant, cette méthode n'est peut-être pas éthiquement acceptable et trop coûteuse pour les éleveurs. Par conséquent, une alternative plus souhaitable pour avoir un meilleur contrôle de la robustesse et de la production de l'animal serait de pouvoir les dissocier (au moment du phénotypage ou de l'évaluation génétique). Pour ce faire, la prise en compte des aspects temporels de la robustesse par utilisation des données de phénotypage à haut débit (désormais disponibles grâce au développement des outils électroniques de phénotypage) pourrait être utile. En effet, en analysant la réponse au cours du temps des animaux face aux perturbations environnementales, un nouveau critère plus pertinent concernant la robustesse pourrait en effet être proposé. Cette approche fait l'objet du chapitre 4 de cette thèse.

Table 5 - Résultats sur l'évolution des corrélations partielles entre la VG prédite de la production et la VG simulée de la robustesse (résistance et résilience) selon différentes situations.

	Cas	Corrélation partielle entre VG prédite et vraie VG de la résistance	Corrélation partielle entre VG prédite et vraie VG de la résilience
BW100	0	0,03	0,04
	1	0,10	0,09
	2	0,14	0,12
	3	0,25	0,05
	4	0,35	0,06
ADG	0	0,03	0,03
	1	0,07	0,06
	2	0,10	0,09
	3	0,18	0,04
	4	0,28	0,05

Les cas 0 à 4 représentent respectivement les simulations où : 0 – les animaux simulés reçoivent des perturbations aléatoires de faible intensité ; 1 - tous les animaux simulés reçoivent la même perturbation ; 2 - tous les animaux simulés reçoivent la même perturbation d'intensité plus élevée (de 2,0 à 2,5) ; 3 - tous les animaux simulés reçoivent la même perturbation d'intensité plus élevée et de début tardif (après le 75e jour) ; 4 - l'intensité de la perturbation de l'ensemble simulé précédent est encore augmentée (de 3,0 à 4,5).

Chapitre 4 - Détection des perturbations et quantification de la robustesse

4.1 Introduction

Dans cette thèse, nous nous concentrons sur la quantification d'une robustesse spécifique des animaux d'élevage à partir de l'analyse de la réponse dynamique de l'animal face à des perturbations environnementales sur un de ses phénotypes longitudinaux dans le cas où aucune information sur les perturbations n'est disponible. Cela nécessite donc de développer une méthode capable de détecter et qualifier de manière précise les perturbations environnementales inconnues que les animaux ont rencontrées. Dans ce chapitre, nous présentons la méthode Up & Down développée pour atteindre cet objectif. Le cœur de ce travail est basé sur l'analyse de l'évolution dynamique des données longitudinales et sur la classification non supervisée. La méthode Up & Down est testée d'abord sur des données simulées (les mêmes que celles utilisées pour l'article 1 – présenté dans le chapitre précédent), puis appliquée sur les poids corrigés et les quantités cumulées d'ingéré des données des porcs en croissance de la station de phénotypage du Rheu, présentées dans le Chapitre 2 - Données d'étude.

4.2 Matériels et méthodes

La méthode Up & Down a été développée pour des phénotypes pour lesquels la dynamique de réponse face à une perturbation est de type « élastique » (voir chapitre 1, sous-section 1.2.2), le changement d'état se traduisant par une baisse (ou une augmentation) de l'évolution du phénotype durant la perturbation puis par une récupération pour rejoindre sa trajectoire théorique (trajectoire que l'animal aurait eu en l'absence de perturbation). Par simplicité, la méthode est présentée ici sous l'hypothèse d'une baisse d'évolution de la valeur du phénotype face à une perturbation environnementale (par exemple : le poids vif ou la consommation alimentaire en présence d'un pathogène) mais peut être facilement adaptée au cas de hausse du phénotype (par exemple : la température corporelle ou la consommation d'eau lors d'une canicule). La méthode Up & Down se divise en 2 grandes parties : la partie Up sert à identifier les éléments (de toutes les échelles : bande, case, individu) qui sont

potentiellement perturbés et la partie Down sert à valider les perturbations qui sont détectées dans la partie Up et indiquer les éléments ayant reçu plusieurs perturbations. Nous estimons ensuite les caractéristiques des perturbations validée (c.-à-d. déterminer le début, la fin et l'intensité de chaque perturbation). Enfin, une nouvelle mesure de robustesse est extraite des phénotypes des animaux identifiés comme ayant subi une perturbation.

4.2.1 Partie Up

Le but de la partie Up est de scinder les différentes échelles (individu, case, bande) de la population en 2 groupes : ceux potentiellement exposés à des perturbations et ceux qui ne le sont pas. L'hypothèse sous-jacente est que les animaux perturbés ont tendance à avoir une diminution de la vitesse d'évolution de leur phénotype durant la période de perturbation, contrairement aux animaux n'ayant subi aucune perturbation. Par conséquent, le critère choisi pour différencier les groupes « perturbés » et « non-perturbés » se fait en fonction du taux de décroissance du phénotype au cours du temps. Pour cette partie Up l'analyse démarre à l'échelle de l'individu, puis 'monte' à l'échelle de la case et enfin à l'échelle de la bande.

i. Afin de limiter l'impact de l'évolution naturelle du phénotype au cours du temps et des erreurs de mesure, la première étape consiste à corriger le phénotype observé de chacun des animaux par les effets fixes et par la médiane de la population par âge (remise au même niveau) et de lisser ce dernier (suppression des bruits). Un exemple de cette correction sur les données d'un phénotype simulé au cours du temps reprenant une courbe de poids vif d'un porc en croissance est présenté dans la Figure 16.

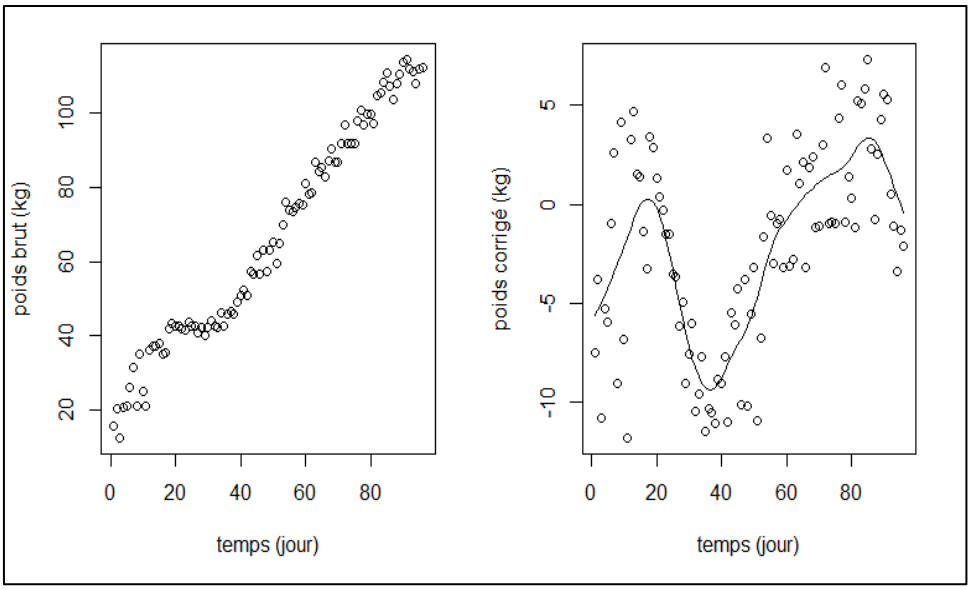


Figure 16 - Evolution du poids vif brut (graphique gauche), corrigé (par les effets fixes et la médiane de la population par âge) et lissé avec la fonction de lissage non-paramétrique de Nadaraya-Watson (graphique de droite).

ii. Dans la deuxième étape, le taux de décroissance maximale du phénotype corrigé et lissé pour chaque individu est extrait. Plus précisément, cette valeur correspond à la valeur minimale de la dérivée première (pente minimale) de la fonction de lissage appliquée dans l'étape précédente, elle représente le 'pire' impact d'une perturbation sur l'individu concerné.

iii. Finalement, en calibrant un modèle de mélange gaussien (*package R : mixtools* (Benaglia et al., 2010)) sur l'ensemble de ces valeurs, les animaux sont séparés en 2 groupes (perturbé et non perturbé) (Figure 17). La procédure d'ajustement du modèle de mélange fait appel à l'algorithme espérance-maximisation (EM) (McLachlan et al., 2019). Plus précisément, le modèle de mélange forme deux courbes gaussiennes différentes qui distinguent la population qui répond négativement à une (ou plusieurs) perturbation(s) et l'autre population qui ne répond pas (ou très peu) à ces perturbations. Le seuil τ qui sépare les deux populations correspond à l'intersection entre ces deux courbes, c'est-à-dire que tous les animaux ayant une pente minimale plus faible que ce seuil seront considérés comme perturbés et ceux au-dessus comme non perturbés. Ce seuil sera utile dans la deuxième partie de la méthode développée (plus de détail dans la partie 'Down'). En l'absence de convergence du modèle de

mélange (nombre maximal d'itérations choisi arbitrairement à 500), la population est considérée comme n'ayant reçu aucune perturbation à l'échelle individuelle.

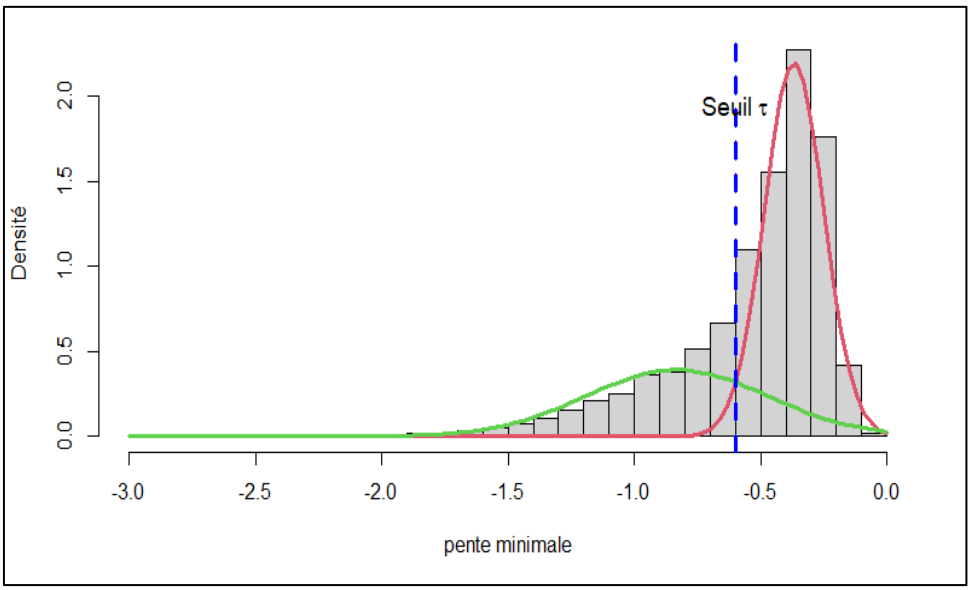


Figure 17 - Modèle de mélange appliqué sur les minimums de pente pour classer la population d'animaux en deux groupes : perturbé (sous la courbe verte) et non-perturbé (sous la courbe rouge) ; la ligne bleue pointillée correspond au seuil τ qui sépare les deux populations.

Le même travail (étape i à iii) est effectué ensuite à l'échelle de la case et enfin à l'échelle de la bande pour identifier les cases et les bandes ayant répondu à au moins une perturbation. Pour les cases, on étudie de la même manière la dynamique d'évolution au cours du temps de la médiane par case des phénotypes corrigés de chaque individu. Plus précisément, la médiane par temps des observations des individus-intra-case est d'abord calculée puis lissée en utilisant la même fonction de lissage que celle de l'échelle des individus. Cela élimine la plupart des perturbations individuelles. En effet, les perturbations individuelles n'influencent pas la médiane par case à un jour donné si à ce jour il y a moins de 50% des individus-intra-case subissant une perturbation individuelle. Pour les bandes, le travail est effectué sur la médiane de la médiane des cases-intra-bande (calculée à l'étape précédente) pour la même raison. Les médianes des jours où il y a moins de 50% des éléments phénotypés ne sont pas calculées car cela biaiserait les résultats finaux. En général, les porcs les plus lourds vont partir à l'abattoir plus tôt que ceux qui sont moins lourds, la médiane du groupe

d'animaux sera influencée en absence de ces animaux. Par conséquent, par exemple, si au jour 97, il y a moins de 50% des individus-intra-case phénotypés, la médiane de case pour ce jour n'est pas prise en compte. De même s'il y a moins de 50% des cases-intra-bande pour lesquelles la médiane est connue, la médiane de la bande ne sera pas calculée.

Dans cette partie Up, le niveau des perturbations détectées est remis à jour après chaque échelle. Par exemple, si un individu détecté comme ayant subi une perturbation à l'échelle de l'individu se retrouve dans une case détectée comme ayant subi une perturbation à l'échelle de la case, alors ce dernier sera considéré comme ayant eu une perturbation de case uniquement (et plus une perturbation individuelle). Le même raisonnement est effectué lorsque l'on passe de l'échelle de la case à l'échelle de la bande.

4.2.2 Partie Down

Une fois que les éléments perturbés (bande, case, individu) sont identifiés, il est nécessaire de les valider dans la partie Down car la partie Up peut probablement classifier de manière erronée certaines perturbations et à l'inverse ne pas en considérer certaines lorsqu'un animal a subi plusieurs perturbations. En effet, dans la partie Up, un animal ne peut avoir subi qu'une perturbation à une seule échelle par construction. De plus, s'il y a plus de la moitié des cases d'une bande qui ont reçu une perturbation de case différentes, la médiane de la médiane des cases-intra-bande pourra être affectée et la phase Up peut classer à tort cette bande comme perturbée (la même erreur peut se reproduire à l'échelle de la case).

La phase Down va donc reprendre les cases et bandes détectées comme perturbées dans la partie Up et va valider (ou invalider) les éléments alors détectés en identifiant les **moments de référence (MR)**. Les MR d'un élément (un élément peut être perturbé à plusieurs reprises) correspondent aux moments où le taux de décroissance du phénotype lissé est localement maximal et peuvent être captés en étudiant l'évolution au cours du temps du phénotype corrigé et lissé. Pour ce faire, tous les moments pour lesquels la pente du phénotype corrigé et lissé est localement minimale et est inférieure au seuil τ concerné (seuil calculé par le modèle de mélange dans la partie Up) sont répertoriés. L'ensemble des minimums locaux est identifié par utilisation de la dérivée seconde (qui est nulle à ces moments-là, négative avant et positive après). Pour clarifier, la première dérivée du phénotype permet de déterminer la valeur du taux de décroissance et les extremums locaux du phénotype ; la dérivée seconde du phénotype permet d'identifier les moments où le taux

de décroissance du phénotype est minimal ou maximal (extremums locaux de la première dérivée). Une représentation graphique de l'identification de MR sur la dérivée première du phénotype corrigé et lissé et sur l'évolution du phénotype corrigé et lissé est présentée respectivement dans les Figure 18 et Figure 19:

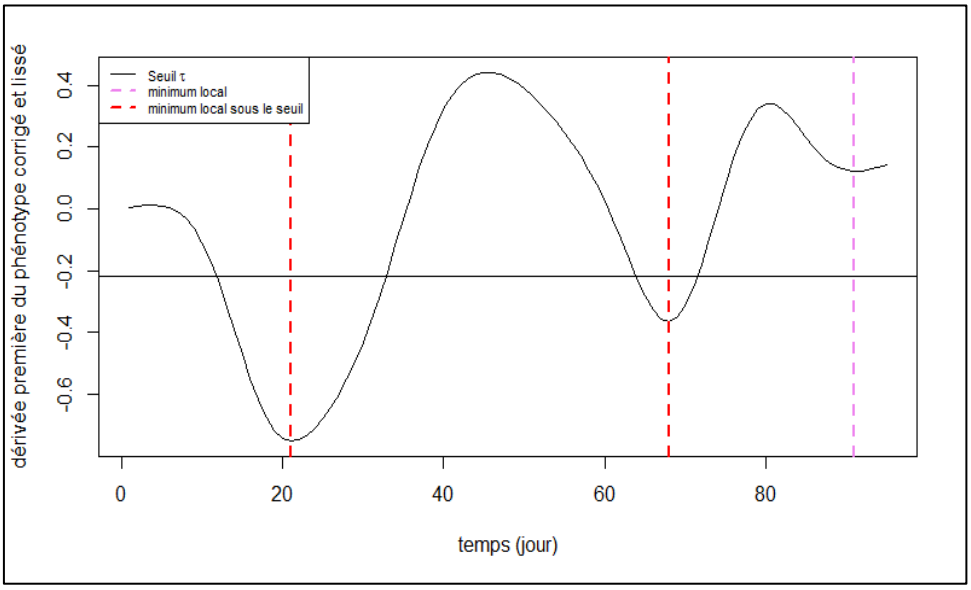


Figure 18 - Identification du (des) moment(s) de référence (MR) sur la dérivée première du phénotype corrigé et lissé

La ligne noire correspond au seuil τ calculé dans la partie Up grâce au modèle de mélange ; toutes les lignes verticales (rouge ou violette) représentent les minimums locaux de la pente (dérivée première) et les lignes rouges représentent les MR (minimums locaux où la valeur de la pente est inférieure au seuil τ).

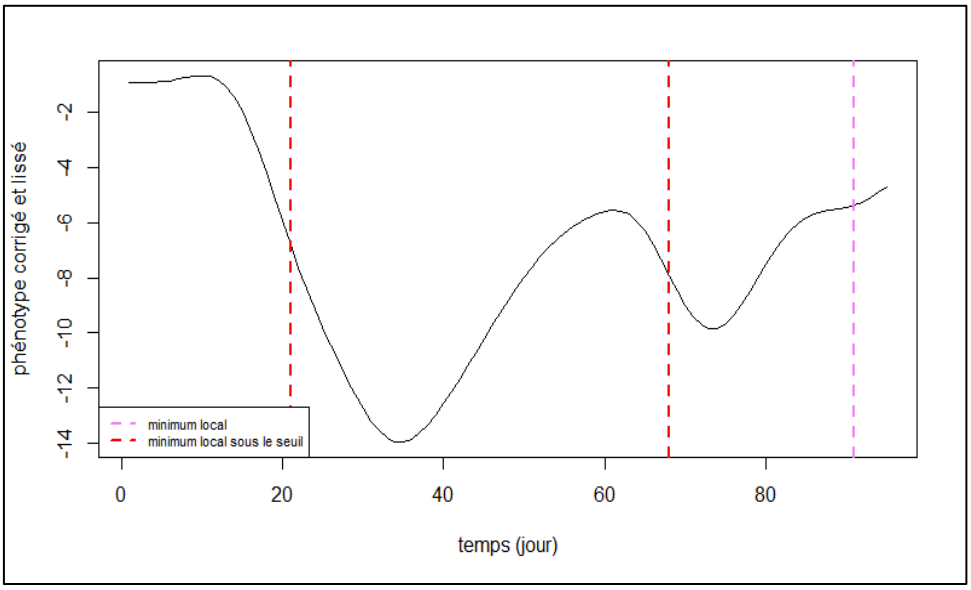


Figure 19 - Identification du (des) moment(s) de référence (MR) sur l'évolution du phénotype corrigé et lissé

Les lignes verticales (rouge et violette) représentent les minimums locaux de la pente (dérivée première) et les lignes rouges représentent les MR (minimums locaux où la valeur de la pente est inférieure au seuil τ calculé dans la partie Up grâce au modèle de mélange).

Pour la partie Down, contrairement à la partie Up, le travail commence à l'échelle de la bande, puis 'descend' à l'échelle de la case. Supposons que le travail est effectué à l'échelle de la bande, l'objectif est de confirmer si les perturbations de bande détectées se situent bien à cette échelle. Pour chaque bande détectée, une 'clusterisation' est appliquée sur les MR des cases-intra-bande. Si la majorité des cases-intra-bande (50%) est perturbée au même moment (c.-à-d. au moins la moitié de l'ensemble des cases ont un MR qui se trouve dans le même cluster), la perturbation de bande est validée. Ensuite, toutes les cases-intra-bande validée ayant plus de 1 MR vont être considérées comme ayant reçu plusieurs perturbations (potentiellement une perturbation à l'échelle de la case) et seront examinés à l'échelle de la case dans la partie Down. En revanche, s'il y a moins de 50% des cases-intra-bande qui sont perturbées au même moment, la bande n'est plus considérée comme perturbée, par conséquent, toutes les cases détectées comme perturbées (à l'échelle de la case dans la partie UP) de cette bande vont être considérées comme perturbées à l'échelle des cases et celles qui n'avaient pas été détectées dans la partie Up vont être considérées comme non-perturbées. Le seuil de 50% a été choisi arbitrairement et peut être modifié afin de varier la qualification de la méthode (c.-à-d. la sensibilité et la spécificité de la méthode, plus de détails concernant ces statistiques sont présentés dans la suite de ce document). Ensuite, le même travail à l'échelle de la case est effectué sur les individus-intra-cases détectées : la perturbation de case est validée si la majorité des individus-intra-case (50%) est perturbée au même moment.

Notez qu'une perturbation de bande peut affecter la médiane des phénotypes des individus-intra-case. Par conséquent, une perturbation de bande peut être détectée à plusieurs échelles (bandes, cases, individus). Afin d'éviter les 'doublons', nous avons eu besoin d'estimer le début et la fin des perturbations détectées (détail dans la section suivante). Plus précisément, à l'étape de la validation des cases, le coefficient de concordance de Kappa (qui sert à mesurer l'accord entre deux évaluateurs) est calculé entre l'estimation du début et de la fin de chaque perturbation à l'échelle de la case et l'estimation du début et de la fin de la perturbation à l'échelle de la bande à laquelle la case concernée appartient. Si la concordance est supérieure à 75% (ce qui signifie une forte chance que la perturbation détectée à l'échelle de la case soit celle de sa bande correspondante), la perturbation n'est plus prise en compte à l'échelle de la case. Le même travail est effectué pour les individus (avec la bande et la case à laquelle l'individu appartient). Pour mieux visualiser les démarches de la méthode Up &

Down, un arbre de prise de décision dichotomique concernant cette méthode est présenté dans Figure 20.

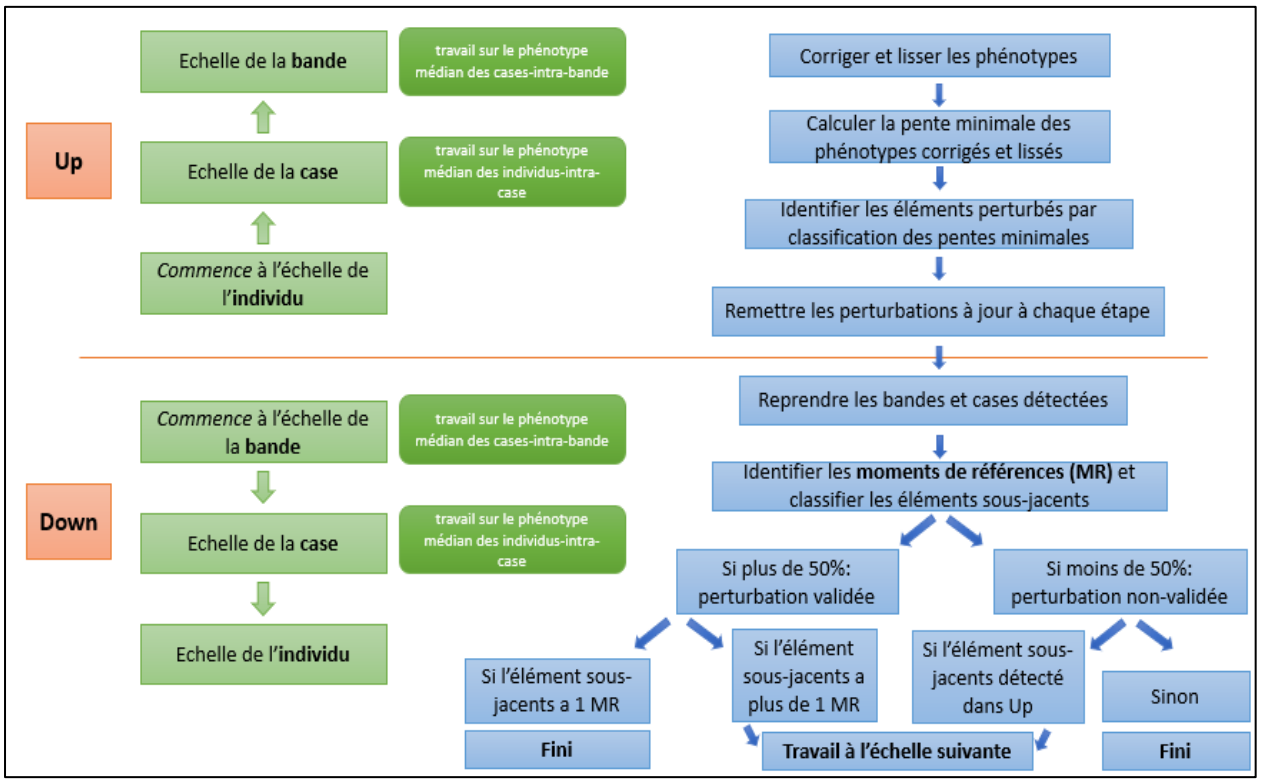


Figure 20 - Arbre de prise de décision dichotomique pour la méthode Up & Down

4.2.3 Caractérisation des perturbations

Dans un deuxième temps l'objectif est de caractériser les perturbations détectées ; c'est à dire identifier le début, la fin de la perturbation et quantifier son intensité (le travail se fait sur les phénotypes lissés corrigés à toutes les échelles, c'est-à-dire sur le phénotype médian des individus-intra-case à l'échelle de la case et sur le phénotype médian-médian des cases-intra-bande à l'échelle de la bande). Théoriquement, au début de la perturbation, l'évolution du phénotype s'infléchit le plus, puis à la fin de la perturbation, du fait de la résilience, le phénotype augmente plus rapidement pour récupérer sa courbe théorique. Une fois qu'il l'a rattrapé, il reprend son évolution normale. Par conséquent, le début de la perturbation correspond au moment où le taux de décroissance du phénotype corrigé lissé est maximal et donc au MR. La fin de la perturbation correspond au moment où le phénotype corrigé lissé arrête de diminuer. Pour déterminer ce point, tous les minimums locaux du phénotype corrigé et lissé sont identifiés par utilisation de la première dérivée (qui est nulle à

ces moments-là, négative avant et positive après). Notez qu'il doit être après un MR (Figure 21). S'il n'y a pas de minimum local après un MR, la fin de la perturbation théorique se situe probablement en dehors de la période d'observation. Dans ce cas, le dernier point d'observation sera considéré comme étant la fin de la perturbation.

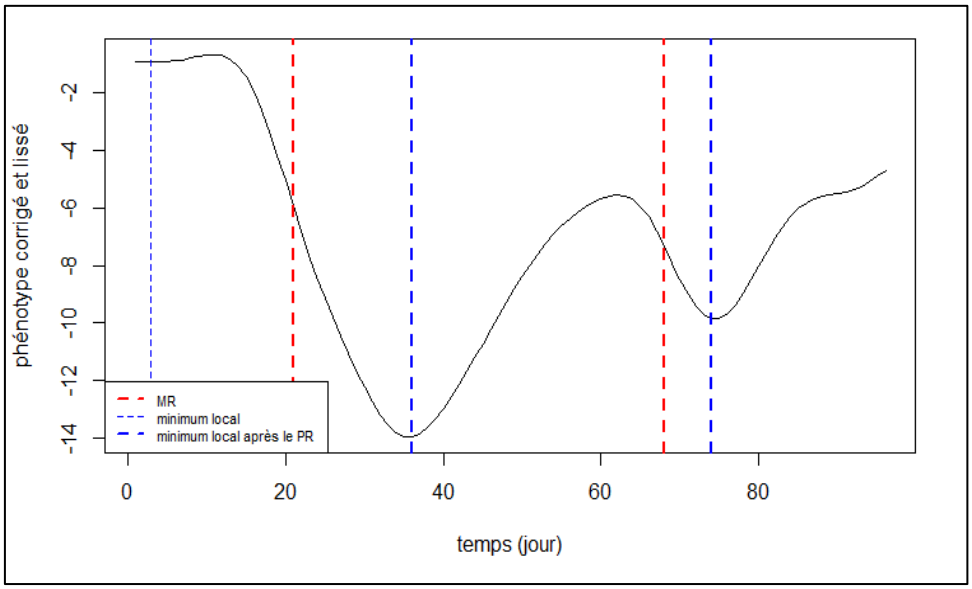


Figure 21 - Identification de la fin de la perturbation sur le phénotype corrigé et lissé.

La ligne rouge représente le début de la perturbation (ou MR) ; les lignes bleus représentent les minimums locaux du phénotype et les lignes bleues foncées correspondent aux fins estimées (les minimum locaux juste après les débuts de la perturbation).

De plus, pour tous les individus ayant reçu au moins une perturbation (peu importe l'échelle de la perturbation : bande, case ou individuel), la fin de la réaction de l'animal est aussi identifiée (elle servira plus tard pour quantifier la robustesse). Elle correspond au moment où le taux de décroissance du phénotype corrigé arrête de diminuer fortement et évolue comme si l'animal n'avait subi aucune perturbation. La raison est qu'à la fin de la réaction (le moment où l'animal rejoint sa courbe théorique non-perturbée), le taux de décroissance redevient « normal ». Pour déterminer ce point, nous avons identifié tous les maximums locaux de la pente du phénotype par utilisation de la seconde dérivée (qui est nulle à ces moments-là, positive avant et négative après). Notez qu'il doit être après la fin de la perturbation. A l'échelle du groupe (bande ou case), la caractérisation des perturbations nécessite une étape supplémentaire. Pour les bandes validées comme ayant reçu une perturbation de bande, les moments de la perturbation (début et fin) ont été calculés par la

médiane des moments des cases-intra-bandes concernées. Le même travail est effectué à l'échelle de la case.

Enfin, à l'échelle de la bande ou de la case, l'intensité de la perturbation est estimée par la valeur de la pente du phénotype médian entre le début et la fin de la perturbation. En effet, cette pente reflète l'intensité de la perturbation du groupe car tous les individus des groupes sont tirés au sort dans la population et sont repartis aléatoirement dans des cases et des bandes différentes (les cases de pleins frères ou de demis frères sont évitées).

4.3 Application de la méthode Up & Down

4.3.1 Données simulées

4.3.1.1 Qualification de la détection des perturbations

Pour qualifier la pertinence de la détection des éléments perturbés de la méthode Up & Down, cette dernière a été appliquée sur les données simulées (présentées dans l'article 1 – chapitre 3 du document) pour lesquelles les perturbations étaient connues (début, fin et intensité. Le set de paramètres sur lequel la méthode a été appliquée était : 20% pour chaque type de perturbation (bande, case et individu), corrélation neutre entre robustesse et production, corrélation neutre entre résistance et résilience, héritabilité modérée pour les composantes de la robustesse et l'intensité des perturbations suivant une loi uniforme $U([0{,}5,2{,}5])$.

La sensibilité (Se), spécificité (Sp), valeur prédictive positive (VPP), valeur prédictive négative (VPN) ont été utilisées pour juger la qualité de la méthode Up & Down pour détecter les perturbations. Les vrais positifs (VP) (resp. vrais négatifs (VN)) sont les éléments (individu, case, bande) qui sont bien détectés comme perturbés (resp. non-perturbés) et les faux positifs (FP) (resp. faux négatifs (FN)) sont les éléments qui sont détectés comme perturbés (resp. non-perturbés) à tort. La sensibilité (Se) est la probabilité de détecter l'élément lorsque l'élément est vraiment perturbé : $\mathrm{Se} = \mathrm{VP}/(\mathrm{VP} + \mathrm{FN})$. La spécificité (Sp) est la probabilité de ne pas détecter l'élément lorsque l'élément n'est vraiment pas perturbé : $\mathrm{Sp} = \mathrm{VN}/(\mathrm{VN} + \mathrm{FP})$. La valeur prédictive positive est la probabilité d'avoir été réellement perturbé lorsqu'on est détecté comme perturbé : $\mathrm{VPP} = \mathrm{VP}/(\mathrm{VP} + \mathrm{FP})$. La valeur prédictive négative est probabilité de ne pas avoir été réellement perturbé lorsque l'on n'est pas détecté comme perturbé : $\mathrm{VPN} = \mathrm{VN}/(\mathrm{VN} + \mathrm{FN})$.

4.3.1.2 Qualification des caractéristiques des perturbations

Pour qualifier les moments des perturbations détectées par la méthode Up & Down, l'écart entre le vrai début de perturbation et le début estimé (edp), l'écart entre la vraie fin de perturbation et la fin de perturbation estimée (efp), l'écart entre la vraie durée de perturbation et la durée de perturbation estimée ($edur$) ont été calculés. En parallèle, la corrélation entre l'intensité estimée et l'intensité théorique des perturbations (cor_int) a été mesurée pour juger la qualité de quantification de la perturbation. Ces statistiques ont été analysées à l'échelle de la bande, de la case et de l'individu.

A cette étape, si plusieurs perturbations ont été détectées pour un même élément à la même échelle (ce qui ne peut être le cas dans les simulations : une bande simulée ne peut subir qu'une seule perturbation de bande et de même pour l'échelle de case et de l'individu), nous avons fait le choix de juger la qualité de la qualification en utilisant uniquement les informations de la perturbation la plus proche temporellement de celle simulée identifiée en utilisant le coefficient de concordance de Kappa.

4.3.1.3 Statistiques par classe d'intensité, moment et durée

L'impact de l'intensité, de la durée, et du moment de la perturbation sur la sensibilité et la qualité de qualification de la méthode Up & Down a été évaluée. Trois classes ont été établies pour l'intensité ('faible', 'moyenne' ou 'forte'), pour la durée ('courte', 'moyenne' ou 'longue') et pour le moment (apparition 'tôt', 'moyenne' ou 'tard'). Ces classes sont présentées dans la Table 6. Leur impact a été évalué par une régression logistique pour la sensibilité et une régression linéaire pour la qualification des caractéristiques des perturbations (edp, efp, $edur$).

Table 6 - Détail des classes d'intensité, durée et moment du début de perturbation.

Classe	A	B	C
Intensité	[0,5, 1,17[	[1,17, 1,83[	[1,83, 2,5]
Durée (nombre de jours)	]1, 8[	[8, 16[	[16, 25]
Moment (j)	]2, 33[	[33, 66[	[66, 100]

4.3.1.4 Quantification de la robustesse

En fonction des informations concernant les perturbations (notamment le début et la fin de la perturbation et la fin de la réaction des animaux), un nouveau caractère de la robustesse a été déterminé. Pour rappel, nous avons défini la robustesse d'un animal d'élevage comme la combinaison entre sa résistance (capacité de minimiser l'impact direct de la perturbation) et sa résilience (capacité de récupérer rapidement afin de rejoindre l'état avant l'exposition de la perturbation). Pour chaque perturbation de groupe qu'un animal a subie (identifiée par la méthode Up & Down), la valeur minimale de la pente du phénotype individuel corrigé et lissé a été utilisée comme un nouveau phénotype de la résistance, et la valeur de la pente du phénotype individuel corrigé et lissé entre la fin de la perturbation et la fin de réaction de l'animal a été utilisée comme un nouveau phénotype de la résilience. Les valeurs de résistance et résilience ne sont pas récupérées dans le cas de perturbation individuelle. En effet dans ce cas-là, la pente minimale du phénotype individuel corrigé et lissé est le reflet de l'intensité de la perturbation et de la robustesse de l'animal qui sont confondues.

Un modèle animal à répétabilité simple (avec l'intensité, la durée de la perturbation et l'âge de l'animal au début de la perturbation comme effets fixes) a été appliqué sur les nouveaux phénotypes de résistance et résilience. Nous avons ensuite estimé les héritabilités et les VG de la résistance et résilience en utilisant le logiciel ASREML 3.0 (Gilmour et al., 2009) et calculé la précision des VG prédites obtenues (corrélation entre VG prédite et simulée). Cette analyse a été effectuée sur 3 sets de paramètres de simulations concernant les types de perturbations : i) cas où 20% des bandes et 20% des cases sont perturbées, ii) cas où 30% des bandes et 30% des cases sont perturbées, iii) cas où 40% des bandes et 40% des cases sont perturbées. Dans tous les cas simulés, aucune perturbation individuelle n'a été simulée car les phénotypes de résistance et résilience ne sont pas analysés en cas de perturbation individuelle. Une corrélation neutre entre robustesse et production, une corrélation neutre entre résistance et résilience et une héritabilité modérée pour la robustesse ont été considérées pour tous les sets de paramètres utilisés.

4.3.2 Données réelles

La méthode Up & Down a été appliquée sur le poids et la consommation alimentaire cumulée des animaux de la station de phénotypage du Rheu (corrigé dans le chapitre 2 du document). A l'échelle de la bande et de la case, les observations à partir du moment où le premier porc du groupe est parti à l'abattoir ont été supprimées pour éviter un changement dans l'évolution des courbes médianes à cause d'une modification du groupe d'individus. Pour rappel, en général, ce sont les porcs les plus lourds qui partent en premier à l'abattoir et la médiane du groupe calculée en absence de ces animaux peuvent fortement être modifiée, ce qui entrainerait des erreurs dans le fonctionnement de la méthode Up & Down.

Enfin, les éléments détectés (bande, case, individu) sur les données réelles ont été comparés avec les éléments qui sont théoriquement perturbés d'après les données santé et météorologiques. De la même manière que les données simulées, seules les perturbations les plus proches entre celles théoriques et détectées ont été utilisées pour mesurer la pertinence de la qualification des caractéristiques (début et fin) des perturbations détectées.

4.4 Résultats

4.4.1 Données simulées

4.4.1.1 Qualité de la détection de perturbation

Une des fonctions de lissage utilisées pour la méthode Up & Down est le lissage non-paramétrique de Nadaraya-Watson (Tsybakov, 2009) avec un noyau gaussien et une taille de fenêtre égale à 10. Dans la suite de ce document, les résultats présentés seront basés sur cette fonction. Les Se, Sp, VPN et VPP de la méthode Up & Down pour détecter des perturbations à l'échelle globale (ensemble des perturbations sans prendre en compte l'échelle des perturbations), bande, case et individu sont présentées dans la Table 7. Une forte spécificité a été obtenue pour chaque échelle (99% à l'échelle de bande, 98% pour les cases et 95% pour les individus). La sensibilité diminue de l'échelle de la bande à celle de l'individu avec une valeur de 0,93 pour la bande, 0,73 pour la case et 0,43 à l'échelle de l'individu.

Table 7 - La sensibilité (Se), spécificité (Sp), valeur prédictive positive (VPP) et valeur prédictive négative (VPN) à l'échelle globale, bande, case et individu pour la méthode Up & Down utilisant la fonction de lissage Nadaraya-Watson (noyau gaussien et taille de fenêtre égale à 10) sur 1000 simulations.

Echelle	Se	Sp	VPP	VPN
Global	0,85	0,93	0,91	0,87
Bande	0,93	0,99	0,97	0,98
Case	0,73	0,98	0,91	0,94
Individu	0,43	0,95	0,69	0,87

4.4.1.2 Qualité des caractéristiques des perturbations détectées

Les valeurs absolues de edp, efp et $edur$ et cor_int sont calculées pour évaluer les caractéristiques des perturbations détectées. Les résultats à l'échelle globale (ensemble des perturbations), bande, case et individu sont donnés dans la Table 8. En médiane, le début estimé est à 3 jours du vrai début, la durée estimée est à 3 jours de la durée théorique, et la fin estimée est à 2 jours de la fin théorique pour toutes les échelles. La corrélation entre l'intensité estimée et théorique est de -0,78 pour la bande, cette valeur diminue légèrement à l'échelle de la case (-0,72) et encore plus à l'échelle de l'individu (-0,50).

Table 8 - Statistiques pour l'évaluation des caractéristiques des perturbations à l'échelle globale, bande, case et individu sur 1000 simulations.

Echelle	Edp			Efp			$Edur$			Cor_int
	Q1	MED	Q3	Q1	MED	Q3	Q1	MED	Q3	
Global	1	3	4	1	2	3	2	3	5	-0,71
Bande	1	3	4	1	2	3	2	3	4	-0,78
Case	2	3	4	1	2	3	2	3	5	-0,72
Individu	1	3	5	1	2	4	1	3	5	-0,50

Les statistiques sont respectivement les écarts en valeur absolue entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp), entre la durée estimée et la vraie durée ($edur$) et la corrélation entre la pente du phénotype lissé entre le début et fin estimés (intensité estimée) et l'intensité théorique (cor_int). Q1, MED, Q3 représentent respectivement le quartile inférieur, la médiane et le quartile supérieur.

4.4.1.3 Statistiques par classe d'intensité, moment et durée

Sensibilité

Les Figure 22, Figure 23, Figure *24* montrent le changement dans Se selon différentes classes d'intensité, durée et moment à l'échelle de la bande, case et individu respectivement. Il y a un effet significatif de l'intensité, de la durée et du début de perturbation pour toutes les échelles. Plus l'intensité est forte et la durée est longue, plus Se augmente (à l'échelle de la bande, nous avons 2,63 fois plus de chance de détecter une perturbation lorsque l'intensité est forte que lorsqu'elle est faible et 4,81 fois plus lorsque la durée est longue que lorsqu'elle est courte). En revanche, Se diminue lorsque la perturbation se produit tardivement (à l'échelle de la bande, la probabilité de détecter une perturbation lorsque le début de la perturbation est tardif est 6,25 fois plus faible que lorsque le début de la perturbation est plus précoce). Les mêmes tendances sont observées à l'échelle des perturbations de case et individu (à l'exception de l'effet du moment sur les perturbations individuelles).

Variable		N	Odds ratio		p
Intensité	[0.5,1.17[	2721		Référence	
	[1.17,1.83[	2676		2.94 (2.32, 3.73)	<0.001
	[1.83,2.5]	2605		2.63 (2.09, 3.31)	<0.001
Durée	[1,8[	2760		Référence	
	[8,16[	2731		15.34 (10.89, 22.33)	<0.001
	[16,25]	2511		4.81 (3.71, 6.32)	<0.001
Début	[2,33[	2493		Référence	
	[33,66[	2734		0.79 (0.56, 1.10)	0.2
	[66,99]	2775		0.16 (0.12, 0.22)	<0.001

Figure 22 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de bande (résultats sur 1000 simulations).

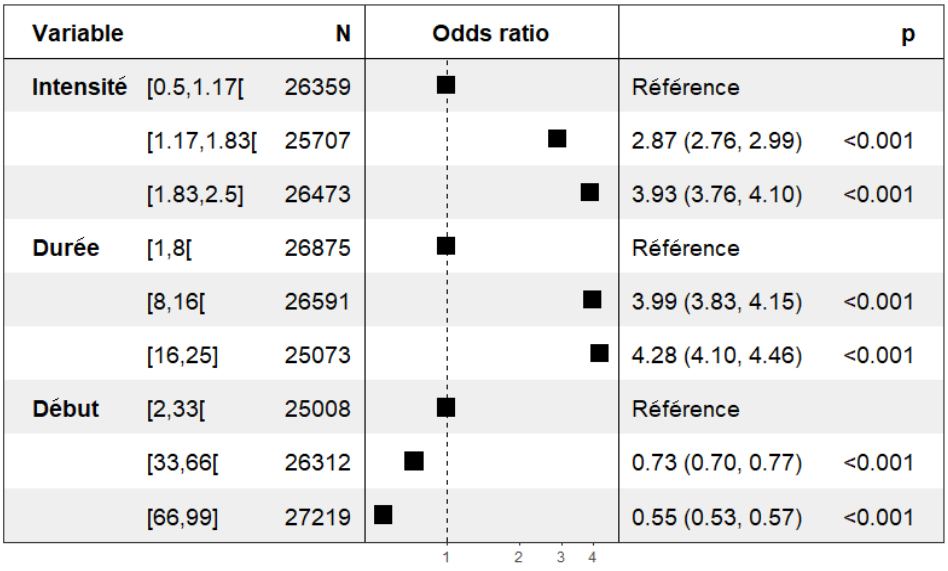


Figure 23 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de case (résultats sur 1000 simulations)

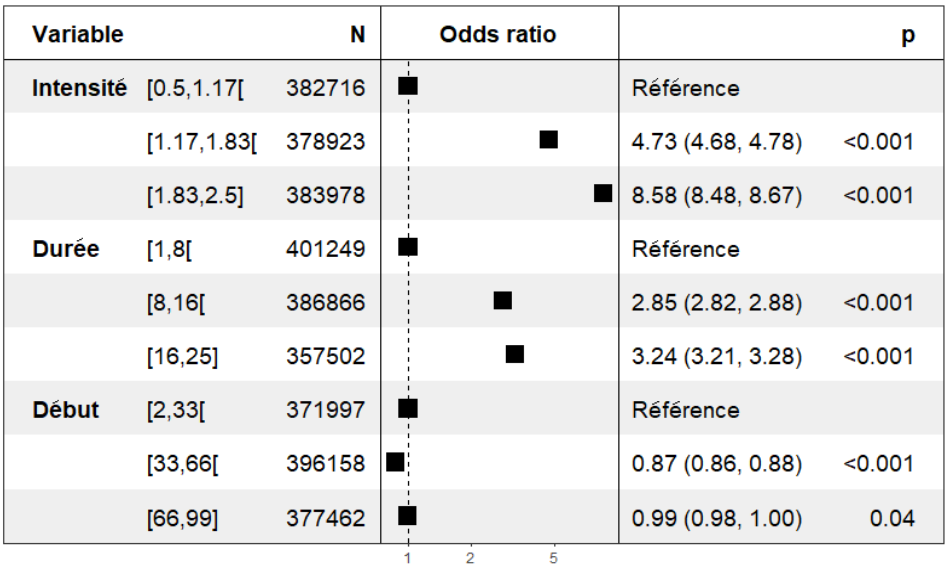


Figure 24 - L'effet de l'intensité, de la durée et du début de la perturbation sur la sensibilité à l'échelle de l'individu (résultats sur 1000 simulations)

La caractérisation des perturbations détectées par classes

La corrélation entre l'intensité estimée et l'intensité vraie de la perturbation varie selon la durée et le début de la perturbation (Figure 25). Plus précisément, cette corrélation (en valeur absolue) est plus forte lorsque la durée de la perturbation est longue que lorsqu'elle

est courte (la corrélation moyenne en valeur absolue passe de 0,70 à 0,97 à l'échelle de la bande, de 0,54 à 0,91 pour la case et de 0,31 à 0,60 pour les individus). En revanche, elle est plus faible lorsque le début est tardif que lorsqu'il est précoce (la corrélation moyenne en valeur absolue passe de 0,81 à 0,74 pour la bande, de 0,75 à 0,69 pour la case et de 0,53 à 0,49 pour les individus).

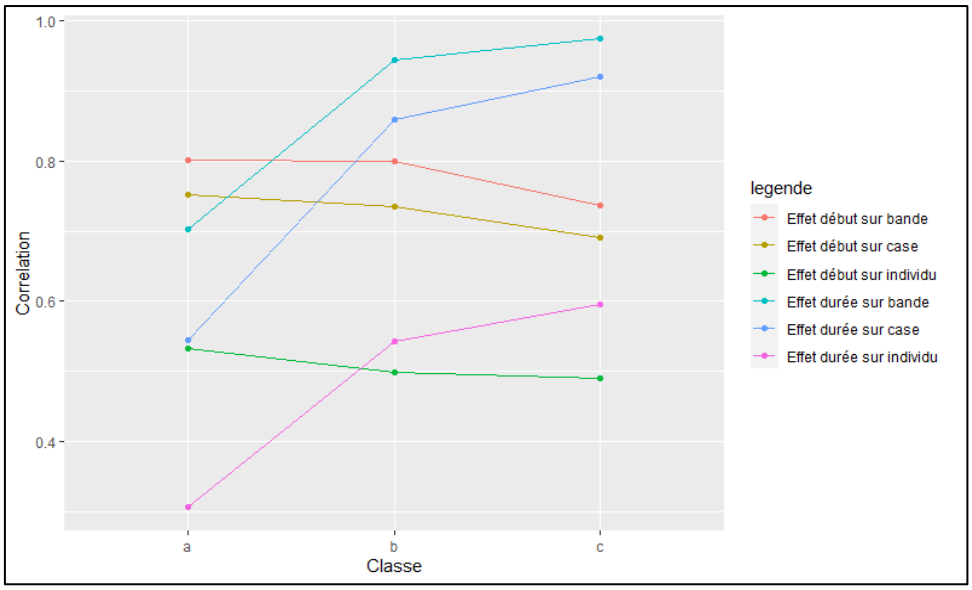


Figure 25 - L'effet de la durée et du moment du début de la perturbation sur la corrélation entre l'intensité estimée et théorique par niveau de perturbation (bande, case, individu).

Aucun effet significatif de l'intensité sur edp, efp et $edur$ n'a été mis en évidence à l'échelle de la bande et de la case (Figure 26). Au contraire, un effet significatif de la durée sur ces statistiques a été observé à toutes les échelles (bande, case, individu) (Figure 27). Une durée longue entraine en général une sous-estimation de la fin et de la durée de la perturbation : en moyenne, efp étant entre 0 et -2 selon l'échelle étudiée (contrairement à des valeurs positives lorsque la durée est courte) et $edur$ étant inférieur à -3 pour toutes les échelles. En ce qui concerne l'effet du moment du début de la perturbation sur les statistiques concernées, nous observons ainsi un effet significatif à toutes les échelles : un début tardif entraine en générale une sous-estimation de efp et $edur$ (Figure 28).

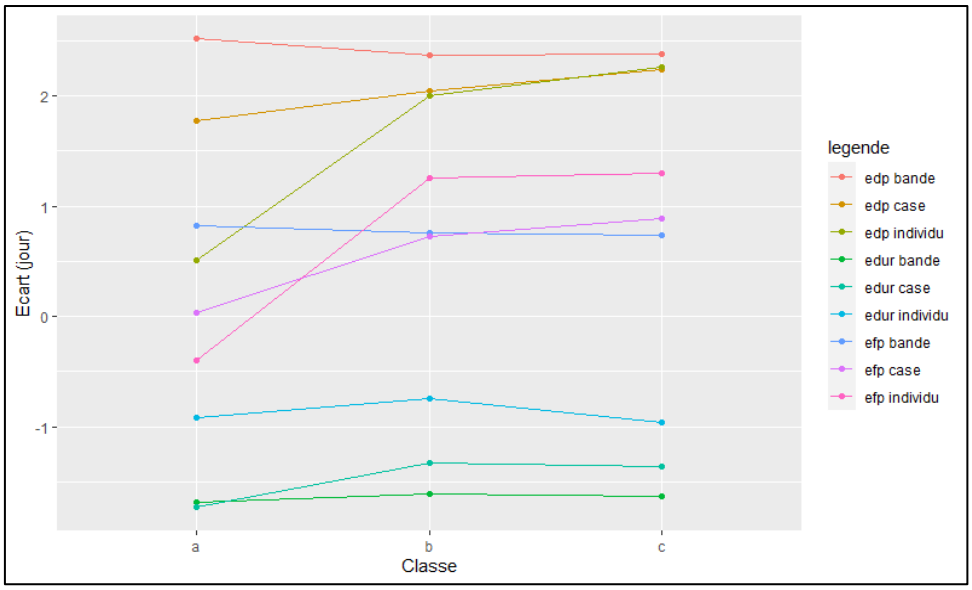


Figure 26 - L'effet de l'intensité de perturbation sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations)

Les classes a, b, c représentent respectivement ([0,5, 1,17[, [1,17, 1,83[, [1,83, 2,5]) pour l'intensité

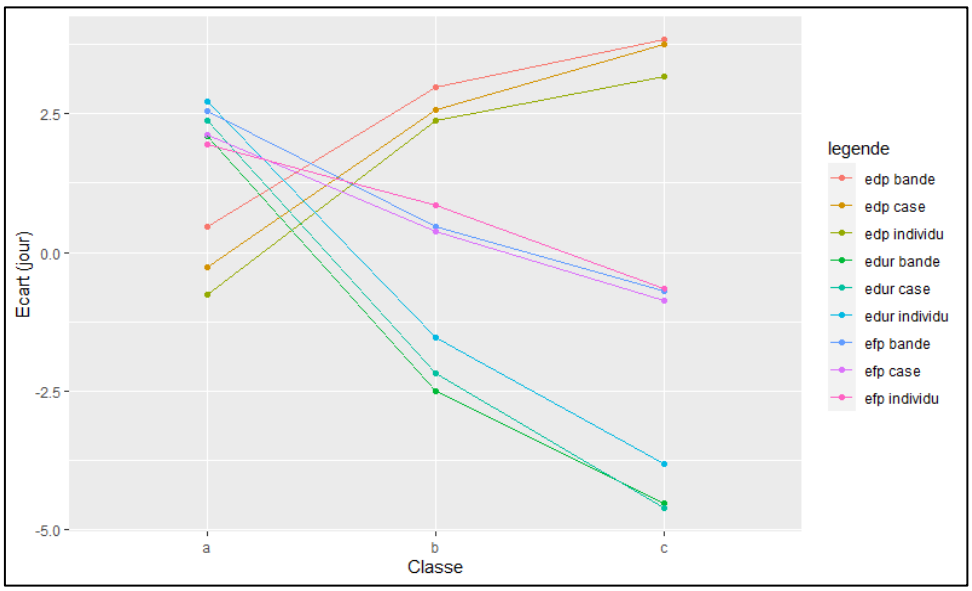


Figure 27 - L'effet de la durée de perturbation sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations)

Les classes a, b, c représentent respectivement ([1, 8[, [8, 16[, [16, 25]) pour la durée.

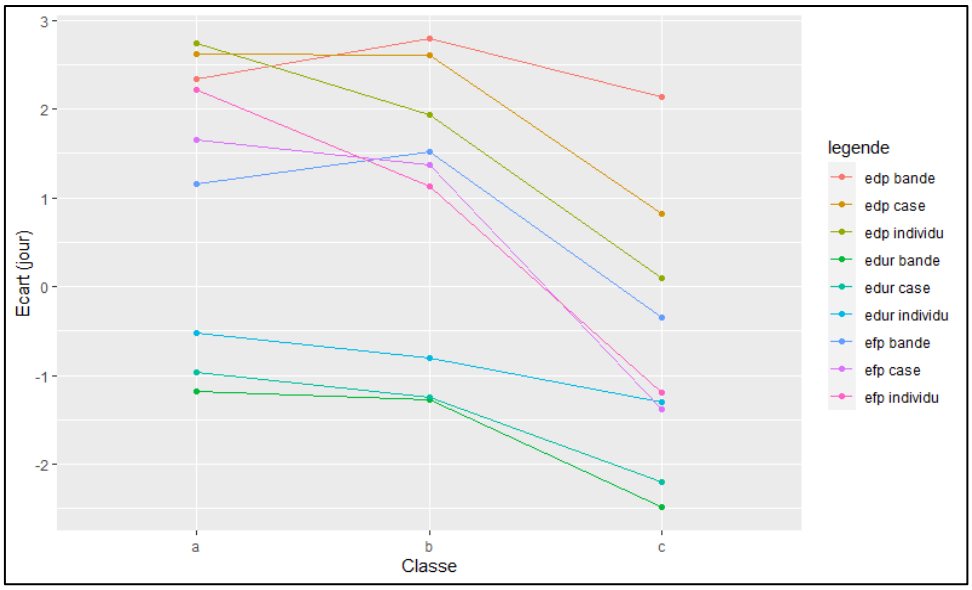


Figure 28 - L'effet du début perturbé sur l'écart en moyenne entre le début estimé et le vrai début (edp), entre la fin estimée et la vraie fin (efp) et entre la durée estimée et la vraie durée ($edur$) à l'échelle de la bande, case et individu (résultats sur 1000 simulations)

Les classes a, b, c représentent respectivement ([2, 33[, [33, 66[, [66, 99]) pour le moment.

4.4.1.4 Quantification de la robustesse

La Table 9 représente l'héritabilité estimée et la qualité des VG prédites pour les deux composantes de la robustesse (résistance et résilience) en fonction du pourcentage de perturbations simulées sur 1000 simulations. Une augmentation du pourcentage de perturbations simulées tend à entrainer une diminution dans l'estimation des héritabilités des deux composantes de la robustesse. En outre, plus le nombre d'éléments perturbés est élevé, plus la précision de la VG prédite pour la résistance et la résilience augmente.

Table 9 - Héritabilité estimée et précision de la VG prédite pour la résistance et résilience en fonction du pourcentage de perturbations (résultats sur 1000 simulations)

% perturbation de bande	% perturbation de case	h^2 résistance	h^2 résilience	précision résistance	précision résilience
20	20	0,12 ± 0,04	0,12 ± 0,06	0,37 ± 0,08	0,11 ± 0,09
30	30	0,10 ± 0,04	0,11 ± 0,05	0,40 ± 0,07	0,13 ± 0,10
40	40	0,09 ± 0,03	0,09 ± 0,04	0,42 ± 0,07	0,13 ± 0,09

4.4.2 Données réelles

Le nombre d'individus, cases et bandes détectés comme perturbés selon le phénotype choisi pour la méthode Up & Down sont présentés dans la Table 10. Plus d'éléments à toutes les échelles (bande, case, individu) ont été détectés en utilisant la CFI que le poids vif des animaux.

Table 10 – Nombre d'éléments détectés selon les phénotypes utilisés (poids vif ou CFI) pour la méthode Up & Down.

	Poids vif		CFI	
Echelle	Nombre d'éléments détectés	% par rapport à la population totale	Nombre d'éléments détectés	% par rapport à la population totale
Bande	2	4	3	6
Case	51	11	80	17
Individu	675	11	1357	22

Selon le phénotype choisi pour la méthode Up & Down, différents éléments (bandes, cases ou individus) sont détectés comme perturbés, cependant, il existe des éléments qui sont détectés à la fois avec les deux phénotypes : 1 bande, 26 cases et 460 individus en commun. La Table 11 représente, pour ces éléments en commun, la différence (médiane des valeurs absolues des écarts) entre les estimations des débuts (ou fins) des perturbations détectées avec le poids et CFI. Nous observons plus de similitudes pour les perturbations individuelles que pour les perturbations à l'échelle de la case. Néanmoins, à l'échelle des bandes, étant donné le faible nombre de bandes détectées en commun entre les deux phénotypes (1 bande), il est difficile de comparer les résultats avec ceux obtenus à d'autres échelles.

Table 11 – La médiane des écarts en valeur absolue entre les débuts (edp) (ou fins (efp)) des perturbations détectées avec le poids et CFI pour toutes les échelles (bande, case, individu).

Echelle	edp	efp
Bande	0	2
Case	4	10
Individu	2	3

Concernant les résultats sur le poids vif des animaux, toutes les bandes détectées comme perturbées avec la méthode Up & Down ont traversé une période où les jours avaient un THI élevé (pour les résultats concernant la CFI, 66,6% des bandes détectées se retrouvent également dans la base de données météorologiques). A l'échelle des cases, 8% et 7% des cases détectées, respectivement pour le poids et la CFI, ont reçu des traitements pour des perturbations de cases. A l'échelle des individus, 51% et 39% des animaux détectés, respectivement pour le poids et la CFI, ont reçu des traitements pour des perturbations individuelles.

Enfin, à propos de la qualité de la caractérisation des perturbations détectées (pour les éléments détectés se trouvant dans la BDD santé ou météorologique), la médiane des écarts en valeur absolue entre les débuts estimés et théoriques varie entre 1,5 et 3 jours selon l'échelle étudiée. Ainsi, lorsqu'un élément est détecté, le moment de la perturbation estimée affectant cet élément correspond au moment enregistré dans les BDD théoriques.

4.5 Discussion

Plusieurs méthodes pour détecter des perturbations environnementales inconnues ont été proposées dans la littérature. Une stratégie possible est de reconstruire la courbe théorique montrant l'évolution non-perturbée du phénotype puis analyser la déviation entre cette courbe et la courbe observée bruitée. Une autre stratégie consiste à utiliser des informations concernant les groupes des animaux pour détecter les jours où les animaux ont été affectés par une ou plusieurs perturbations. Différentes méthodes concernant ces stratégies ont été présentées en détail dans le chapitre « Etat de l'art ».

Durant cette étude, plusieurs méthodes de lissage ont été testées (autres que la fonction de lissage non-paramétrique Nadaraya-Watson). Il s'agit des polynômes de degrés différents, B-splines, moyennes mobiles, régression par polynômes locaux, exponentiel (Wood, 2006). Les résultats concernant ces fonctions de lissages sont présentés dans la partie 'Analyses et résultats supplémentaires'. Tout au long de ce document, les résultats de la fonction Nadaraya-Watson donnant les meilleurs résultats en termes de sensibilité/spécificité par rapport aux autres fonctions ont été présentés.

En outre, dans cette étude, pour séparer la population en deux groupes différents (perturbé et non-perturbé), nous avons utilisé comme indicateur la valeur maximale du taux de décroissance du phénotype corrigé et lissé des animaux (en basant sur l'hypothèse que les animaux d'élevage ont tendance à ralentir leur production durant une perturbation environnementale). Cependant, il existe d'autres alternatives pour déterminer les éléments qui ont été affectés par une ou plusieurs perturbations. Il est possible d'utiliser les indicateurs proposés par Berghof *et al.* (2018) (l'erreur quadratique moyenne, la symétrie et l'autocorrélation des résidus) ou par Garcia-Baccino *et al.* (2021) (coefficient de variation de l'ensemble de la population par jour). Toutes ces statistiques ont aussi été testées durant cette étude de thèse et les détails concernant les résultats de ces différentes méthodes sont présentés dans la partie 'Analyses et résultats supplémentaires'.

D'après les résultats sur simulation, Se augmente de l'échelle individuelle à l'échelle de la bande. Cette amélioration s'explique par le fait que les phénotypes à l'échelle des groupes ont été obtenus en utilisant la médiane des phénotypes de l'échelle sous-jacente. Cela permet, d'une part, de diminuer le bruit et de corriger les phénotypes des robustesses individuelles, d'autre part. En effet, à l'échelle individuelle, le taux de décroissance maximale du phénotype est fonction de l'intensité de la perturbation et de la robustesse de l'animal. L'utilisation de la médiane permet de corriger ces réponses individuelles, sous l'hypothèse d'une répartition homogène des individus intra groupe ce qui est recommandé en élevage de sélection. Néanmoins, tous les groupes n'étant pas parfaitement identiques, la médiane a été préférée à la moyenne pour limiter l'influence d'animaux de robustesse extrême sur le phénotype calculé à l'échelle du groupe. L'identification des perturbations à l'échelle du groupe permet de comparer les individus ayant subi la même perturbation, ce qui est l'objectif

des sélectionneurs (en revanche, il semble difficile de quantifier la robustesse des animaux ayant subi des perturbations individuelles différentes).

La méthode Up & Down fourni aussi une excellente spécificité pour toutes les échelles (entre 95% et 98%). Ce résultat est très rassurant puisque la proportion d'animaux pour lesquels une mesure de la robustesse sera obtenue à tort (puisque n'ayant pas subi de perturbation) sera faible et ne devrait avoir que peu d'impact sur l'estimation des VG de robustesse des animaux. Les statistiques pour qualifier les caractéristiques des perturbations (edp, efp, $edur$) et la corrélation entre l'intensité théorique et estimée (cor_int) sont aussi correctes. Les écarts sont inférieurs ou égaux à 3 jours pour toutes les échelles et cela montre que les moments des perturbations estimés sont proches des vrais moments. La pente minimale du phénotype corrigé et lissé que nous avons utilisée pour i) calculer l'intensité des perturbations et ii) mesurer la robustesse des animaux d'élevage a donc de fortes chances d'être pertinente. En effet, cela se confirme au niveau du calcul de l'intensité : les corrélations entre l'intensité simulée et estimée sont supérieures à 0,72 à l'échelle de groupe. De plus, le moment où la pente minimale est calculée pour estimer l'intensité est compris entre le vrai début et la vraie fin de la perturbation dans 94% de fois. Ces résultats montrent que la méthode Up & Down est pertinente non seulement dans la détection des perturbations, mais elle l'est aussi dans la qualification des moments perturbés et l'intensité des perturbations. Cela est très important car la quantification de la robustesse des animaux d'élevage dépend fortement de la précision des pentes individuelles calculées entre les moments perturbés.

Les résultats par différentes classes d'intensité, durée et début de perturbation ont montré des tendances sur la détection des perturbations. Une augmentation de l'intensité et de la durée de la perturbation améliore Se, en revanche, un début tardif de la perturbation entraine une diminution de Se. Pour l'intensité, cette tendance peut être expliquée premièrement par le fait qu'un animal subissant une perturbation de forte intensité a plus de chance de l'exprimer (en fonction de la pente minimale) que lorsque l'intensité est faible (la perturbation aurait peu d'impact sur l'animal). Ensuite, le fait d'utiliser un modèle de mélange afin de séparer la population en deux parties (perturbée et non-perturbée) est aussi une explication pour cette tendance : les deux courbes du modèle se superposent et il est possible d'avoir des animaux étant sous la courbe « perturbée » mais qui ont une pente minimale supérieure au seuil séparant les deux groupes et qui vont donc être considérés comme non-perturbés. Cela arrive plus souvent si l'intensité de la perturbation est faible. L'effet de la

durée sur Se peut s'expliquer par les paramètres du lissage utilisés. En effet, la taille de la fenêtre de la fonction Nadaraya-Watson a été fixée à 10 pour être cohérente avec la période d'observation de 100 jours (pour clarifier, cette fonction consiste à calculer des moyennes glissantes avec un noyau gaussien). Il pourrait être possible d'améliorer la détection des perturbations de courte durée en modifiant cette taille de fenêtre (une plus petite taille pourrait entrainer une meilleure détection des perturbations de très courte durée). Enfin, une perturbation tardive a moins de chance d'être détectée vraisemblablement parce que les animaux n'ont pas eu le temps d'exprimer de réponse face à la perturbation avant la fin de la période de test qui est de 100j.

La durée et le moment de perturbation ont aussi un impact significatif sur les statistiques pour la qualification des perturbations. En présence d'une perturbation de longue durée, la méthode a tendance à identifier un début de perturbation plus tardif, une fin plus précoce et en conséquence une durée plus courte que les vraies caractéristiques de la perturbation. Une explication possible à ce phénomène est une fois de plus le changement du lissage. Une illustration de cette explication est donnée dans la Figure 29 illustrant la réponse d'un animal à deux perturbations intervenant au même moment et de même intensité mais de durée différente : lorsque la durée est longue (24 jours – lissage en noir), la méthode a tendance à surestimer le début et sous-estimer la fin de la perturbation par rapport à lorsque la durée est courte (6 jours – lissage en rouge) : l'écart entre le début estimé et théorique est de 1 et 0 respectivement pour le cas de durée longue et courte ; l'écart entre la fin estimée et théorique est de -2 et 1 respectivement pour le cas de durée longue et courte.

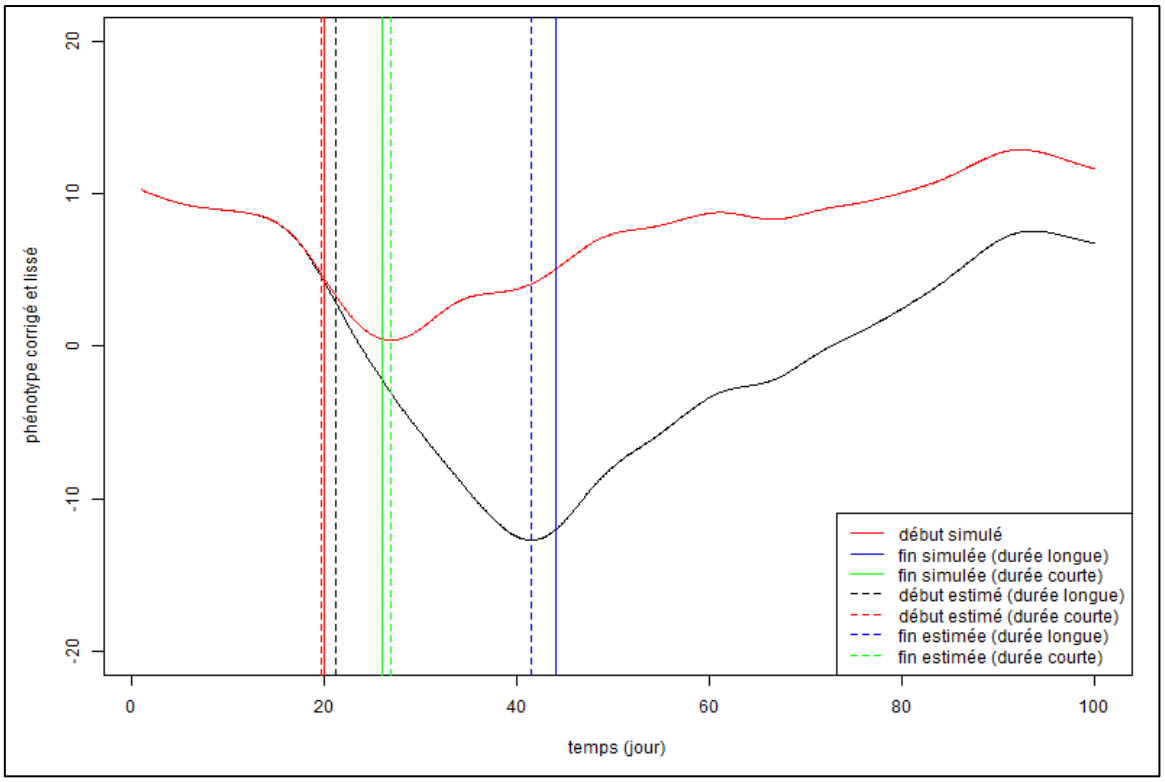


Figure 29 - Effet de la durée de la perturbation sur la caractérisation de la perturbation sur le poids vif corrigé et lissé.

Les deux courbes de lissages noires et rouges représentent respectivement l'évolution du phénotype dans le cas où la perturbation est longue (24 jours) et courte (6 jours) ; les lignes verticales pleines rouge, bleue et verte représentent respectivement le début théorique de la perturbation (commun pour les deux cas), la fin théorique de la perturbation longue et courte ; les lignes verticales pointillées noire, rouge, bleue et verte représentent respectivement le début estimé dans le cas de perturbation longue et courte, et la fin estimée dans le cas de perturbation longue et courte.

Lorsque qu'une perturbation se produit tardivement, la fin de la perturbation est identifiée comme plus précoce que la fin réelle de la perturbation et donc sa durée est également sous-estimée. Cela est dû, une fois de plus, au fait que la période d'observation a une limite de 100 jours. La fin de la perturbation est hors de la période d'observation et donc estimée par défaut à 100j. Un exemple de ce phénomène est présenté dans la Figure 30: lorsque le moment est tardif (jour 93 – lissage en noir), la méthode a tendance à sous-estimer la fin de la perturbation par rapport à lorsque la perturbation arrive plus tôt (jour 27 – lissage en rouge) : l'écart entre la fin estimée et théorique est de -3 et 1 respectivement pour le cas de moment tard et tôt.

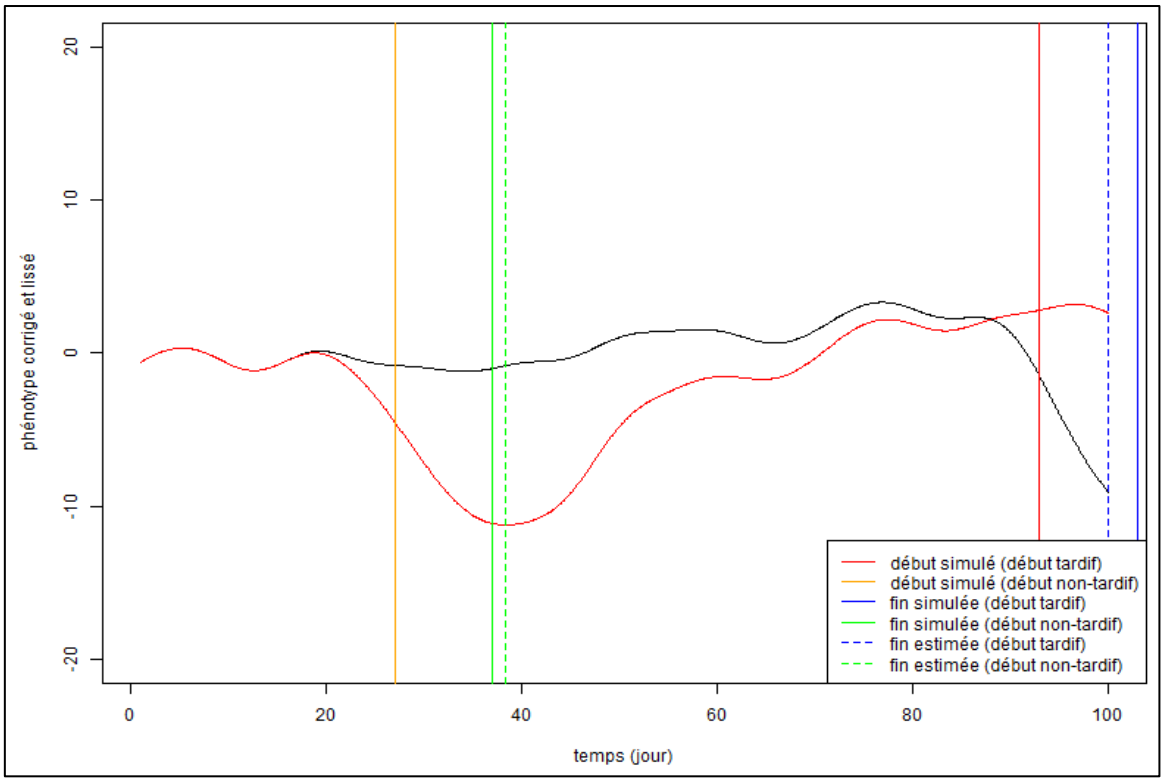


Figure 30 - Effet du moment perturbé sur la caractérisation de la perturbation sur le poids vif corrigé et lissé.

Les deux courbes de lissages noires et rouges représentent respectivement l'évolution du phénotype dans le cas où la perturbation arrive tardivement (jour 93) et pas tardivement (jour 27) ; les lignes verticales pleines rouge, orange, bleue et verte représentent respectivement le début théorique de la perturbation dans le cas début tardif et non, la fin théorique de la perturbation dans le cas début tardif et non; les lignes verticales pointillées bleue et verte représentent respectivement la fin estimée dans le cas de perturbation tardive et non.

Pour les données réelles, la méthode Up & Down ne détecte pas systématiquement les mêmes éléments (bande, case, individu) perturbés selon les phénotypes choisis (poids vifs ou consommation alimentaire cumulée). La différence entre ces différents cas peut être expliquée par le fait que le poids et la CFI un animal ne réagissent probablement pas de la même manière et au même moment face à une perturbation. En effet, les stratégies d'allocation et d'acquisition des ressources peuvent varier en fonction des perturbations et des individus. La Figure 31 montre un exemple où un animal est détecté comme perturbé avec son poids mais pas avec sa CFI. La baisse légère de sa CFI (non détectée) a eu une répercussion forte sur son gain de poids, les ressources ayant été probablement préférentiellement

utilisées pour d'autres fonctions. En revanche, la Figure 32 montre le cas inverse où un animal est détecté comme perturbé avec sa CFI mais pas avec son poids. L'animal a fortement diminué son ingéré en maintenant une croissance relativement convenable probablement au détriment des autres fonctions selon la théorie de l'allocation des ressources (ce dernier a peut-être diminué son activité). Dans ces cas où seul un des phénotypes a répondu à la perturbation de manière significative, nous pouvons quand même indiquer que l'animal était dans un environnement perturbé. Par conséquent, plusieurs nouvelles questions scientifiques se posent :

- Est-ce qu'il est possible de mesurer la robustesse d'un animal d'élevage sur un seul phénotype sachant qu'il a été dans un environnement perturbé ? Cette robustesse serait uniquement spécifique pour ce phénotype et pas pour les autres phénotypes.
- Ou est-ce que la quantification de la robustesse doit absolument s'effectuer simultanément sur plusieurs phénotypes ? Ce qui revient à la question posée au début de ce document (chapitre 'Etat de l'art')

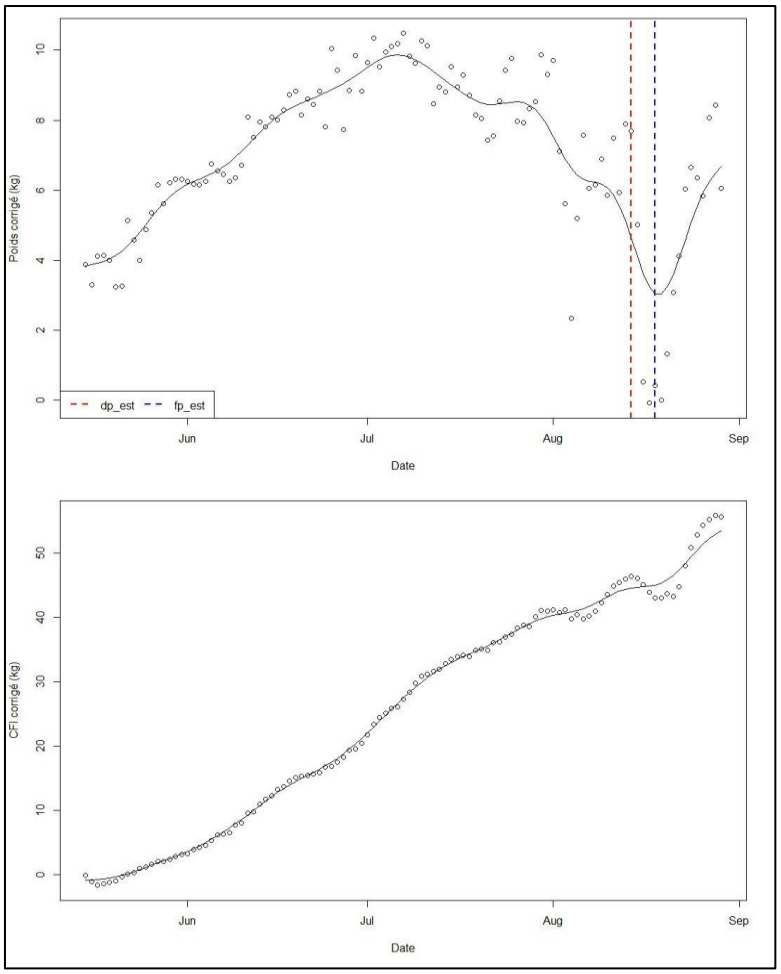


Figure 31 - Exemple d'un animal détecté comme perturbé avec son poids (en haut) vers mi-aout mais pas avec sa CFI (en bas).

Les lignes rouge et bleue représentent respectivement le début et la fin de la perturbation détectée. Les lignes noires représentent les courbes de lissage.

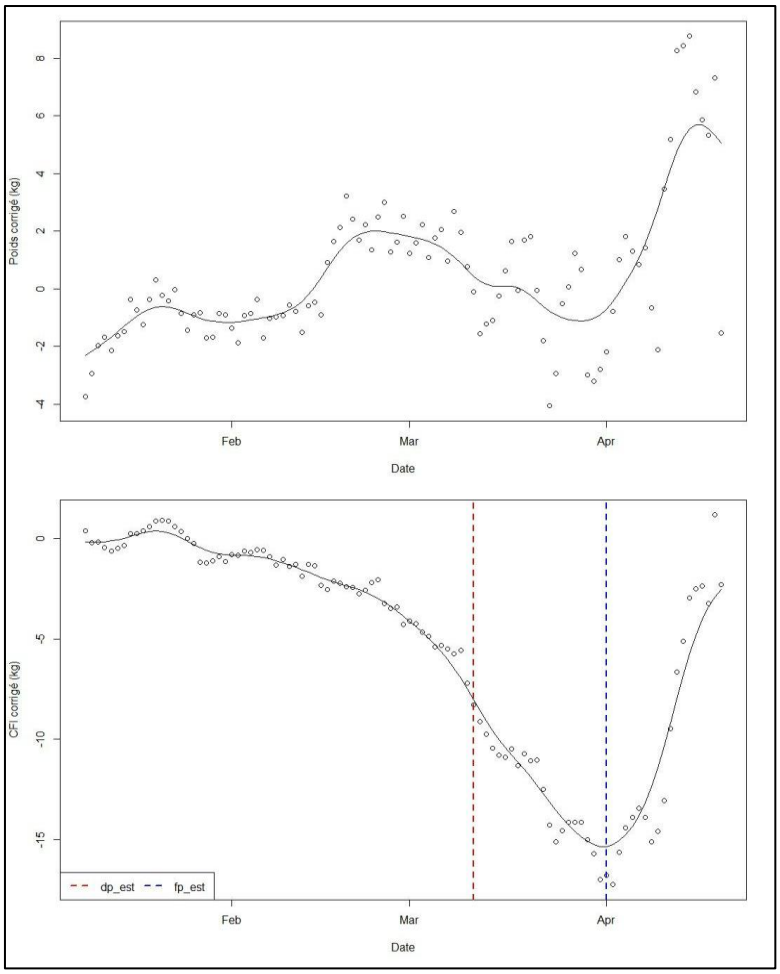


Figure 32 - Exemple d'un animal détecté comme perturbé avec sa CFI (en bas) vers début Mars mais pas avec son poids (en haut).

Les lignes rouge et bleue représentent respectivement le début et la fin de la perturbation détectée. Les lignes noires représentent les courbes de lissage.

Lorsqu'un élément est détecté comme perturbé à la fois avec les deux phénotypes, il y a une ressemblance dans la caractérisation de la perturbation détectée : la médiane des différences en valeur absolue entre les débuts estimés en utilisant le poids et la CFI est inférieure ou égale à 4 pour toutes les échelles. Par conséquent, dans la pratique, il est possible de valider avec sûreté comme perturbés tous les éléments (bandes, cases, individus) qui sont détectés en même temps avec différents phénotypes car il y a une forte chance que l'élément soit réellement perturbé étant donnée la cohérence dans l'estimation du début des perturbations entre les phénotypes d'étude.

Des pistes d'évolution de la méthode Up & Down sont envisageables :

- Dans la partie Down, le seuil de 50% a été choisi arbitrairement comme seuil pour valider les éléments à l'échelle d'un groupe. Ce seuil peut être modifié ce qui aura pour conséquence de modifier la Se et la Sp de la méthode. L'effet de ce dernier pourra être évalué par une courbe ROC (receiver operating characteristic). Le choix du seuil optimal sera alors choisi en collaboration avec les sélectionneurs. Dans le cas de la sélection d'animaux peu sensibles aux perturbations, il est probablement plus raisonnable de maximiser la spécificité de la détection des perturbations, car il serait moins "grave" de ne pas sélectionner tous les animaux robustes que de sélectionner incorrectement des animaux non-robustes.
- Les perturbations ont été simulées de telle manière qu'à chaque échelle (bande, case, individu), une seule perturbation de l'échelle concernée pouvait se produire. Il serait possible de considérer un nombre aléatoire de perturbations par échelle pour approfondir l'étude de la méthode Up & Down.
- Sur les données réelles, plusieurs phénotypes ont été pris en compte pour détecter des perturbations inconnues, cependant les analyses ont été menées séparément pour chaque phénotype. Il est envisageable, de travailler simultanément avec plusieurs phénotypes et de prendre en compte l'interaction entre ces derniers en utilisant un modèle de mélange multivarié dans la partie Down par exemple.
- Enfin, la fonction de lissage utilisée pour la méthode Up & Down était la fonction non paramétrique Nadaraya-Watson avec une taille de fenêtre égale à 10. Ce choix a été fait car c'est la fonction qui donnait les meilleurs résultats (par rapport aux autres fonctions testées) en termes de sensibilité et de spécificité. Cependant, pour d'autres types de phénotypes choisis pour détecter et caractériser des perturbations environnementales inconnues, d'autres choix concernant la taille de la fenêtre de la fonction Nadaraya-Watson ou d'autres fonctions de lissage plus adaptées peuvent être envisagés. Pour ce faire, une analyse graphique des courbes lissées serait importante pour choisir la fonction de lissage la mieux adaptée aux données d'étude. L'idée serait de choisir la fonction qui ne lisse ni trop (risque d'éliminer les informations concernant les perturbations), ni trop peu (risque de laisser trop de bruits et rendre difficile la détection des perturbations).

4.6 Conclusion

Les résultats présentés tout au long de cette étude, tant sur des données simulées que sur des données réelles, ont prouvé la pertinence de la méthode Up & Down pour détecter et caractériser les perturbations. Par utilisation des données simulées, la méthode développée a permis de détecter les perturbations avec une bonne spécificité tout en gardant une sensibilité raisonnable. De plus, la qualité des caractéristiques estimées (début, fin et intensité) des perturbations détectées était correcte (l'écart médian entre les moments perturbés estimés et simulés était inférieur ou égal à 3 jours pour toutes les échelles et la corrélation entre l'intensité estimée et simulée était supérieure ou égale à 0,72 à l'échelle de la bande et de la case). Pour les données réelles, la méthode Up & Down a été appliquée en utilisant deux phénotypes différents (le poids vif et la consommation alimentaire cumulée). Ainsi, il est possible de confirmer « avec certitude » les éléments qui sont détectés comme perturbés à la fois avec les deux phénotypes d'étude car il existe une similitude entre les caractéristiques estimées des éléments détectés à la fois avec les deux phénotypes : l'écart médian entre les moments de référence (débuts des perturbations) selon le poids et la CFI était inférieur à 4 jours pour toutes les échelles. D'autre part, les éléments détectés peuvent être validés avec les données de santé et météorologiques qui fournissent des perturbations potentielles affectant les animaux individuellement ou à l'échelle du groupe : la médiane des écarts en valeur absolue entre les débuts estimés des perturbations détectées et les débuts théoriques est correcte et varie entre 1,5 et 3 jours. Enfin, à partir des résultats pertinents issus de l'application de la méthode Up & Down sur la détection et la caractérisation des perturbations environnementales inconnues, un nouveau critère de la robustesse (résistance et résilience) a été proposé : la pente minimale du phénotype pour la résistance, et la pente entre la fin de la perturbation et la fin de réaction de l'animal du phénotype pour la résilience. Un modèle animal a été appliqué pour estimer les paramètres et VG de ces nouveaux phénotypes sur trois sets de simulations en variant le pourcentage de perturbations de groupe. Les résultats ont montré qu'une augmentation du pourcentage de perturbations simulées avait tendance à entrainer une diminution dans l'estimation des héritabilités et une augmentation de la précision des VG prédites des deux composantes de la robustesse. Ce sujet nécessite des recherches plus approfondies.

4.7 Analyses et résultats complémentaires

4.7.1 Identification du début des perturbations

La méthode Up & Down a été présentée en utilisant le moment où le taux de décroissance est maximal comme indicateur du moment du début de la perturbation. Le moment où le phénotype commence à ralentir avait également été envisagé comme un indicateur du début de la perturbation. Pour déterminer ce moment, tous les maximums locaux du phénotype lissé et corrigé sont identifiés par utilisation de la première dérivée (qui est nulle à ces moments-là, positive avant et négative après) et celui qui précède MR est considéré comme le début de la perturbation. (Figure 33).

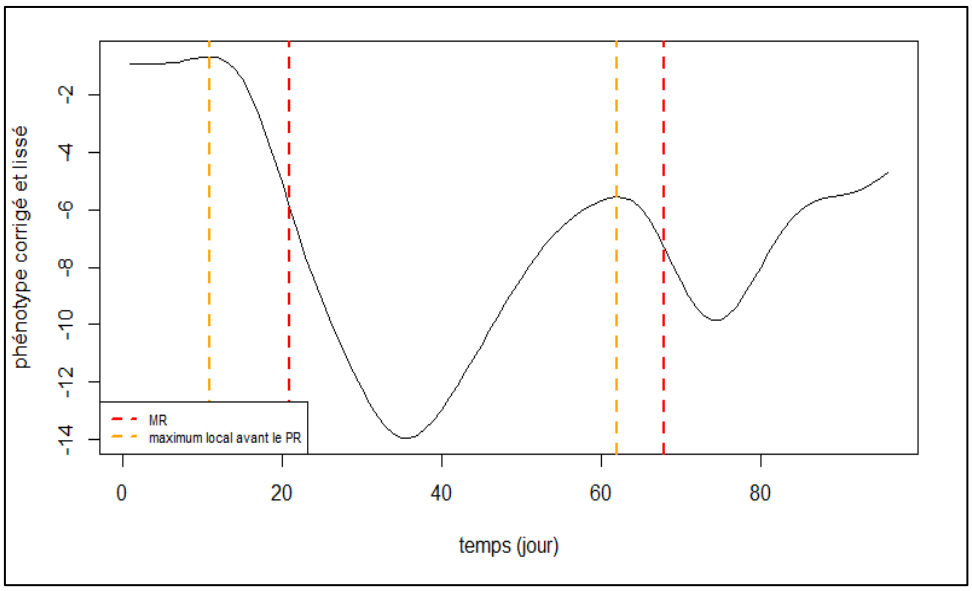


Figure 33 - Identification du moment des maximums locaux du phénotype corrigé et lissé.

Les lignes rouges représentent les débuts utilisés dans cette thèse (ou MR) ; les lignes oranges correspondent aux moments des maximums locaux du phénotype corrigé et lissé (maximums locaux juste avant les MR).

Bien que théoriquement plus judicieux, ce critère n'a pas été conservé pour présenter la méthode Up & Down car ce dernier avait tendance à identifier un début de perturbation trop précoce avec un écart au moment du vrai début plus important que le moment où la pente est minimale (Table 12). Ce décalage est certainement la conséquence du lissage de la courbe (il n'y a pas forcément de maximum local avant la pente de la perturbation). Il pourrait être amélioré en modifiant les paramètres de ce dernier.

Table 12 – La médiane des écarts entre le début simulé et le moment du maximum local du phénotype corrigé et lissé ($edp2$) et le moment du taux de décroissance maximale du phénotype corrigé et lissé ($edp1$) à l'échelle de la bande, case et individu.

Echelle	$edp1$	$edp2$
Bande	3	9
Case	3	8
Individu	3	8

4.7.2 Méthodes de lissages

Durant cette étude, plusieurs méthodes de lissage ont été testées, notamment des polynômes de degrés différents, B-splines (10 nœuds), moyennes mobiles, régression par polynômes locaux, exponentiel ou le lissage non-paramétrique de Nadaraya-Watson (Wood, 2006). Les résultats concernant la qualification de la détection des perturbations (sensibilité et spécificité) pour toutes les méthodes de lissage testées se trouvent dans la Table 13. La fonction Nadaraya Watson présentait les résultats les plus pertinents par rapport aux autres fonctions et a donc été conservée pour la suite de l'étude.

Table 13 - La sensibilité et spécificité en pourcentage pour toutes les méthodes de lissage testées à l'échelle globale, bande, case et individu.

Echelle		s8	s10	s12	s14	s16	p1	p2	p3	bs	ses1	ses2	flo	fnada
Général	Se	85	86	86	85	83	56	69	70	86	68	66	83	85
	Sp	94	95	96	96	97	97	96	93	78	89	79	93	93
Bande	Se	94	92	91	90	88	64	74	78	89	86	80	83	93
	Sp	99	99	99	99	99	99	99	99	97	98	97	99	99
Case	Se	64	66	66	66	66	58	65	58	56	49	40	60	73
	Sp	99	99	99	99	99	88	95	97	95	94	90	98	98
Individu	Se	36	40	42	42	41	35	39	34	40	11	11	39	43
	Sp	97	98	98	99	99	82	90	94	90	96	93	96	95

s8-s16 représentent les fonctions de moyenne mobile de fenêtre 8 à 16 ; p1-3 représentent les polynômes de degré 1 à 3 ; ses1-2 représentent les fonctions de lissage exponentielles de facteur de lissage 0,1 et 0,2 respectivement ; bs, flo, fnada représentent respectivement une fonction bsplines (10 nœuds), régression par polynômes locaux et Nadaraya-Watson.

4.7.3 Statistiques pour la partie Up

En plus de la pente minimale issue des différentes fonctions de lissages les critères utilisés dans la littérature comme indicateurs de la réponse des animaux face aux perturbations ont aussi été testés pour identifier les animaux, cases et bandes perturbés (spécifiquement pour séparer la population en deux sous-populations dans la partie Up). Il s'agissait des critères issus des résidus d'une régression linéaire d'ordre 1 (variance, autocorrélation ou symétrie (Berghof et al., 2018)) et le coefficient de variation de l'ensemble de la population par jour (Garcia-Baccino et al., 2021). Cependant, ces statistiques n'ont pas été pertinentes par rapport à l'utilisation de la pente minimale et par conséquent n'ont pas été utilisées au cours de cette thèse. Les résultats concernant la qualification de la détection des perturbations (sensibilité et spécificité) de ces dernières méthodes sont présentés dans la Table 14**.**

Table 14 – La sensibilité et spécificité en pourcentage à l'échelle de la bande, de la case et de l'individu pour les statistiques complémentaires concernant les résidus et le coefficient de variation.

Echelle		R1	R2	R3	KURT	SYM	AUTO	RMED	CV
Bande	Se	84	85	80	14	17	76	81	26
	Sp	93	93	93	72	79	90	89	69
Case	Se	56	53	47	08	06	55	53	08
	Sp	92	92	92	78	87	93	80	88
Individu	Se	25	20	15	05	00	33	22	00
	Sp	94	94	94	89	94	93	93	94

R1, R2, R3 et RMED représentent la variance des résidus pour polynôme d'ordre 1,2,3 et la médiane du groupe ; KURT, SYM et AUTO représente l'aplatissement, la symétrie et l'autocorrélation des résidus avec un polynôme d'ordre 1 ; CV représente le coefficient de variation. Notons que ce sont les résultats pour seulement la partie Up de la méthode Up & Down.

Conclusion générale et perspectives

Le projet de thèse s'inscrivait dans une action menée conjointement par INRAE et Alliance R&D, financeurs de la thèse. L'objectif principal de cette thèse était de proposer un nouveau critère pertinent de la robustesse « spécifique » des animaux d'élevage, en lien avec la réponse à travers le temps des animaux face à une ou plusieurs perturbations environnementales. Pour atteindre cet objectif, une méthode basant sur l'analyse des données de phénotypage haut-débit a été développée afin de détecter et caractériser les perturbations inconnues affectant les animaux d'élevage, et de proposer une nouvelle mesure de la robustesse.

Dans la première partie de la thèse, nous avons mesuré l'impact des perturbations de faible intensité et de courte durée sur l'estimation des paramètres et valeurs génétiques des animaux d'élevage en utilisant des données simulées de porcs en croissance. Les résultats suggéraient que la sélection dans un environnement bien contrôlé avec des perturbations faibles et aléatoires ne permettait pas d'améliorer la robustesse des animaux alors qu'elle était moins efficace pour améliorer les caractères de production que la sélection sans perturbations environnementales. Il est donc désirable de séparer la production et la robustesse pour améliorer la sélection animale. Ce travail a permis de comprendre la composition des valeurs génétiques estimées des phénotypes d'intérêt dans un environnement bien contrôlé et de montrer l'importance de séparer la robustesse et la production afin de pouvoir sélectionner simultanément les animaux pour ces deux fonctions. Il faut pour cela être capable de mesurer la robustesse d'une part et d'estimer sans biais les paramètres génétiques de production d'autre part.

Pour mesurer la robustesse, nous avons proposé la méthode Up & Down qui consiste à caractériser les perturbations puis quantifier la robustesse des animaux sachant ces informations sur les perturbations. Cette dernière a été validée sur données simulées puis appliquée sur les données de porcs en croissance fournies par l'Alliance R&D. La méthode a donné de bons résultats en termes de sensibilité et spécificité de détection des perturbations sur simulation. Appliquée aux données réelles, nous avons montré qu'il pouvait être pertinent d'utiliser les résultats de la méthode sur plusieurs phénotypes pour identifier les

perturbations. Un package R de la méthode Up & Down est en cours de développement pour une diffusion en Open Access.

Les nouveaux phénotypes de robustesse proposés sont la pente minimale du phénotype corrigé et lissé traduisant la résistance de l'animal et la pente du phénotype corrigé et lissé entre la fin de la perturbation et la fin de réaction traduisant la résilience de l'animal. Nous avons montré que la précision des VG de ces composantes de la robustesse variait en fonction des conditions environnementales, les précisions étant plus importantes lorsque le milieu est plus fortement perturbé. Pour juger de l'intérêt de ces nouveaux phénotypes, il sera nécessaire de les estimer sur les données réelles et d'en effectuer l'analyse génétique afin de pouvoir à terme les introduire dans les objectifs de sélection. Pour ce faire, il est nécessaire tout d'abord de connaitre la valeur économique de ce phénotype. Par exemple, Poppe *et al.* (2020) ont montré en bovins laitiers que les animaux résilients permettaient une augmentation du bénéfice de l'éleveur par amélioration de la longévité, de la production et par réduction des coûts d'élevage. De plus, des analyses complémentaires devraient être réalisées afin d'ajuster au mieux la pondération pour atteindre un progrès génétique pour tous les caractères d'intérêt. Une des stratégies est d'effectuer des simulations sur plusieurs générations d'animaux en testant plusieurs sets de paramètres (concernant la population, les types de perturbations, etc.) pour calculer le gain cumulé selon le poids attribué à la robustesse et identifier le meilleur compromis entre la robustesse et les autres caractères d'intérêt (Ghaderi-Zefreh et al., 2022).

Concernant l'évaluation génétique des animaux d'élevage, dans cette étude, nous avons utilisé des données uniques (phénotype mesuré à un moment donné) pour estimer les paramètres et valeurs génétiques des caractères de production des animaux (poids à 100 jours d'âge et le gain moyen quotidien). Nous pouvons envisager de faire la même chose en utilisant des données longitudinales. Ainsi, une amélioration de l'estimation génétique des phénotypes d'intérêt en utilisant des informations supplémentaires des perturbations détectées pourrait avoir lieu. Cela implique de réévaluer les paramètres et les valeurs génétiques des animaux en utilisant des courbes tronquées au moment du début de la première perturbation rencontrée par l'animal. En effet, même si les observations ne sont pas complètes, il est possible de fournir des prédictions stables des phénotypes futurs grâce à un modèle de régression

aléatoire (David et al., 2015) à partir des courbes tronquée. Le bénéfice d'une sélection à partir des phénotypes tronqués pourra alors être évalué par simulation.

Références

Bahlo C, Dahlhaus P, Thompson H and Trotter M 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. Computers and Electronics in Agriculture 156, 459–466.

Benaglia T, Chauveau D, Hunter DR and Young DS 2010. mixtools: An R Package for Analyzing Mixture Models. Journal of Statistical Software 32, 1–29.

Berckmans D 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. Revue Scientifique Et Technique (International Office of Epizootics) 33, 189–196.

Berghof TVL, Poppe M and Mulder HA 2018. Opportunities to Improve Resilience in Animal Breeding Programs. Frontiers in Genetics 9, 692.

Bidanel JP, Silalahi P, Tribout T, Canario LL, Ducos A, Garreau HH, Gilbert H, Larzul C, Milan D, Riquet J, Schwob SS, Mercat M-JM-J, Hassenfratz C, Bouquet A, Bazin C and Bidanel J 2020. Cinquante années d'amélioration génétique du porc en France : bilan et perspectives. INRA Productions Animales 33, 1.

Bishop S 2012. A consideration of resistance and tolerance for ruminant nematode infections. Frontiers in Genetics 3, 168.

Bishop SC, Fleming RH, McCormack HA, Flock DK and Whitehead CC 2000. Inheritance of bone characteristics affecting osteoporosis in laying hens. British Poultry Science 41, 33–40.

Blanc F, Ollion E, Puillet L, Delaby L, Ingrand S, Tichit M and Friggens NNC 2013. Evaluation quantitative de la robustesse des animaux et du troupeau : quels principes retenir ? In 20. Rencontres autour des Recherches sur les Ruminants, Rencontres autour des Recherches sur les Ruminants, pp. 265–272. INRA. Institut de l'Elevage, Paris, France.

Bodin L, Bolet G, Garcia M, Garreau H, Larzul C and David I 2010. Robustesse et canalisation : vision de généticiens. Productions Animales 23.

Bytyqi H, Ødegård J, Mehmeti H, Vegara M and Klemetsdal G 2007. Environmental Sensitivity of Milk Production in Extensive Environments: A Comparison of Simmental, Brown Swiss, and Tyrol Grey Using Random Regression Models. Journal of Dairy Science 90, 3883–3888.

Clapperton M, Diack AB, Matika O, Glass EJ, Gladney CD, Mellencamp MA, Hoste A and Bishop SC 2009. Traits associated with innate and adaptive immunity in pigs: heritability and associations with performance under different health status conditions. Genetics, selection, evolution: GSE 41, 54.

Clément A, Marmier F, Kamissoko D, Gourc D, Wioland L, Govaere V and Cegarra J 2018. Robustesse, résilience : une brève synthèse des définitions au travers d'une analyse structurée de la littérature. In MOSIM'18 - 12ème Conférence internationale de Modélisation, Optimisation et SIMulation, p. 8. ISAE and IMT Mines Albi, Toulouse, France.

Codrea MC, Højsgaard S and Friggens NC 2011. Differential smoothing of time-series measurements to identify disturbances in performance and quantify animal response characteristics: An example using milk yield profiles in dairy cows. Journal of Animal Science 89, 3089–3098.

Colditz IG, Hine BC, Colditz IG and Hine BC 2016. Resilience in farm animals: biology, management, breeding and implications for animal welfare. Animal Production Science 56, 1961–1983.

David I, Ruesche J, Drouilhet L, Garreau H and Gilbert H 2015. Genetic modeling of feed intake. Journal of Animal Science 93, 965–977.

Dekkers, P. K. Mathur, and E. F. Knol 2011. Genetic Improvement of the Pig.

Dekkers J, Rowland RRR, Lunney JK and Plastow G 2017. Host genetics of response to porcine reproductive and respiratory syndrome in nursery pigs. Veterinary Microbiology 209, 107–113.

Dragicevic AZ 2016. From robustness to resilience: A network Price identity approach. Ecological Complexity 28, 47–53.

Ducos A, Douhard F, Savietto D, Sautier M, FILLON V, Gunia M, Rupp R, Moreno-Romieux C, Mignon-Grasteau S, Gilbert H and Fortun-Lamothe L 2021. Contributions de la génétique animale à la transition agroécologique des systèmes d'élevage. INRAE Productions Animales 34, 79–96.

Dumont B, Fortun-Lamothe L, Jouven M, Thomas M and Tichit M 2013. Prospects from agroecology and industrial ecology for animal production in the 21st century. Animal 7, 1028–1043.

Engblom L, Stalder KJ, Mabry JW and Schwab CR 2010. Lifetime Reproductive Traits in Landrace, Yorkshire and Crossbred Sows. Iowa State University Animal Industry Report 7.

Foury A, Devillers N, Sanchez M-P, Griffon H, Le Roy P and Mormède P 2005. Stress hormones, carcass composition and meat quality in Large White×Duroc pigs. Meat Science 69, 703–707.

Foury A, Geverink NA, Gil M, Gispert M, Hortós M, Furnols MF i, Carrion D, Blott SC, Plastow GS and Mormède P 2007. Stress neuroendocrine profiles in five pig breeding lines and the relationship with carcass composition. animal 1, 973–982.

Fraser D, Weary D, Pajor E and Milligan B 1997. A Scientific Conception of Animal Welfare that Reflects Ethical Concerns. Ethics and Animal Welfare Collection.

Frey M, Hofer A and Künzi N 1997. Comparison of models with a fixed or a random contemporary group effect for the genetic evaluation for litter size in pigs. Livestock Production Science 48, 135–141.

Friggens NC, Adriaens I, Boré R, Cozzi G, Jurquet J, Kamphuis C, Leiber F, Lora I, Sakowski T, Statham J and De Haas Y 2022. Resilience: reference measures based on longer-term consequences are needed to unlock the potential of precision livestock farming technologies for quantifying this trait. Peer Community Journal 2.

Friggens NC, Blanc F, Berry DP and Puillet L 2017. Review: Deciphering animal robustness. A synthesis to facilitate its use in livestock breeding and management. animal 11, 2237–2251.

Friggens NC, Sauvant D and Martin O 2010. Vers des définitions opérationnelles de la robustesse s'appuyant sur des faits biologiques. L'exemple de la nutrition. INRAE Productions Animales 23, 43–52.

Garcia-Baccino CA, Marie-Etancelin C, Tortereau F, Marcon D, Weisbecker J-L and Legarra A 2021. Detection of unrecorded environmental challenges in high-frequency recorded traits, and genetic determinism of resilience to challenge, with an application on feed intake in lambs. Genetics Selection Evolution 53, 4.

Garreau H, Bolet G, Larzul C, Robert-Granié C, Saleil G, SanCristobal M and Bodin L 2008. Results of four generations of a canalising selection for rabbit birth weight. Livestock Science 119, 55–62.

Ghaderi-Zefreh, Doeschl-Wilson, Riggio, Matika, and Pong-Wong 2022. Selection for robustness: exploring the value of genomic prediction, reaction norm models and phenotyping strategies.

Gilmour AR, R.B.J. G, Cullis BR and Thompson R 2009. Asreml User Guide Release 3.0. VSN International Ltd, Hemel Hemptead, HP1 1ES, UK.

de Goede DM, Gremmen B and Blom-Zandstra M 2013. Robust agriculture: Balancing between vulnerability and stability. NJAS - Wageningen Journal of Life Sciences 64–65, 1–7.

Gross WB 1976. Plasma Steroid Tendency, Social Environment and Eimeria necatrix infection. Poultry Science 55, 1508–1512.

Gunia M, David I, Hurtaud J, Maupin M, Gilbert H and Garreau H 2018. Genetic Parameters for Resistance to Non-specific Diseases and Production Traits Measured in Challenging and Selection Environments; Application to a Rabbit Case. Frontiers in Genetics 9.

Heckendorn F, Bieber A, Werne S, Saratsis A, Maurer V and Stricker C 2017. The genetic basis for the selection of dairy goats with enhanced resistance to gastrointestinal nematodes. Parasite (Paris, France) 24, 32.

Henderson CR 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model. Biometrics 31, 423–447.

Henderson CR 1982. Analysis of Covariance in the Mixed Model: Higher-Level, Nonhomogeneous, and Random Regressions. Biometrics 38, 623–640.

Hennessy DP, Stelmasiak T, Johnston NE, Jackson PN and Outch KH 1988. Consistent capacity for adrenocortical response to ACTH administration in pigs. American journal of veterinary research 49, 1276–1283.

Hermesch S, Li L, Doeschl-Wilson AB and Gilbert H 2015. Selection for productivity and robustness traits in pigs. Animal Production Science 55, 1437.

Hess AS, Islam Z, Hess MK, Rowland RRR, Lunney JK, Doeschl-Wilson A, Plastow GS and Dekkers JCM 2016. Comparison of host genetic factors influencing pig response to infection with two North American isolates of porcine reproductive and respiratory syndrome virus. Genetics, selection, evolution 48, 43.

Homma C, Hirose K, Ito T, Kamikawa M, Toma S, Nikaido S, Satoh M and Uemoto Y 2021. Estimation of genetic parameter for feed efficiency and resilience traits in three pig breeds. Animal 15, 100384.

IFIP 2013. Memento de l'éleveur de porc.

Iung LH de S, Carvalheiro R, Neves HH de R and Mulder HA 2020. Genetics and genomics of uniformity and resilience in livestock and aquaculture species: A review. Journal of Animal Breeding and Genetics 137, 263–280.

Kanis E, De Greef KH, Hiemstra A and van Arendonk JAM 2005. Breeding for societally important traits in pigs1. Journal of Animal Science 83, 948–957.

Khanh PCP, Dinh Chinh N, Cham TT, Vui PT and Tan TD 2016. Classification of cow behavior using 3-DOF accelerometer and decision tree algorithm. In 2016 International Conference on Biomedical Engineering (BME-HUST), pp. 45–50.

Knap PW 2005. Breeding robust pigs. Australian Journal of Experimental Agriculture 45, 763–773.

Knott SA, Cummins LJ, Dunshea FR and Leury BJ 2008. Rams with poor feed efficiency are highly responsive to an exogenous adrenocorticotropin hormone (ACTH) challenge. Domestic Animal Endocrinology 34, 261–268.

Köhn F, Sharifi AR, Malovrh Š and Simianer H 2007. Estimation of genetic parameters for body weight of the Goettingen minipig with random regression models1. Journal of Animal Science 85, 2423–2428.

Kolmodin R, Strandberg E, Madsen P, Jensen J and Jorjani H 2002. Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. Acta Agriculturae Scandinavica, Section A — Animal Science 52, 11–24.

Kristensen HH and Cornou C 2011. Automatic detection of deviations in activity levels in groups of broiler chickens – A pilot study. Biosystems Engineering 109, 369–376.

Laghouaouta H, Pena RN, Ros-Freixedes R, Reixach J, Díaz M, Estany J, Armengol R, Bassols A and Fraile L 2021. A Methodology to Quantify Resilience in Growing Pigs. Animals 11, 2970.

Lawrence AB and Wall E 2014. Selection for 'environmental fit' from existing domesticated species.

Leenhouwers JI, Knol EF, de Groot PN, Vos H and van der Lende T 2002. Fetal development in the pig in relation to genetic merit for piglet survival1. Journal of Animal Science 80, 1759–1770.

Lenoir G, Flatres-Grall L, Friggens NC and David I 2022. Robustness scores in fattening pigs based on routinely collected phenotypes: determination and genetic parameters. Journal of Animal Science 100, skac157.

Martin P, Raoul J and Bodin L 2014. Effects of the FecL major gene in the Lacaune meat sheep population. Genetics Selection Evolution 46, 48.

Mazé-Guilmo E, Loot G, Páez DJ, Lefèvre T and Blanchet S 2014. Heritable variation in host tolerance and resistance inferred from a wild host-parasite system. Proceedings. Biological Sciences 281, 20132567.

McLachlan GJ, Lee SX and Rathnayake SI 2019. Finite Mixture Models. Annual Review of Statistics and Its Application 6, 355–378.

Mengistu SB, Mulder HA, Bastiaansen JWM, Benzie JAH, Khaw HL, Trinh TQ and Komen H 2022. Fluctuations in growth are heritable and a potential indicator of resilience in Nile tilapia (Oreochromis niloticus). Aquaculture 560, 738481.

Meyer K 2009. Factor-analytic models for genotype × environment type problems and structured covariance matrices. Genetics Selection Evolution 41, 21.

Michel V, Peinnequin A, Alonso A, Buguet A, Cespuglio R and Canini F 2007. Effect of glucocorticoid depletion on heat-induced Hsp70, IL-1β and TNF-α gene expression. Brain Research 1164, 63–71.

Mormède P, Foury A, Terenina E and Knap PW 2011. Breeding for robustness: the role of cortisol. Animal: An International Journal of Animal Bioscience 5, 651–657.

Nguyen-Ba H, Milgen J van and Taghipoor M 2020. A procedure to quantify the feed intake response of growing pigs to perturbations. animal 14, 253–260.

Nordenfelt LY 1995. On the Nature of Health: An Action-Theoretic Approach. Springer Science & Business Media.

Phocas F, Belloc C, Bidanel J, Delaby L, Dourmad JY, Dumont B, Ezanno P, Fortun-Lamothe L, Foucras G, Frappat B, González-García E, Hazard D, Larzul C, Lubac S, Mignon-Grasteau S, Moreno CR, Tixier-Boichard M and Brochard M 2016. Review: Towards the agroecological management of ruminants, pigs and poultry through the development of sustainable breeding programmes: I-selection goals and criteria. Animal: An International Journal of Animal Bioscience 10, 1749–1759.

Phocas F, Bobe J, Bodin L, Charley B, Dourmad J-Y, Friggens N, Hocquette J-F, Bail P-YL, Duval E, Mormède P, Quéré P and Schelcher F 2014. Des animaux plus robustes : un enjeu majeur pour le développement durable des productions animales nécessitant l'essor du phénotypage fin et à haut débit. INRA Productions Animales 27, 181.

Phocas F, Brochard M, Larroque H, Laggrifoul G, Labatut J and Guerrier J 2013. État actuel et perspectives d'évolution des objectifs de sélection chez les ruminants. p. 4 p. Institut de l'Elevage.

Poppe M, Veerkamp RF, van Pelt ML and Mulder HA 2020. Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. Journal of Dairy Science 103, 1667–1684.

Putz AM, Harding JCS, Dyck MK, Fortin F, Plastow GS, Dekkers JCM and PigGen Canada 2019. Novel Resilience Phenotypes Using Feed Intake Data From a Natural Disease Challenge Model in Wean-to-Finish Pigs. Frontiers in Genetics 9, 660.

Rauw WM and Gomez-Raya L 2015. Genotype by environment interaction and breeding for robustness in livestock. Frontiers in Genetics 6, 310.

Revilla M, Lenoir G, Flatres-Grall L, Muñoz-Tamayo R and Friggens NC 2021. Quantifying growth perturbations over the fattening period in swine via mathematical modelling.

Robson DS 1959. A Simple Method for Constructing Orthogonal Polynomials When the Independent Variable is Unequally Spaced. Biometrics 15, 187–191.

Rodenburg TB, Buitenhuis AJ, Ask B, Uitdehaag KA, Koene P, van der Poel JJ, van Arendonk JAM and Bovenhuis H 2004. Genetic and Phenotypic Correlations Between Feather Pecking and Open-Field Response in Laying Hens at Two Different Ages. Behavior Genetics 34, 407–415.

Ros M, Sorensen D, Waagepetersen R, Dupont-Nivet M, SanCristobal M, Bonnet J-C and Mallard J 2004. Evidence for Genetic Control of Adult Weight Plasticity in the Snail Helix aspersa. Genetics 168, 2089–2097.

Roy PL, Ducos A and Phocas F 2019. Quelles performances pour les animaux de demain ? Objectifs et méthodes de sélection. INRAE Productions Animales 32, 233–246.

Rutgers B and Heeger R 1999. Inherent worth and respect for animal integrity. Van Gorcum.

Sánchez-Molano E, Kapsona VV, Ilska JJ, Desire S, Conington J, Mucha S and Banos G 2019. Genetic analysis of novel phenotypes for farm animal resilience to weather variability. BMC Genetics 20, 84.

Sauvant D and Martin O 2010. Robustesse, rusticité, flexibilité, plasticité... les nouveaux critères de qualité des animaux et des systèmes d'élevage : définitions systémique et biologique des différents concepts. INRAE Productions Animales 23, 5–10.

Schaeffer LR 2004. Application of random regression models in animal breeding. Livestock Production Science 86, 35–45.

Scheffer M, Carpenter SR, Dakos V and van Nes EH 2015. Generic Indicators of Ecological Resilience: Inferring the Chance of a Critical Transition. Annual Review of Ecology, Evolution, and Systematics 46, 145–167.

Shariati MM, Korsgaard IR and Sorensen D 2009. Identifiability of parameters and behaviour of MCMC chains: a case study using the reaction norm model. Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie 126, 92–102.

Siwek M, Buitenhuis B, Cornelissen S, Nieuwland M, Knol EF, Crooijmans R, Groenen M, Parmentier H and van der Poel J 2006. Detection of QTL for innate: Non-specific antibody levels binding LPS and LTA in two independent populations of laying hens. Developmental & Comparative Immunology 30, 659–666.

Sorensen D and Waagepetersen R 2003. Normal linear models with genetically structured residual variance heterogeneity: a case study. Genetics Research 82, 207–222.

Star L, Ellen E, Uitdehaag K and Brom F 2008. A plea to implement robustness into a breeding goal: Poultry as an example. Journal of Agricultural and Environmental Ethics 21, 109–125.

Tortereau F, Marie-Etancelin C, Weisbecker J-L, Marcon D, Bouvier F, Moreno-Romieux C and François D 2020. Genetic parameters for feed efficiency in Romane rams and responses to single-generation selection. animal 14, 681–687.

Tsybakov A 2009. Introduction to Nonparametric Estimation.

Van Gameren V, Weikmans R and Zaccaï E 2014. L'adaptation au changement climatique. La Découverte.

Van Vleck LD 1987. Contemporary groups for genetic evaluations. Journal of Dairy Science 70, 2456–2464.

Veerkamp RF, Mulder HA, Calus MPL, Windig JJ and Napel J ten 2009. Statistical genetics to improve robustness of dairy cows. In Proceedings Association for the Advancement of Animal Breeding and Genetics, Barossa Valley, Australia, September 28 - October 1, 2009, pp. 406–413.

Verbeke WAJ and Viaene J 2000. Ethical Challenges for Livestock Production:Meeting Consumer Concerns about Meat Safety and AnimalWelfare. Journal of Agricultural and Environmental Ethics 12, 141–151.

Waddington CH 1960. Experiments on canalizing selection*. Genetics Research 1, 140–150.

Wilmink JBM 1987. Comparison of different methods of predicting 305-day milk yield using means calculated from within-herd lactation curves. Livestock Production Science 17, 1–17.

Wood SN 2006. Generalized Additive Models: an introduction with R. 397.

Wybo J-L 2008. The Role of Simulation Exercises in the Assessment of Robustness and Resilience of Private or Public Organizations. In Resilience of Cities to Terrorist and other Threats (eds. H.J. Pasman and I.A. Kirillov), NATO Science for Peace and Security Series Series C: Environmental Security, pp. 491–507. Springer Netherlands, Dordrecht.

Zumbach B, Misztal I, Tsuruta S, Sanchez JP, Azain M, Herring W, Holl J, Long T and Culbertson M 2008. Genetic components of heat stress in finishing pigs: Parameter estimation. Journal of animal science 86, 2076–81.

Formations

Catégorie : Éthique et intégrité scientifique

- Ethique et Recherche (28 novembre 2019 - 28 novembre 2019) - Toulouse (8 heures)

Catégorie : Méthodologie de la thèse recherche documentaire et publication

- Papirus (15 janvier 2021 - 12 mars 2021) INRAE - Toulouse

Catégorie : Scientifique

- Data Science pour l'agriculture (17 mai 2021 - 21 mai 2021) acta – distanciel (20 heures)
- Evaluating Animal Breeding Programmes (06 novembre 2019) Wageningen University - Pays-Bas (25 heures)
- Genetic Models for Animal Breeding (06 novembre 2019) Wageningen University - Pays-Bas (25 heures)
- Linux (12 octobre 2020) INRAE - Toulouse

Valorisations

Publications scientifiques

Vincent Le, Tom Rohmer, Ingrid David. Impact of environmental disturbances on estimated genetic parameters and breeding values for growth traits in pigs. Journal Animal, 2022, ⟨10.1016/j.animal.2022.100496⟩

Article en cours d'écriture

Vincent Le, Tom Rohmer, Ingrid David. Identifying and characterizing disturbances from high-throughput phenotyping data.

Communications à des congrès

Vincent Le, Tom Rohmer, Ingrid David. Identifying and characterizing disturbances from high-throughput phenotyping data. EAAP, Aug 2021, Davos, Switzerland. (Poster)

Vincent Le, Tom Rohmer, Florence Ytournel, Loïc Flatres-Grall, Bruno Ligonesche, Ingrid David. Evaluation de l'impact des perturbations sur l'estimation des paramètres et la prédiction des valeurs génétiques. JRP, Feb 2021, Paris, France. (Poster)

Résumé

Dans un contexte de changement climatique, de transition agro-écologique et de nouvelles demandes de la société pour le bien-être animal, les animaux de demain seront élevés dans des milieux suboptimaux ou changeants auxquels ils devront pouvoir s'adapter. Il est donc nécessaire de sélectionner des animaux robustes. La robustesse est un caractère complexe et difficile à mesurer. Le but de cette thèse était de proposer un nouveau critère de robustesse pour la sélection par utilisation des données de phénotypage haut-débit. Nous avons dans un premier temps démontré, par simulation, que la sélection animale à partir des phénotypes de production enregistrés dans un environnement fortement contrôlé tel que rencontré en élevage de sélection porcine, ne permettait pas une sélection conjointe sur la production et la robustesse. En effet, en simulant une sélection sur le poids à 100j de contrôle en engraissement, nous avons comparé les paramètres et **valeurs génétiques (VG)** pour ce phénotype dans deux situations : en l'absence ou en présence de perturbations. Bien que l'héritabilité du caractère ait diminué en présence de perturbations (de l'ordre de 28 %), les VG prédites ne correspondaient pas à une combinaison du potentiel génétique de production et de robustesse (les corrélations partielles entre ces dernières et le vrai potentiel de robustesse n'étant pas significativement différentes de 0). Il était donc nécessaire de proposer une méthode de mesure de la robustesse qui puisse être appliquée en élevage de sélection porcine. Nous avons développé pour cela la méthode Up & Down qui consiste, en étudiant la dynamique d'évolution d'un phénotype au cours du temps à des échelles différentes : individuelle ou groupe d'animaux (case et bande), à identifier et caractériser les perturbations à ces différentes échelles puis mesurer les réponses individuelles face à ces perturbations, c.-à-d. quantifier la robustesse. Cette méthode a été validée dans un premier temps par simulation : la sensibilité de détection de perturbation variait selon l'échelle sur laquelle cette dernière intervenait (93%, 73% et 43% à l'échelle de la bande, la case et individuelle, respectivement) tout en conservant une bonne spécificité (supérieur à 95% à toutes les échelles). Les jours de début et fin des perturbations étaient correctement estimés (écart en valeur absolue médian de 2 à 3 jours entre les estimations et les vraies valeurs selon l'échelle) tout comme l'intensité de la perturbation (corrélation entre l'intensité vraie et sa valeur prédite de 0,78, 0,72, 0,50 à l'échelle de la bande, la case et individuelle respectivement). La méthode Up & Down a ensuite été appliquée aux données de consommation cumulée et poids

de 6298 porcs en croissance de la station de phénotypage du Rheu. Les perturbations identifiées ont été confirmées par les données météorologiques et de santé enregistrées dans l'élevage : les moments estimés des perturbations détectées correspondaient relativement bien aux moments où les conditions météorologiques étaient délétères ou des moments où des problèmes de santé avaient été identifiés dans l'élevage (les écarts variaient entre 1,5 et 3 jours selon l'échelle). Enfin, un nouveau critère des composantes de la robustesse (résistance et résilience) a été proposé : la pente minimale du phénotype pour la résistance, et la pente du phénotype entre la fin de la perturbation et la fin de réaction de l'animal pour la résilience. Un modèle animal a été appliqué pour estimer les paramètres et VG de ces nouveaux phénotypes sur trois sets de simulations en variant le pourcentage de perturbations de groupe. A partir de ces nouveaux caractères, les résultats ont montré qu'une augmentation du pourcentage de perturbations simulées avait tendance à entrainer une diminution des héritabilités et une augmentation de la précision des VG prédites des deux composantes de la robustesse.

Abstract

In a context of climate change, agro-ecological transition and new societal demands for animal welfare, tomorrow's animals will be raised in suboptimal or changing environments to which they will need to be able to adapt. In this context, it is necessary to select robust animals (i.e. not sensitive to environmental variations). Robustness is a complex trait that is difficult to measure. The goal of this thesis was to propose a new robustness criterion for selection using high throughput phenotyping data (data measured repeatedly over time). We first demonstrated, by simulation, that animal selection based on production phenotypes recorded in a highly controlled environment as encountered in pig breeding, did not allow a joint selection on production and robustness. By simulating a selection on the weight at 100 d of control in fattening, we compared the genetic parameters and breeding values for this phenotype in two situations: in the absence or presence of environmental disturbances. Although the heritability of the trait decreased in the presence of disturbances (about 28%), the EBV did not correspond to a combination of genetic potential for production and robustness (the partial correlations between the latter and the true robustness potential were not significantly different from 0). It was therefore necessary to propose a method to measure robustness that could be applied in pig breeding (i.e. when pigs are raised in pens and batches). We have developed the Up & Down method which consists, by studying the dynamics of a phenotype over time at different scales: individual or group of animals (pen and batch), in identifying and characterizing the disturbances at these different scales and then measuring the individual responses to these disturbances, i.e. quantifying the robustness. This method was first validated by simulation: the sensitivity of disturbance detection varied according to the scale on which the disturbance occurred (93%, 73% and 43% at the batch, pen and individual scales respectively) while maintaining good specificity (greater than 95% at all levels). The beginning and end of the disturbances were correctly estimated (median absolute value difference of 2 to 3 days between the estimates and the true values depending on the scale) as was the intensity of the disturbance (correlation of 0,78, 0,72, 0,50 at the batch, pen and individual scales respectively). The Up & Down method was then applied to the cumulative feed intake and weight data of 6298 growing pigs from the phenotyping station of Le Rheu. The identified disturbances were confirmed by the weather and health data recorded in the farm: the estimated times of the detected disturbances corresponded

relatively well to the times of the recorded disturbances (the differences varied between 1,5 and 3 days depending on the scale). Finally, a new criterion for the components (resistance and resilience) of robustness was proposed: the minimum slope of the phenotype for resistance, and the slope between the end of the disturbance and the end of the animal's response for resilience. An animal model was applied to estimate the genetic parameters and breeding values of these new phenotypes on three sets of simulations by varying the percentage of group disturbances. From these new traits, the results showed that an increase in the percentage of simulated disturbances tended to lead to a decrease in the estimation of heritabilities and an increase in the EBV accuracy of both components of robustness.