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 :

Agrosystèmes, Écosystèmes et Environnement

Présentée et soutenue par :

Mme MARIE-ANNE-DOMINIQUE VEDY-ZECCHINI

le jeudi 9 janvier 2020

Titre :

Analyse et modélisation des effets du système de culture et de la situation de production sur les principaux bioagresseurs du tournesol

---

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

AGroécologie, Innovations, TeRritoires (AGIR)

Directeur(s) de Thèse :

M. JEAN-NOEL AUBERTOT

Rapporteurs :

M. DIDIER ANDRIVON, INRA RENNES

M. MARC BARDIN, INRA AVIGNON

Mme MURIEL MORISON, INRA VERSAILLES GRIGNON

Membre(s) du jury :

M. FREDERICK GARCIA, INRA TOULOUSE, Président

M. JEAN-NOEL AUBERTOT, INRA TOULOUSE, Membre

M. JEAN-PHILIPPE DEGUINE, CIRAD SAINT-DENIS, Membre

M. XAVIER PINOCHET, TERRES INOVIA, Membre

AUTEURE : Marie-Anne VEDY-ZECCHINI
DIRECTEUR DE THESE : Jean-Noël AUBERTOT
CO-ENCADRANTE DE THESE : Marie-Hélène ROBIN
LIEU ET DATE DE LA SOUTENANCE : Centre INRAE, Auzeville, le 9 janvier 2020
TITRE : Analyse et modélisation des effets du système de culture et de la situation de production sur les principaux bioagresseurs du tournesol

---

RÉSUMÉ. Le tournesol présente de nombreux atouts agroécologiques (résistance aux stress abiotiques, faible Indice de Fréquence de Traitement, bon précédent cultural, entretien de la biodiversité des pollinisateurs), mais ses surfaces et son rendement stagnent depuis près de 30 ans. Ces stagnations sont notamment dues à un cortège de bioagresseurs dont il est la cible. Dans un contexte d'objectif de réduction de l'utilisation des pesticides, les méthodes non chimiques de prophylaxie (ou préventives) doivent être valorisées face aux bioagresseurs. Or, il est difficile de prévoir l'effet du changement d'une ou plusieurs pratiques culturales sur l'ensemble de bioagresseurs. Les outils de modélisation sont des outils indispensables pour aborder cette question. Trois bioagresseurs ont été considérés lors de cette thèse : deux champignons phytopathogènes, le phomopsis (*Diaporthe helianthi*) et le phoma (*Leptosphaeria lindquistii*), en raison des connaissances et des jeux de données disponibles ; et une plante parasitant exclusivement le tournesol, l'orobanche (*Orobanche cumana*) en raison de l'importance agronomique des dommages qu'elle entraîne. Elle est récemment apparue en France et prolifère rapidement. Ces trois bioagresseurs peuvent avoir un haut niveau de nuisibilité. Le travail de thèse aborde trois questions. i) Comment gérer la production d'inoculum primaire de ces bioagresseurs ? ii) Comment synthétiser les savoirs disponibles sur les effets des pratiques agricoles et de la situation de production sur les cycles biologiques de ces bioagresseurs ? iii) Comment améliorer les méthodes de modélisation mobilisées ? Des synthèses des littératures scientifique et technique ont été réalisées. L'expertise de phytopathologistes, de malherbologues et d'agronomes a été élicitée et formalisée. Des méthodes algorithmiques originales ont été développées pour adapter deux modèles génériques à différents bioagresseurs. Le modèle dynamique générique SimMat a été mobilisé pour aborder la première question. Ce modèle prédit les effets de la pluviométrie et de la température sur la reproduction sexuée de champignons ascomycètes. Il a été adapté au phomopsis et au phoma du tournesol. La seconde question

a été abordée à l'aide de l'approche IPSIM (Injury Profile SIMulator), qui permet de développer des modèles qualitatifs prédisant les effets du système de culture et de la situation de production sur les dégâts causés par un ou plusieurs bioagresseurs. Deux modèles qualitatifs ont été créés : un pour le phomopsis, un pour le phoma. Des modèles préliminaires ont été proposés pour le phoma et l'orobanche. L'ensemble de ces modèles a été adapté aux bioagresseurs grâce à des méthodes innovantes pour améliorer leurs qualités prédictives. En particulier, une méthode d'hybridation de l'expertise et des informations contenues dans les jeux de données a été proposée. La thèse a donc permis de développer un ensemble de modèles contribuant à la gestion des stress biotiques du tournesol, et des innovations méthodologiques génériques pour la modélisation. On dispose désormais d'outils opérationnels pour contribuer à la mise en place de stratégies de protection agroécologique du tournesol.

MOTS-CLES : *Helianthus annuus*, *Diaporthe helianthi*, *Leptosphaeria lindquistii*, *Orobanche cumana*, *modèle qualitatif*, *modèle épidémiologique*, *optimisation*, *Protection AgroEcologique des Cultures*

AUTHOR: Marie-Anne VEDY-ZECCHINI
THESIS SUPERVISORS: Jean-Noël AUBERTOT
THESIS CO-TUTOR: Marie-Hélène ROBIN
DEFENSE PLACE AND DATE: Centre INRAE, Auzeville, le 9 janvier 2020
TITLE: Analysis and modeling of the effects of cropping systems and production situations on the main pests of sunflower

---

ABSTRACT. Sunflower has many agroecological benefits (resistance to abiotic stresses, low Treatment Frequency Index, maintenance of pollinator biodiversity), but its surface area and yield have stagnated for nearly 30 years. These stagnations are due in particular to a collection of pests. In a context of pesticide reduction, non-chemical prophylactic (or preventive) methods must be promoted against pests. However, it is difficult to predict the effects of one or more cropping practice(s) on pest development. Models are essential tools to reach this objective. Three pests were considered in this thesis: two phytopathogenic fungi, phomopsis stem canker (*Diaporthe helianthi*) and phoma (*Leptosphaeria lindquistii*), because of the available knowledge and datasets; and a plant parasitizing exclusively sunflower, broomrape (*Orobanche cumana*) because of the damages it generates. Indeed, it has recently appeared in France and is rapidly expanding. These three pests can generate significant crop losses. The conducted work addressed three questions: i) How to manage primary inoculum production of these pests? ii) How to value the available knowledge on the effects of cropping practices and production situations on the biological cycles of these pests? iii) How to improve the modeling methods used? Syntheses of scientific and technical literature have been conducted. The expertise of plant pathologists and agronomists has been elicited and formalized. Original algorithms have been developed in order to adapt two generic models to different pests. The generic SimMat dynamic model was used to address the first question. This model predicts the effects of rainfall and temperature on the sexual reproduction of ascomycete fungi. It has been adapted to phomopsis and phoma of sunflower. The second question was addressed using the IPSIM (Injury Profile SIMulator) approach, which allows the development of qualitative models predicting the effects of cropping practices and production situations on injuries caused by one or several pest(s). Two qualitative models were created for phomopsis stem canker and phoma and the model for phomopsis stem canker was evaluated. A preliminary model has been proposed for

broomrape. Four of these models have been adapted using innovative methods to improve their predictive qualities. In particular, a method for hybridizing expert knowledge and the information contained in the datasets has been developed. The thesis therefore made it possible to develop a set of models contributing to the management of biotic stresses of sunflower, and generic methodological breakthroughs for modeling. Operational tools are now available to contribute to the implementation of agroecological protection strategies for sunflower.

KEYWORDS: *Helianthus annuus*, *Diaporthe helianthi*, *Leptosphaeria lindquistii*, *Orobanche cumana*, *qualitative model*, *epidemiological model*, *optimization*, *Agroecological Crop Protection*

Remerciements

J'ai attendu avec impatience de pouvoir écrire cette partie ! Je vais la comparer à un modèle (imparfait comme tout modèles, donc navrée pour les oublis !) basée sur une équation dont j'assume l'hétérogénéité.

Thèse = $\int_{01/01/2016}^{31/12/2019}$ les financements + les membres du jury et du comité de pilotage + mon entourage professionnel + les divers clubs + mes amis + ma famille

Je souhaite remercier les financeurs de cette thèse : Terres Inovia, la région Occitanie et l'INRAE.

Je remercie également les membres du jury pour m'avoir fait l'honneur d'accepter de lire, d'évaluer et de commenter ce travail, ainsi que les membres de mon comité de pilotage (Emmanuelle Bret-Mestries, Ivan Sache, Regis Sabbadin, Corentin Barbu, Vincent Faloya et Felicity Vear) pour m'avoir aiguillé au cours de la thèse.

Je vais maintenant décomposer les termes "mon entourage professionnel" : mon directeur de thèse * mon encadrante * Terres Inovia * AGIR * ACTA * MIAT * CIRAD * LIPM * ENSAT * PURPAN

Je remercie mon directeur de thèse, Jean-Noël Aubertot pour avoir accepté de m'encadrer, pour ta confiance, ta bienveillance, ta patience. Je te remercie aussi plus personnellement pour ton écoute lors de périodes difficiles et pour ton humour (même quand tu me taquinais "j'aime bien la faire rougir !"). Pour les soirées d'écriture de mon premier article "construction de SimMat/démolition d'ASPHODEL" qui finissaient en jeux de mot, blagues de toto, calisson d'Aix et chocolat... Je repense aux nombreux matchs de badminton, au cours desquels tu m'as mis la pâté (sauf une fois où je me souviens bien t'avoir smashé sur le crâne !). Je remercie Marie-Hélène Robin, pour m'avoir poussé dans mes raisonnements et la pédagogie grâce à ta vision d'agronome et d'enseignante, pour les nombreuses séances d'écriture. Je te remercie aussi plus personnellement pour le voyage à Florence pendant lequel nous avons fait connaissance.

Je remercie les membres de Terres Inovia que j'ai eu le plaisir de côtoyer. Merci à Emmanuelle Bret-Mestries, sans qui cette thèse n'aurait pas vu le jour. Merci d'avoir fait la demande de financement, d'avoir apporté ton expertise avec dynamisme. Merci de m'avoir poussé dans la précision et la rigueur et pour m'avoir épaulé en fin de thèse Je remercie Christophe Jestin et les

membres du Résorobanche pour leur aimable collaboration. Je remercie Luc Champolivier, Bernard Garric pour les pauses café au cours desquelles vous m'avez bien fait rire et taquiné sur mon régime alimentaire. Merci aussi à Célia Pontet pour sa sympathie depuis mon stage.

Je remercie les membres de l'unité AGIR pour les bons moments. Merci aux chercheurs de VASCO pour tous leurs conseils, aux techniciens de laboratoire pour leurs aides lors de mes expé, toujours dans la bonne humeur, et les équipes administrative et informatique pour leur efficacité. Je remercie spécialement Philippe Debaeke, le premier chercheur que j'ai abordé à AGIR et qui m'a fait basculer du côté lumineux (ou obscur ?) de la Force du tournesol. Merci pour ta gentillesse et ta disponibilité à chaque fois que j'avais une question sur la physio du tournesol, ou un article en roumain datant d'une époque antédiluvienne à te demander. J'ai pris beaucoup de plaisir à t'aider à dépoussiérer ta bibliothèque (malgré quelques éternuements) et je te lègue mon fauteuil de compète pour ta fin de carrière ! Merci à Célia Seassau et Gregory Dechamp-Guillaume pour leurs expertises. Merci à Jay-Ram Lamichhane, mon premier co-bureau pour les échanges franco-anglo-italien, les repas risotto et les sorties ! Je te souhaite une belle carrière et de voir ton fils s'épanouir ! Merci à Camilo Corrales Muños (Don Camilito) pour ta grande gentillesse, les nombreux fou-rire et les debugages LateX ! Merci à Pierre Casadebaig pour les discussions R, à Noémie Gaudio pour les moments de rigolades, à Claire Jouany (avec qui j'ai eu le plaisir de faire un bout de voyage à Florence), la passionnée de poule ! Mes collègues de bureaux de B205 et 206, Yolaine Hily, Julie Brechemier, Sandrine Longis, Andrea, Nicolas Meyer, Néila Aït-Kaci et Rémi Mahmoud pour les discussions sur les sciences, la musique et la politique ! Pour les séances de graphisme et coupe de buisson, pour les partages de plantes et de calissons/gâteaux/citrouilles/chocolats (avec et sans nanoparticules !)... Je remercie les collègues des bureaux alentours, Gwen Christiansen (don't stop me nooow !), Federica Bove, Amélia Landre, Félix Attia, Camille Launay, Julien Quénon, Antoine Couëdel, Loïc Viguie, Sixtine Cueff, ainsi que les anciennes docs de Magellan qui avaient instauré une dynamique sympa en début de thèse (Camille Lacombe, Sandrine Allain, Hélène Cristofari).

Cette thèse aurait été bien différente sans les membres du MIAT que j'ai considérée comme ma propre unité. Merci à Gauthier Quesnel avec qui j'ai collaboré pendant mon stage et tout le long de ma thèse. Pour ta bonne humeur constante ! Merci à Stéphane Couture qui m'a permis de faire cette thèse et pour ton soutien. Merci à Marie-Josée Cros (Marie-Jo) pour tes nombreux coups de main sur LateX ! Merci Robert Faivre pour tous tes conseils, ton soutien chocolaté ! Eric Casellas pour les débuggages sous R. Merci à Matthias Zytnicki, Simon de Givry, Claire Raynal, Nathalie Vialaneix, Ronan Trépos. Un grand merci à Sylvain Jasson, mon ancien prof qui a accueilli avec

bienveillance la doc "impertinente" que j'ai été à la table MIAT chaque vendredi. Merci aussi à Damien Leroux le relou et ses jeux de mots qui ont fini par me contaminer, à cause de toi je me suis essayée à la vannerie, Damien Berry, Etienne Auclair (le faux platiste), Alyssa Imbert, Sebastian Lecoz, Manon Ruffini, Lise Pomies, Floreal Cabanettes (parce que tu le valais bien), Léo Lamy... pour toutes les pauses café et bar. Vous avez réussi : la MIATisation a fait son oeuvre...

Merci à François Brun (d'ACTA). Pour ses bons conseils, sa disponibilité et les bons moments.

Merci à Jean-Philippe Deguine (CIRAD) pour les conversations intellectuelles (ou non) et les trafics de photos-dossiers !... Merci aussi à Natacha Motisi, pour sa bonne humeur et son humour !

Merci à l'équipe Tournesol du LIPM, pour leur collaboration. Merci les non-perms du LIPM et d'ailleurs, pour les soirées ciné, jeux de société : Pauline, Ariane, Rémy, Romain...

Je remercie l'ENSAT et PURPAN, en particulier Rhoda El Kouri, Christian Chervin, Benoit Van Der Rest, Jean-Pierre Sarthou, Jean-Noël et Marie-Hélène qui m'ont permis d'essayer l'enseignement dans les meilleures conditions possibles. J'en garde un très bon souvenir.

"Les divers club" : merci à Xavier Barlet et Olivier Catrice du club apiculture del'ADAS pour tout ce que vous m'avez appris ! Merci au club escalade mené par Sylvain Foissac. Merci au club badminton (ADAS). Merci à l'OAP de Mandoline et Guitare de Toulouse que je suis depuis plus de 10 ans, pour les nombreux concerts et les voyages !

Merci à mes amies Sophie Vora-Malpel et Elodie-France Turatsinze pour les soirées muséum et les sorties shopping/salon de thé/ciné/restau ! Merci pour la relecture de mon manuscrit ! Merci à Charles (pti Bigouden mu..., promis je ne donne pas la suite) pour les bons moments (Dame canard, 22lassy, mon pif bleu à Luminopolis, *Dialanthe*, ma transformation en minion violet en fin de thèse, j'en passe et pas des meilleures !). Merci à Sam Antonetti pour ta sympathie, pour le sport et les nombreuses soirées où tu m'as fait rire. Merci à Stéphane Donatien, qui me soutient depuis près de 10 ans et qui a trouvé comme seule question pour mes travaux de thèse "et pourquoi dans un champ de tournesol, il y en a toujours un qui n'est pas orienté comme les autres ?".

Je remercie ma famille pour son soutien inconditionnel depuis toujours. Ma mère Béatrice à qui ces 3 années ont rappelé des souvenirs (Zecchini, 1992), ma très chère soeur Flore à qui je souhaite une belle thèse et une carrière brillante. C'est grâce à vous que je suis arrivée jusqu'au bout. Merci à ma grand-mère Monique Zecchini chez qui je me suis ressourcée. En souvenir des devoirs de math que tu me faisais faire. Merci d'avoir rerere...lu mon manuscrit. Merci à mon parrain russe Xavier Jollivet (dyadya).

À ma mère,

à ma soeur Flore

et à ma grand-mère Monique Zecchini,

avec toute ma reconnaissance



Table des matières

Liste des figures 1

Liste des tableaux 5

Liste des abréviations 11

Glossaire 25

I Introduction

1 La réduction des pesticides : un enjeu d'actualité 31

2 Le tournesol : une culture phare pour la transition agroécologique 33

2.1 Le contexte économique de la culture de tournesol 33

2.2 Dispositifs de recherche sur le tournesol 36

2.3 Des intérêts agronomiques et environnementaux multiples... 37

2.4 ... et des freins limitants la compétitivité de la culture 40

2.4.1 Les bioagresseurs principaux 41

2.4.2 Les bioagresseurs secondaires 44

2.5 Les bioagresseurs considérés 45

2.5.1 Choix des bioagresseurs étudiés 45

2.5.2 Cycle de vie des bioagresseurs étudiés 45

3 Les méthodes de protection de cultures : la PIC et la PAEC... 53

3.1 La Protection Intégrée des Cultures 53

3.2 La Protection AgroEcologique des Cultures 56

3.3 Les méthodes de lutte contre les bioagresseurs d'intérêt 57

4 Rôle de la modélisation en protection des cultures 59

4.1 Exemples de modèles pour quelques cultures 60

4.1.1 Les modèles épidémiologiques 60

4.1.2 Les modèles de dégâts et de dommages 61

4.2 Les modèles pour le tournesol 64

4.2.1 Les modèles épidémiologiques 64

4.2.2 Les modèles de culture, les modèles de dégats et de nuisibilité 64

5 Objectifs de la thèse 65

6 Organisation du manuscrit 67

II Apports méthodologiques à la modélisation

1 Construction des modèles 71

1.1 Rappel des différentes étapes de construction d'un modèle 71

1.2 Les étapes de construction abordées au cours de la thèse 76

2 Présentation de l'approche de modélisation IPSIM 77

2.1 Le modèle générique IPSIM 77

2.2 Évaluation des modèles IPSIM dans la thèse 80

3 Comparaison des performances de l'approche IPSIM et d'une méthode statistique 83

3.1 Objectifs de l'étude 83

3.2 Matériel et méthodes 83

3.2.1 Description du principe de *Classification And Regression Tree* 83

3.2.2 Méthode de comparaison . . . 84
3.2.3 Modèles existants utilisés . . . 85
3.3 Résultats 86
3.4 Discussion 89
4 Méthode innovante améliorant la qualité prédictive des modèles IPSIM . . . 91
4.1 Principe de la méthode 91
4.1.1 Démarche algorithmique . . . 91
4.2 Validation informatique de l'algorithme améliorant la qualité prédictive des modèles IPSIM 92
4.2.1 Description d'IPSIM - Tournesol - Phoma sur tige et du jeu de données . . . 93
4.2.2 Résultats . . . 98
5 Optimisation de paramètres de modèles dynamiques simples . 101
III Modèles SimMat tournesol
1 Application à SimMat et *D. helianthi* . . . 109
2 Application à SimMat et *Leptosphaeria lindquistii* . . . 155
2.1 Introduction 155
2.2 Matériel et méthodes 156
2.2.1 Présentation du modèle SimMat . . . 156
2.2.2 Valeurs initiales des paramètres . . . 157
2.2.3 Jeux de données pour la calibration de SimMat-*L. lindquistii* . . . 157
2.2.4 Estimation des paramètres . . . 161
2.3 Résultats 162
2.3.1 Évaluation de la qualité de prédiction initiale de SimMat-*L. lindquistii* sur les dynamiques de taux de maturation de pseudothèces . . . 162
2.3.2 Identification de la combinaison des paramètres de SimMat-*L. lindquistii* à ajuster pour minimiser les erreurs de prédiction des dynamiques de maturation des pseudothèces . . . 165

2.3.3 Optimisation des paramètres de SimMat-*L. lindquistii* en utilisant le jeu de données entier 165

2.4 Discussion 165

2.4.1 Jeu de données utilisé pour l'estimation des paramètres 165

2.4.2 Résultats obtenus lors de l'évaluation et de la procédure d'optimisation des paramètres 166

2.4.3 Conclusion 167

IV **Modèles IPSIM tournesol**

1 **IPSIM - Tournesol - Phomopsis** 171

1.1 Design and evaluation of IPSIM-Sunflower-Phomopsis, a model that predicts the severity of Phomopsis stem canker on Sunflower 171

1.2 Amélioration de la qualité de prédiction d'IPSIM - Tournesol - Phomopsis 219

2 **IPSIM - Tournesol - Dessèchement précoce** 225

2.1 Identification, organisation des attributs, définition de leurs valeurs et de leurs poids 225

2.1.1 Modèle IPSIM - Tournesol - Dessèchement précoce sous DEXi 225

2.1.2 Commentaires du modèle, synthèse bibliographique 234

2.2 Description du convertisseur 239

2.3 Évaluation future du modèle IPSIM - Tournesol - Dessèchement précoce 245

3 **IPSIM - Tournesol - Orobanche** 247

3.1 Construction sur le modèle IPSIM - Tournesol - Orobanche 247

3.2 Commentaires sur le modèle, synthèse bibliographique 253

3.2.1 Degré d'infestation de l'orobanche 255

3.2.2 Risque au niveau régional 255

3.2.3 Climat 255

3.2.4 Pratiques culturales 256

3.3 Variables à développer 258

3.4 Convertisseur à développer 260

V Discussion et perspectives

1 Rappel des principaux résultats 265

2 Limites des modèles mis en place 267

2.1 Des modèles évolutifs 267

2.2 Des modèles encore davantage tournés vers la PIC que la PAEC 268

3 Avancées méthodologiques de la thèse 269

3.1 Généricité de la méthode d'optimisation de la qualité de prédiction de modèle quantitatif de faible complexité 269

3.2 Les modèles IPSIM : améliorations méthodologiques pour leur construction 270

3.2.1 IPSIM : des modèles statiques mais non figés 270

3.2.2 Compromis entre résolution et complexité 270

3.2.3 Pertinence des convertisseurs 271

4 Perspectives pour la modélisation pour le tournesol et la PAEC . 273

4.1 Avenir et utilisation future des modèles mis en place 273

4.1.1 SimMat-*Diaporthe* 273

4.1.2 IPSIM - Tournesol - Phomopsis 274

4.1.3 IPSIM - Tournesol - Phoma 274

4.2 Apport des modèles IPSIM existants 274

4.3 Réflexions sur la mise en place et l'optimisation de futurs modèles IPSIM277

4.4 IPSIM - Tournesol : prise en compte des interactions entre bioagresseurs278

4.5 Couplage des modèles SimMat, IPSIM et Sunflower-Pest pour la prédiction d'une partie du cycle épidémique, du profil de dégâts et des dommages associés 280

5 Conclusion 285

Bibliographie

Bibliographie 289

Annexes

Liste des figures

I.1 Évolution des surfaces de colza, de tournesol et de soja en France de 1977 à 2017 (d'après Terres Univia, 2017) 34

I.2 Évolution des rendements de colza, de tournesol et de soja en France de 1977 à 2017 (d'après Terres Univia, 2017) 34

I.3 Surfaces de tournesol en France en 2017 (d'après Terres Univia, 2017) . . . 35

I.4 Évolution des productions de colza, de tournesol et de soja en France de 1993 à 2017 (d'après Terres Univia, 2017) 36

I.5 Cycle de vie du tournesol et périodes d'observation des symptômes des bioagresseurs. Le stade est atteint lorsque 50% des plantes ont atteint le stade. () : échelle BBCH (Terres Inovia, 2019) 39

I.6 Cycle de vie de : *D. helianthi* (A) et *L. lindquistii* (B) (d'après Maric *et al.*, 1988 ; Chollet *et al.*, 2002) 48

I.7 Cycle de vie de *O. cumana* (d'après Maric *et al.*, 1988 ; Chollet *et al.*, 2002, Terres Inovia) 50

I.8 Schéma générique de la protection d'une culture face à un bioagresseur (Attoumani-Ronceux *et al.*, 2010) 55

I.9 Origines de la PAEC (Deguine *et al.*, 2016) 57

I.10 Représentation d'une fonction de dommage (Aubertot *et al.*, 2010) . . . 60

I.11 Schéma organisationnel du manuscrit 68

II.1 Étapes de construction d'un modèle 71

II.2 Principe de la validation croisée sans corrélation spatio-temporelle et de l'ajustement final. n_{max} est le nombre de paramètres du modèle et n est chaque entier entre 0 et n_{max} 75

II.3 Structure générique d'IPSIM telle que proposée par Aubertot et Robin (2013) 78

II.4 Illustration du calcul des poids dans un modèle DEXi composé de deux attributs agrégés composés respectivement de 3 et 4 valeurs. Les points bleus reliés représentent la fonction d'utilité définis par les règles de décision. En rouge, l'hyperplan créée par DEXi pour calculer les poids (d'après Bohanec, 2015). Val. = Valeur. 79

II.5 Boites à moustaches composées des 500 κ pondérés quadratiquement obtenues pour différentes tailles du jeu de données du système chayotte avec A : la méthode IPSIM. B : la méthode CART 87

II.6 Boites à moustaches composées des 1000 κ pondérés quadratiquement obtenues pour différentes tailles du jeu de données pour le système blé avec A : la méthode IPSIM. B : la méthode CART 88

II.7 Structure préliminaire de l'arbre IPSIM - Phoma sur tige représentée sous DEXi à gauche. Les attributs agrégés sont représentés en gras. Valeurs adoptées par chaque attribut à droite. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 94

II.8 Évolution du kappa de Cohen pondéré quadratiquement avec le nombre de règles d'agrégation ajustées simultanément. En rouge le kappa maximum obtenu en validation-croisée. En bleu tous les autres kappas 99

II.9 Illustration de la méthode algorithmique de Nelder-Mead sur un espace à 2-dimensions avec A : la mise en place d'un simplexe composé de trois points x_p (la pire valeur), x_s (la seconde valeur) et x_m (la meilleure valeur) et du centroïde du segment créé par x_p et x_m ; B : l'étape de réflection ; C : l'étape d'expansion ; D : l'étape de rétraction à l'extérieur du triangle ; E : l'étape de rétraction à l'intérieur du triangle et F : le rétrécissement. Légende : • représente les points de simplexe et le centroïde ; en bleu, le simplexe présent ; en rouge, le nouveau simplexe ; en pointillés : les modifications en cours ; P1 et P2 les axes des paramètres. 106

III.1 Distribution des taux de maturation des pseudothèces observés dans le jeu de données (189 observations issues de 9 départements et pour 5 années). 161

III.2 Résidus des prédictions de dynamique de SimMat-*L. lindquistii* sur le jeu de données entier (189 observations) avec : les valeurs initiales des paramètres (A) ; les valeurs finales des paramètres après optimisation. 164

IV.1 Évolution du kappa de Cohen pondéré quadratiquement en fonction du nombre de règles d'agrégation ajustées simultanément pour le modèle IPSIM - Tournesol - Phomopsis. En rouge, le kappa maximum obtenu en validation-croisée. 219

IV.2 Structure d'IPSIM-*O. cumana*. À gauche, la structure de l'arbre avec en gras, les attributs agrégés et en non-gras, les attributs basiques. À droite, les valeurs des attributs 248

V.1 Modèle générique IPSIM - X - Interactions sous DEXi (Aubertot et Robin, 2013) 279

V.2 Capture d'écran du modèle Sunflower-Pest sous la plateforme ModelBuilder (`xpest.inra.fr`). À gauche, le modèle Sunflower-Pest avec des rectangles représentants les variables météorologiques, le stade de développement de la culture, le pool de biomasse créée et les compartiments de la culture. Les ovales représentent les bioagresseurs du tournesol. Les flèches en pointillés représentent les processus biologiques (effets des variables climatiques sur la physiologie de la culture et les transferts de biomasse). À droite, le tableau de bord permet de définir les variables, les paramètres avec leurs équations, leurs valeurs et leurs unités. Les captures d'écran sont issues des travaux de M.A. Vedy-Zecchini (2016) et A. Costa (2017). 284

Liste des tableaux

I.1 Récapitulatif des caractéristiques des maladies provoquées par les principaux champignons phytopathogènes et de *Orobanche cumana* sur le tournesol (d'après Desanlis, 2013 ; Jestin, 2016) 42

I.2 Classification systémique de *D. helianthi* et *L. lindquistii* (d'après Agrios, 1936 ; Blackwell *et al.*, 2012) 46

I.3 Classification systémique de *O. cumana* (d'après Krupp *et al.*, 2015 49

I.4 Évaluation de l'effet de composantes du système contre le phoma, le phomopsis et l'orobanche du tournesol (d'après Debaeke *et al.*, 2017b). + : Défavorable à la maladie, - : Favorable à la maladie 57

II.1 Matrice de confusion entre une variable catégorielle observée et une variable catégorielle simulée. N : niveau maximal de la variable. x : valeur intermédiaire de la variable. X : un nombre de situation quelconque. En rouge, les situations sont les plus mal prédites ; en orange, les erreurs sont intermédiaires ; en vert, les situations sont bien prédites (cas d'une variable ordinale). 80

II.2 Table d'agrégation de l'incidence du phoma sur tige composée des 12 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 95

II.3 Table d'agrégation des pratiques culturales composée des 6 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 95

II.4 Table d'agrégation du contrôle cultural composée des 12 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 96

II.5 Table d'agrégation de l'atténuation par l'état du couvert composée des 4 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 96

II.6 Table d'agrégation de la gestion de l'inoculum primaire composée des 18 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige. 97

II.7 Description du jeu de données utilisé pour optimiser le modèle IPSIM - Phoma sur tige préliminaire. Numéro de département, année, nombre de parcelles observées et nombre et niveau d'incidences observées. 98

II.8 Valeurs des kappas de Cohen pondérés quadratiquement obtenus avec l'algorithme codé sous C++ et codé sous R sans ajustement, puis avec l'ajustement d'une, puis de deux règles d'agrégation. 99

III.1 Paramètres, définitions et unités des paramètres de SimMat. 156

III.3 Présentation du jeu de données (189 observations) utilisé pour calibrer le modèle SimMat-*L. lindquistii*. Provenance géographique des résidus infectés (09 : Ariège, 11 : Aude, 18 : Cher, 31 : Haute-Garonne, 32 : Gers, 33 : Gironde, 35 : Ille-et-Vilaine, 47 : Lot-et-Garonne). Les taux de maturation séparés d'un « ; » correspondent à des parcelles différentes du même site et de la même année . . 160

III.4 Critères utilisés dans l'étude, type de qualité et étape de chaque symbole. SM : SimMat, *L. l.* : *Leptosphaeria lindquistii* ; C.V. : Cross-Validation. 162

III.5 Résultats obtenus avec la procédure d'optimisation des paramètres de SimMat-*L. lindquistii*. Intervalle de valeurs testées, valeurs de paramètres avant estimation, origine des valeurs, valeurs de paramètres après estimation, et leurs unités. En gras : nouvelles valeurs obtenues après optimisation. 163

III.6 Valeurs de MSE, $Biais$ et $Efficience$ calculées pour la qualité de prédiction initiale de SimMat-*L. lindquistii*, pour la qualité de prédiction obtenue après la procédure d'optimisation et pour la qualité d'ajustement. 163

IV.1 Synthèse des sorties de l'algorithme pour optimiser le modèle IPSIM - Tournesol - Phomopsis en ajustant quatre règles d'agrégation. Les valeurs initiales et optimisées sont les résultats des règles d'agrégation. 220

IV.2 Table d'agrégation de l'incidence du modèle IPSIM - Tournesol - Phomopsis avant optimisation de la qualité de prédiction telle que présentée sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables. 221

IV.3 Table d'agrégation de la pression du phomopsis du modèle IPSIM - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables. 221

IV.4 Table d'agrégation des Pratiques culturales du modèle IPSIM - Sunflower - Tournesol après optimisation de la qualité de prédiction telles que présentées sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables. 222

IV.5 Local and global weights of attributes of IPSIM-Sunflower-Phomopsis stem canker 223

IV.6 Structure et échelle (Scale) des attributs (Attribute) d'IPSIM - Tournesol - Dessèchement précoce. À gauche, la structure de l'arbre. À droite, les valeurs des attributs. En rouge, les valeurs sont défavorables au tournesol, en noir, les valeurs sont intermédiaires et en vert les valeurs sont favorables au tournesol 226

IV.7 Attributs d'IPSIM - Tournesol - Dessèchement précoce et leurs poids locaux et globaux normalisés. Niv. = Niveau 227

IV.8 Table d'agrégation de la sortie du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi avant optimisation. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 228

IV.9 Table d'agrégation de l'incidence du phoma du collet du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 229

IV.10 Table d'agrégation de la pression du phoma à l'échelle de l'exploitation du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 230

IV.11 Table d'agrégation du sol et du climat du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 230

IV.12 Table d'agrégation des pratiques culturales du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 231

IV.13 Table d'agrégation des conditions pour le dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 231

IV.14 Table d'agrégation du sol et du climat pour les conditions permettant l'expression du dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 232

IV.15 Table d'agrégation des pratiques culturales pour les conditions pour l'expression du dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 232

IV.16 Table d'agrégation du contrôle cultural du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 233

IV.17 Table d'agrégation de l'atténuation par l'état du couvert du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet. 233

IV.18 Synthèse des connaissances sur le phoma au collet et le dessèchement précoce, le sens de l'effet et les sources bibliographiques 238

IV.20 Variables d'entrée d'IPSIM, valeurs des variables d'entrée d'IPSIM, Variable(s) d'entrée utilisée(s) dans le convertisseur, valeurs des variables d'entrée du convertisseur et leurs unités 241

IV.21 Valeurs de k_c pour chaque période phénologique du tournesol (Debaeke, communication personnelle) 244

IV.22 Convertisseur définissant le statut azoté du tournesol (en unité, u) en fonction du reliquat azoté (en unité, u), la profondeur du sol et la fertilisation azotée . 245

IV.23 Table des poids des attributs du modèle d'IPSIM-*O. cumana*. De gauche à droite : les poids locaux normalisés et globaux normalisés. 249
IV.24 Table d'agrégation du degré d'infestation du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 250
IV.25 Table d'agrégation des pratiques culturales du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 250
IV.26 Table d'agrégation du contrôle cultural du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 251
IV.27 Table d'agrégation de la gestion du stock grainier du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche 251
IV.28 Table d'agrégation du développement de l'orobanche post-fixation du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 251
IV.29 Table d'agrégation de l'effet du climat du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 252
IV.30 Table d'agrégation de l'effet du risque régional à l'échelle de la parcelle du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche). 252
IV.31 Synthèse des connaissances sur l'orobanche avec les attributs, le sens de l'effet et les sources bibliographiques 254
V.1 Caractéristiques des modèles IPSIM construits pour le blé, le tournesol, le pommier, la chayotte et le caféier 276

Liste des abréviations

ACP **A**nalyse en **C**omposantes **P**rincipales. 277

BSV **B**ulletin de **S**anté du **V**égétal. 76, 273, 285

CART **C**lassification **A**nd **R**egression **T**ree. 2, 83, 85–90, 271, 272

CESBIO **C**entre d'**E**tudes **S**patiales de la **BIO**sphère. 37

DEXi **D**ecision **EX**pert. 2, 3, 5–9, 77, 79, 94–97, 221, 222, 228–233, 250–252, 275, 279

EI **E**cole d'**I**ngénieurs de Purpan. 37, 171, 274, 282

ENSAT **E**cole **N**ationale **S**upérieure **A**gronomique de **T**oulouse. 37, 171, 237, 274

IFT **I**ndice de **F**réquence de **T**raitements phytosanitaires. 40

IMI **IMI**dazlinones. 257

INRAE **I**nstitut **N**ational de la **R**echerche pour l'**A**griculture, l'**A**limentation et l'**E**nvironnement. 36, 158, 161, 171, 237, 245, 247, 258, 265, 275

IPSIM **I**njury **P**rofile **SIM**ulator. 2, 3, 5–9, 15, 17, 60, 61, 65, 66, 76–80, 83–98, 219, 221, 222, 225–234, 239, 242, 243, 245, 247–252, 255, 263, 266, 267, 270, 272, 274–278, 282

LIPM **L**aboratoire des **I**nteractions **P**lantes **M**icro-organismes. 36, 37

LPT **L**imite **P**assage **T**racteur. 236, 261

MIAT **M**athématiques et **I**nformatique **A**ppliquées de **T**oulouse. 92, 106, 270

MSE **M**ean **S**quared **E**rror. 101, 102, 163, 282

OAD **O**util d'**A**ide à la **D**écision. 31, 32, 56, 72, 76, 79, 273, 274

OLS **O**rdinary **L**east **S**quares. 102

PAEC **P**rotection **A**gro**E**cologique des **C**ultures. 1, 56, 57, 268, 275, 285

PIC **P**rotection **I**ntégrée des **C**ultures. 53, 56, 57, 268, 285

SRPV **S**ervice **R**égional de la **P**rotection des **V**égétaux. 158

UE **U**nion **E**uropéenne. 31

UMR AGIR **U**nité **M**ixte de **R**echerche **AG**roécologie, **I**nnovation et te**R**ritoires. 36

VASCO **VA**riétés **S**ystèmes de **C**ulture pour une production agr**O**-écologique. 36, 37

VBA **V**isual **B**asic **A**pplication. 80, 273

VTH **V**ariété **T**olérante aux **H**erbicides. 257

Glossaire

adventice : plante poussant dans une parcelle sans y avoir été intentionnellement semée. 40, 53, 54, 59, 60, 257

agroécologie : alternative à une agriculture intensive basée sur l'artificialisation des cultures par l'usage d'intrants de synthèse (engrais, pesticides...) et d'énergies fossiles. Elle promeut des systèmes de production agricole valorisant la diversité biologique et les processus naturels (cycles de l'azote, du carbone, de l'eau, équilibres biologiques entre organismes ravageurs et auxiliaires des cultures...). (https://dicoagroecologie.fr/encyclopedie/agroecologie/) 59

allélopathie : ensemble des interactions biochimiques réalisées par les plantes entre elles, ou avec des microorganismes (https://dicoagroecologie.fr/encyclopedie/allelopathie/). 54

ascomycètes : embranchement des Mycètes réalisant leur reproduction sexuée avec des ascospores dans des asques (Agrios, 1936). 45

ascospores : spores issues de la méiose puis de la mitose. Elles sont produites dans les asques contenus dans les structures reproductrices (périthèces, pseudothèces, cléistothèces, apothécies) chez les ascomycètes (http://ephytia.inra.fr/fr/C/21335/Di-gno-Leg-Peritheces). 45, 46, 60, 61, 156–158, 166, 167, 235, 236, 242, 266, 273, 279

asque : cellule zygote contenus dans les structures reproductrices (périthèces, pseudothèces, cléistothèces, apothécies) chez les ascomycètes. Elles peuvent être unituniqués ovoïdes à cylindriques. Il existe deux types d'asques : persistants libérant leurs ascopores à travers un pore apical et les asques déhiscents libérant leurs spores passivement par déhiscence de la paroi (http://www.univ-brest.fr/esiabscientifique/Mycologie/Classificati

on/Ascomycota/Classification+Alexopouloss) 46

bioagresseur : organisme pouvant engendrer des dommages sur des cultures. Il peut s'agir d'agents pathogènes responsables de maladies, de ravageurs ou de plantes adventices (Aubertot *et al.*, 2010). 1, 3, 13, 14, 29, 31, 37, 39, 41, 44, 45, 47, 49, 51, 53–61, 65–67, 77, 85, 90, 235, 237, 265, 266, 269–271, 275, 277–280, 282, 284, 285

biogéographie : étude de la répartition des organismes dans un espace 56

contamination : séquence de la maladie où l'agent pathogène a pénétré dans l'hôte à partir de ses propres ressources (Rapilly, 1991). 43, 53, 54, 64, 236, 261

dégât : toute déviation visible ou mesurable par rapport à une plante saine (symptôme) causée par la présence d'un bioagresseur sur une culture (jaunissement, nécroses,. . .) (Zadoks, 1993 ; Aubertot *et al.*, 2005). 53, 59, 60, 64, 65, 77, 275, 278

dommage : (aussi appelé nuisibilité, Aubertot *et al.*, 2010), perte de récolte, quantitative ou qualitative, due à l'activité d'un bioagresseur donné.. 1, 14, 44, 53, 60, 61, 64, 76, 280, 282

écologie du paysage : étude de la variation spatiale dans les paysages à différentes échelles 56

évitement : aussi appelée esquive, cette stratégie consiste à éviter la concordance entre la phase de contamination du bioagresseur et la période de sensibilité de la culture. Le principal levier est le raisonnement de la date de semis, à coupler avec un choix de variétés adéquat (Aubertot *et al.*, 2010). 53, 236, 244

haustorium : organe d'une plante ou d'un champignon parasite, qui s'insère dans les tissus de la plante hôte, pour absorber l'eau et les nutriments contenus dans les cellules (Agrios, 1936). 256

hydathode : structure sur la surface foliaire avec une ou plusieurs ouvertures qui évacuent l'eau de la feuille (Agrios, 1936). 44

incidence : fréquence de plantes subissant des dégâts à l'échelle d'une parcelle (Savary *et al.*, 1994). 5–7, 44, 85, 92, 93, 95, 98, 99, 219–223, 229, 234–236, 265, 267, 274

infection : pénétration de l'hôte par l'agent pathogène (Agrios, 1936). 41, 46, 56, 59, 65, 227, 235, 236, 240, 242, 261, 273

inoculum : terme générique qui caractérise tout élément du parasite capable de contaminer un hôte (Rapilly, 1991). 6, 53, 59, 65, 97, 100, 109, 155, 165, 235, 274

lutte biologique : lutte utilisant des organismes vivants pour prévenir ou réduire les dommages causés par des bioagresseurs. 43, 58, 261, 268

lutte chimique : lutte consistant à utiliser des produits phytosanitaires pour la protection des cultures (Aubertot *et al.*, 2010). 31, 43, 54, 56, 57, 223, 224, 227, 236

lutte physique : lutte consistant à utiliser des moyens mécaniques, thermiques, électromagnétiques ou pneumatiques pour le contrôle des bioagresseurs (Exemple du désherbage mécanique en grandes cultures) (Panneton *et al.*, 2000). 58

pratiques agricoles : ensemble des actes techniques mis en oeuvre sur une parcelle par un agriculteur. 45, 59, 75–77

pseudothèce : stroma (matrice) de mycélium à l'intérieur de laquelle se trouve des cavités où se forment les asques permettant la reproduction sexuée chez les Loculoascomycètes de la branche des Ascomycètes (Nasraoui, 2003 ; Agrios, 1936). 2, 15, 45, 46, 61, 64, 157–159, 161, 162, 165, 166, 273

pycnide : structure de reproduction asexuée sphérique ou en forme de flasque chez les ascomycètes. Elle contient des conidiophores et produit des conidies (Agrios, 1936). 45

race : ensemble des souches, ou des variétés possédant en commun le même type de pouvoir pathogène par rapport à une ou plusieurs variétés de plante-hôtes. Elles sont définies par leur virulence ou leur avirulence par rapport à une source de résistance (Agrios, 1936, `http://ephytia.inra.fr/fr/C/16226/Hypp-encyclopedie-en-protection-des-plantes-race-rhyzomorphe`). 43, 256, 257

réceptive : qualifie une plante-hôte dont les tissus sont, en dehors de tout aspect de résistance, dans un état où un parasite peut les contaminer plus ou moins intensément. Cette propriété est propre à un couple plante-bioagresseur et son effet est mesurable au niveau de la maladie et/ou de l'épidémie (Rapilly, 1991). 236

rotation : suite des cultures qui se succèdent sur une parcelle dans un sens ordonné, et se reproduisant de manière semblable au cours du temps. Lorsque cette dernière condition n'est pas respectée, on utilise plutôt le terme de succession culturale (Vilain, 1989). 40, 41, 53, 247, 249, 257

sévérité : pourcentage de surface atteinte sur une plante subissant des dégâts de bioagresseur (Savary *et al.*, 1994). 41

situation de production : ensemble des caractéristiques physiques, chimiques et biologiques d'un champ cultivé (en dehors du peuplement cultivé) et de son environnement, ainsi que par les déterminants socio-économiques et techniques susceptibles d'influencer les décisions de l'agriculteur. Dans cette définition, l'environnement de la parcelle fait référence au climat

et aux éléments du territoire susceptibles d'interagir avec la parcelle considérée (autres parcelles, bordure, forêts, ...) via des flux biotiques ou abiotiques (Aubertot et Robin, 2013). 45, 60, 65, 76, 77, 253

système de culture : ensemble des modalités techniques mises en oeuvre sur des parcelles traitées de manière identique. Chaque système de culture se définit par la nature des cultures, leur ordre de succession (rotation) et les itinéraires techniques appliqués à ces différentes cultures (Sebillotte, 1990). 40, 65, 77, 253

Introduction

1 La réduction des pesticides : un enjeu d'actualité 31

2 Le tournesol : une culture phare pour la transition agroécologique 33

2.1 Le contexte économique de la culture de tournesol

2.2 Dispositifs de recherche sur le tournesol

2.3 Des intérêts agronomiques et environnementaux multiples...

2.4 ... et des freins limitants la compétitivité de la culture

2.5 Les bioagresseurs considérés

3 Les méthodes de protection de cultures : la PIC et la PAEC... 53

3.1 La Protection Intégrée des Cultures

3.2 La Protection AgroEcologique des Cultures

3.3 Les méthodes de lutte contre les bioagresseurs d'intérêt

4 Rôle de la modélisation en protection des cultures 59

4.1 Exemples de modèles pour quelques cultures

4.2 Les modèles pour le tournesol

5 Objectifs de la thèse 65

6 Organisation du manuscrit 67

1. La réduction des pesticides : un enjeu d'actualité

Au cours du vingtième siècle, l'agriculture française a été marquée par les progrès de la mécanisation et de la sélection variétale, ainsi que par l'utilisation d'intrants issus de la chimie de synthèse. La lutte chimique a connu un essor particulièrement important à partir de la fin de la seconde guerre mondiale, avec l'apparition des pesticides de synthèse. Ces derniers se sont révélés plus efficaces que les pesticides d'origine naturelle et les ont progressivement remplacés. Cette intensification de l'agriculture française a permis de répondre aux besoins alimentaires d'après-guerre. L'utilisation massive de produits phytosanitaires au cours de ces cinq dernières décennies a conduit à la pollution des écosystèmes et à des problèmes de santé humaine (Tegtmeier et Duffy, 2004 ; Geiger *et al.*, 2010). De plus, les pesticides perdent leur efficacité du fait de l'apparition de nombreux cas de résistances chez différents bioagresseurs (Lamichhane *et al.*, 2016). En 2009, l'**U**nion **E**uropéenne (UE) a mis en place la directive européenne 2009/128 qui impose à chaque État membre de mettre en place un plan d'actions pour une utilisation des pesticides compatible avec le développement durable (`https://agriculture.gouv.fr/reduction-des-pesticides-les-indicateurs-de-risque-harmonises-etablis-au-niveau-europeen`). En France il s'est agi de la mise en place de versions successives du plan Ecophyto (`www.agriculture.gouv.fr/le-plan-ecophyto-quest-ce-que-cest`, Lamichhane *et al.*, 2019). En 2008, un premier plan (Ecophyto I) avait été mis en place pour réduire de 50% l'utilisation des pesticides en 10 ans. Après l'échec de ce premier plan, un deuxième plan (Ecophyto II) a été mis en place pour réduire de 25% l'utilisation des pesticides d'ici 2020, puis de 50% d'ici 2025 (Anonyme, 2015).

L'axe 1 de ce plan porte sur l'évolution des pratiques culturales. Une des manières envisagées pour atteindre cet objectif est la mise en place et l'utilisation d'**O**util d'**A**ide à la **D**écisions (OADs).

Ces derniers sont de plus en plus utilisés en agriculture. Ils peuvent être définis de manière très globale comme des outils permettant à l'utilisateur de prendre une décision face à une situation pour répondre à une problématique donnée (Gorry et Scott-Morton, 1971). Dans le contexte Ecophyto II, les OADs ont pour but d'encourager et d'aider les agriculteurs à réduire leur utilisation des pesticides de synthèse.

2. Le tournesol : une culture phare pour la transition agroécologique

2.1 Le contexte économique de la culture de tournesol

Le tournesol est la quatrième huile produite dans le monde derrière les huiles de palme, de soja et de colza (Pilorgé, 2016). La France arrive en cinquième position en termes de surface en Europe en 2017. En France, le tournesol a connu une brusque augmentation de ses surfaces en 10 ans (plus de 1.000.000 ha à la fin des années 1980 et au début des années 1990, Figure I.1), puis il a connu une diminution progressive de 500.000 ha en 20 ans pour stagner depuis 10 ans en deuxième position chez les oléagineux (entre 500.000 et 750.000 ha en 2018, estimation d'Agreste d'après SAA), derrière le colza. Son rendement stagne depuis 40 ans entre 15 et 30 $q.ha^{-1}$, alors que les rendements du colza et le soja augmentent progressivement. En effet, il est en troisième position pour le rendement (22,6 $q.ha^{-1}$), derrière le soja et le colza (respectivement 26 $q.ha^{-1}$ et 28,7 $q.ha^{-1}$, Figure I.2).

Le tournesol est cultivé essentiellement en région Occitanie et Nouvelle-Aquitaine (Figure I.3).

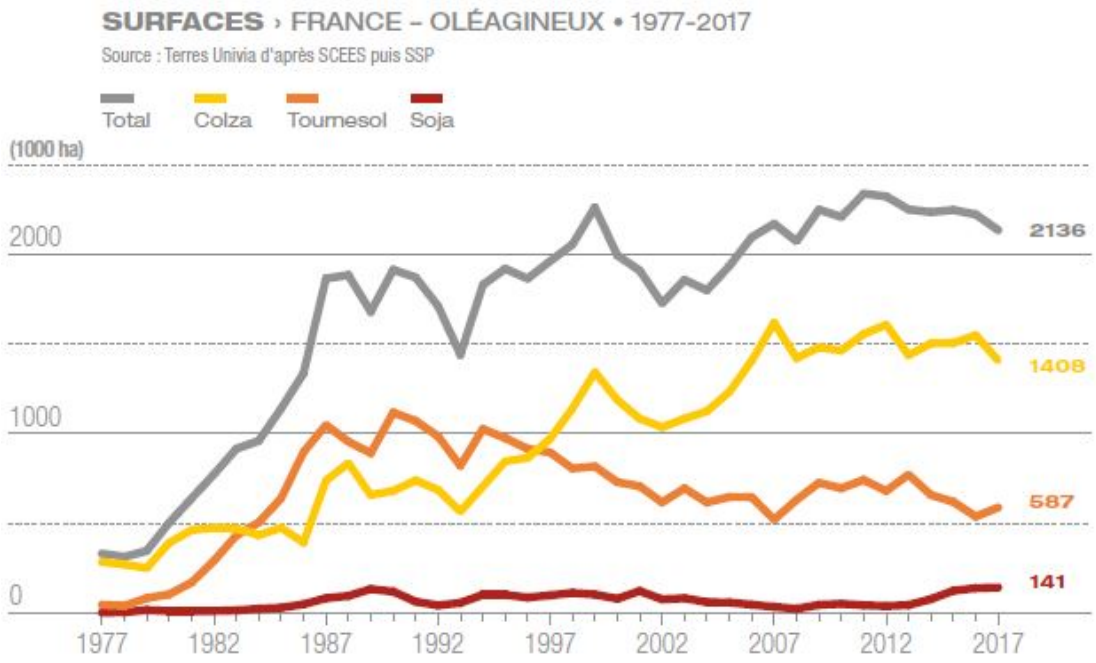


Figure I.1 – Évolution des surfaces de colza, de tournesol et de soja en France de 1977 à 2017 (d'après Terres Univia, 2017)

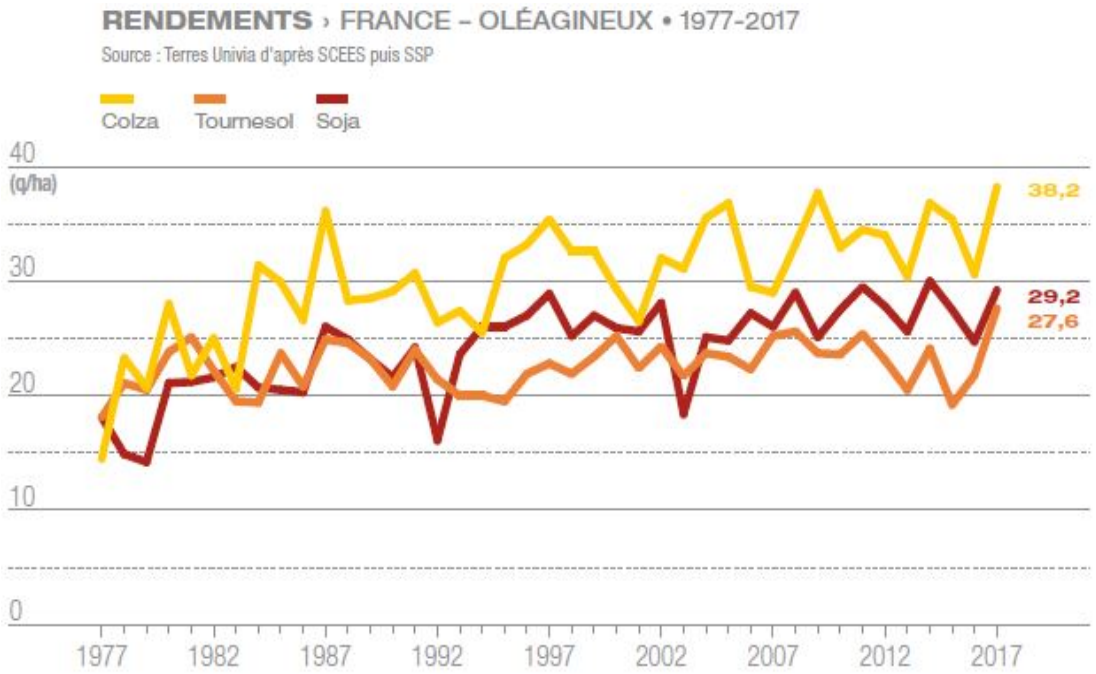


Figure I.2 – Évolution des rendements de colza, de tournesol et de soja en France de 1977 à 2017 (d'après Terres Univia, 2017)

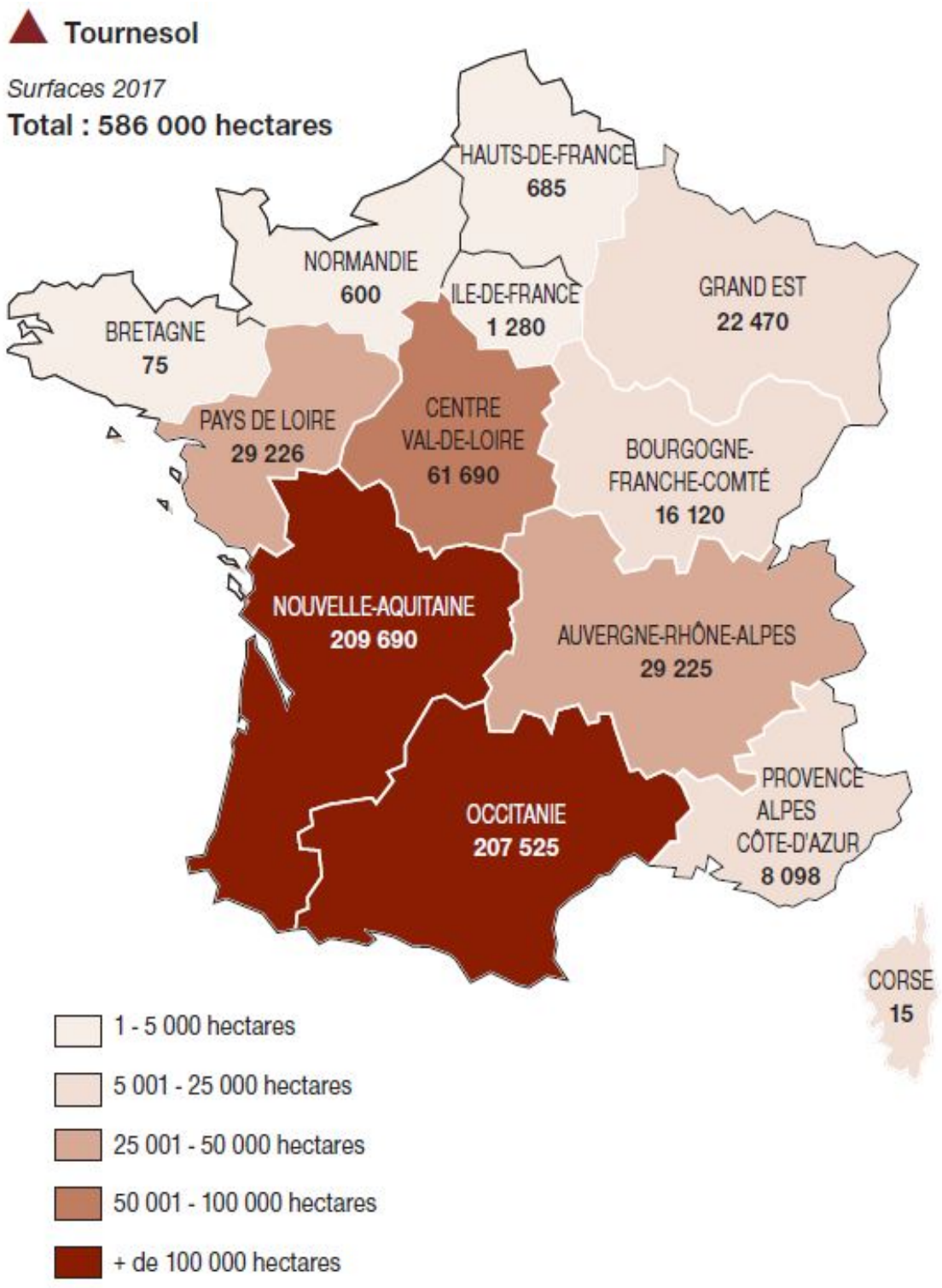


Figure I.3 – Surfaces de tournesol en France en 2017 (d'après Terres Univia, 2017)

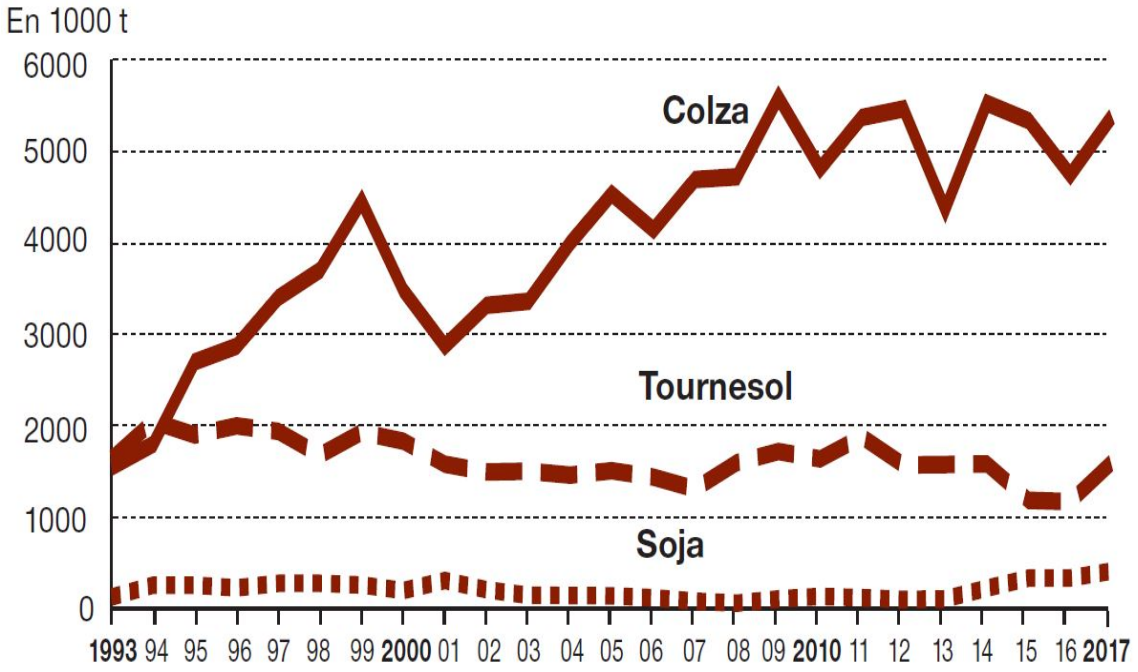


Figure I.4 – Évolution des productions de colza, de tournesol et de soja en France de 1993 à 2017 (d'après Terres Univia, 2017)

Cette diminution des surfaces depuis le début des années 1990 est due à des changements de pratiques culturales des agriculteurs qui ont cessé de cultiver le tournesol dans les sols fertiles et le cultivent désormais dans des sols de moins bonne qualité. Il en découle que les progrès génétiques réalisés sur les variétés de tournesol ces dernières décennies (Vear et Muller, 2011) servent à maintenir un rendement stagnant entre 15 et 30 q.ha^{-1} (Figure I.2) dans ces sols, en condition de stress biotiques et abiotiques (Vear, 2016), alors que le rendement de variétés très productives pourrait atteindre plus de 40 q.ha^{-1} (Skoric *et al.*, 2007). La baisse de surface et la stagnation du rendement du tournesol depuis le début des années 1990 entraînent une diminution progressive de la production de tournesol en France (Figure I.4), donc une diminution de sa compétitivité. Cette dernière doit être augmentée en améliorant le niveau de production du tournesol et sa stabilité.

2.2 Dispositifs de recherche sur le tournesol

En France, la recherche sur la culture du tournesol est principalement conduite sur le pôle toulousain. L'espèce fait l'objet de partenariats scientifiques entre plusieurs laboratoires de recherche (l'équipe **VA**riétés **S**ystèmes de **C**ulture pour une production agr**O**-écologique (VASCO) de l'**U**nité **M**ixte de **R**echerche **AG**roécologie, **I**nnovation et te**R**ritoires (UMR AGIR) et l'équipe Tournesol du **L**aboratoire des **I**nteractions **P**lantes **M**icro-organismes (LIPM) de l'**I**nstitut **N**ational de la

Recherche pour l'**A**griculture, l'**A**limentation et l'**E**nvironnement (INRAE), l'institut Terres Inovia , des réseaux de semenciers et de selectionneurs (Euralis, Caussades semences, RAGT, Maïsadour, Promosol), les **E**cole d'**I**ngénieurs de Purpan (EI) et de l'**E**cole **N**ationale **S**upérieure **A**gronomique de **T**oulouse (ENSAT) et le **C**entre d'**E**tudes **S**patiales de la **BIO**sphère (CESBIO). Les agronomes et les modélisateurs travaillant sur le tournesol au sein de l'équipe VASCO ont mené des études sur l'impact des pratiques culturales et du contexte pédoclimatique sur le tournesol et ses bioagresseurs et sur son rendement. Le LIPM étudie la génomique du tournesol et notamment les interactions d'un point de vue génétique entre le tournesol et des bioagresseurs émergeants. Terres Inovia est un institut technique qui a pour mission d'améliorer la compétitivité des oléagineux, des protéagineux et du chanvre. Cet institut mène des enquêtes sur l'évolution des pratiques culturales en France et a mis en place plusieurs services pour le conseil agricole concernant les différentes cultures d'intérêt depuis de nombreuses années. Les partenariats entre ces différents organismes ont permis de mener des études de sélection variétale, de mener des enquêtes sur les pratiques culturales et de mener des projets de recherches scientifiques.

2.3 Des intérêts agronomiques et environnementaux multiples...

Le tournesol est considéré comme rustique car tolérant aux stress abiotiques :

— c'est une culture d'été qui est peu exigeante en eau (Champolivier *et al.*, 2011) : son rendement maximal peut être atteint si seulement 70% de ses besoins en eau sont couverts (Merrien et Milan, 1992). En comparaison, le maïs (autre culture d'été) maximise sa production uniquement si 100% de ses besoins sont couverts (Merrien et Grandin, 1990 ; Nolot et Debaeke, 2003). Sa surface foliaire s'ajuste à la disponibilité en eau (Blanchet et Merrien, 1990) et un indice foliaire de 2,5 à 3 à la floraison est optimal pour sa productivité (Merrien et Milan, 1992). Il supporte bien le stress hydrique pré-floraison puisque des essais en champs ont prouvé qu'une réduction de sa consommation en eau ne réduit pas sa biomasse (Merrien et Milan, 1992). En revanche, le stress hydrique post-floraison peut le pénaliser dans certaines conditions et sa sénescence sera plus rapide. Dans un contexte de changement climatique où des périodes de sécheresse de plus en plus rudes sont attendues, notamment dans le sud-ouest de la France, cette résistance au stress hydrique est un avantage (Debaeke *et al.*, 2017a).

— Il est peu exigeant en azote, ses besoins s'élèvent à 140 kg.ha^{-1} d'azote pour un rendement de 30 q.ha^{-1} (Merrien et Milan, 1992). En comparaison, les besoins du colza s'élèvent à environ 280 kg.ha^{-1} pour un rendement de 40 q.ha^{-1}. Deux-tiers de l'azote absorbé par le

tournesol proviennent des reliquats azotés du sol. Son système racinaire pivotant est capable d'exploiter des horizons à plus d'un mètre de profondeur (Merrien et Milan, 1992).

Le cycle du tournesol est composé de 5 phases : germination-levée, phase végétative, développement du bouton floral, floraison et maturation (Figure I.5). Il est généralement semé aux mois d'avril-mai et récolté en août-septembre.

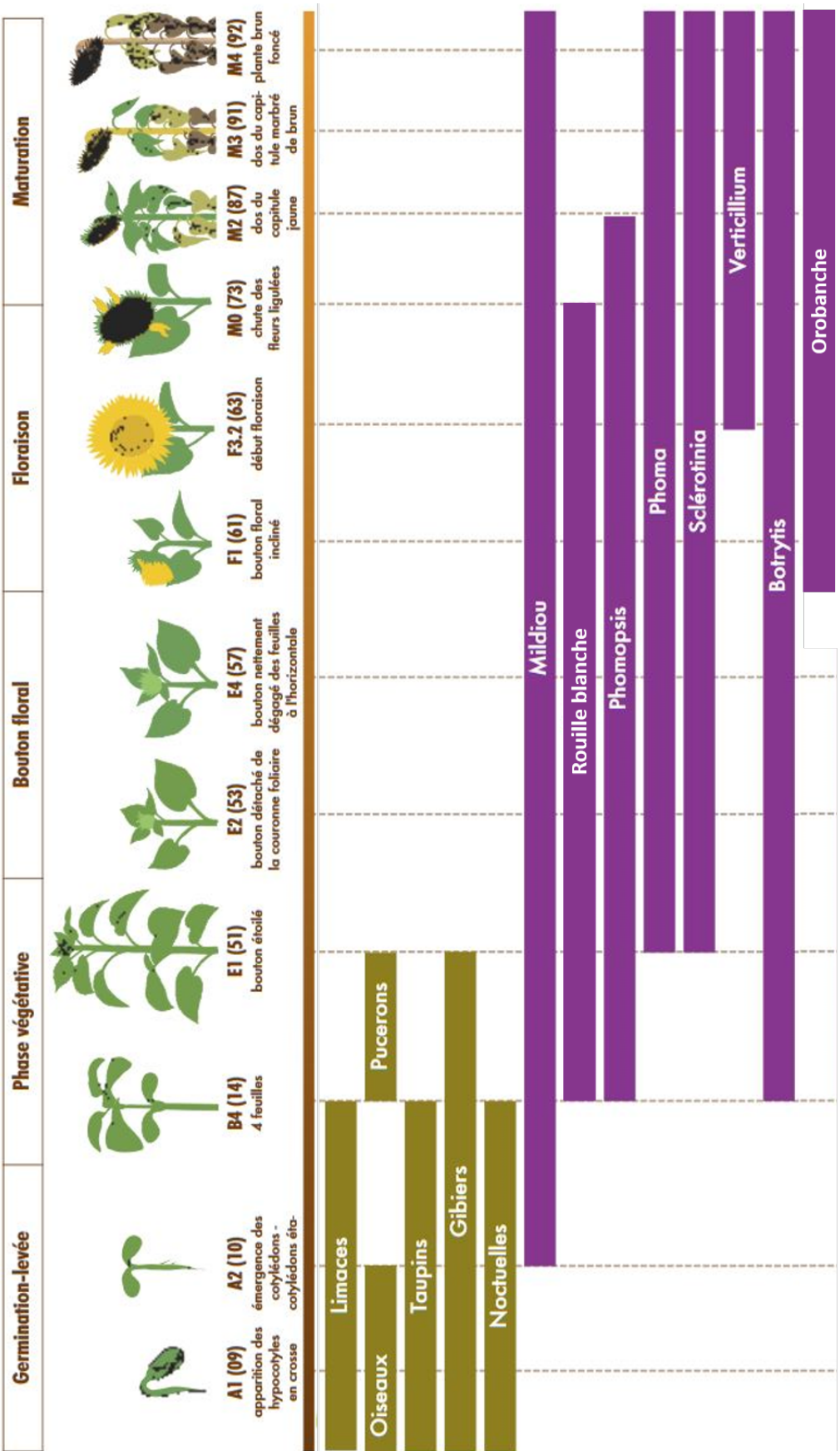


Figure I.5 – Cycle de vie du tournesol et périodes d'observation des symptômes des bioagresseurs. Le stade est atteint lorsque 50% des plantes ont atteint le stade. () : échelle BBCH (Terres Inovia, 2019)

Le tournesol possède un cycle court, ce qui le rend adaptable à de larges gammes de situations pédoclimatiques et est peu exigeant en chaleur (base de température de 6°C), ce qui permet une certaine souplesse dans la date de semis (semis possible du 1er avril au 20 mai, Lecomte et Nolot, 2011).

Le tournesol fait parti des cultures à faible **I**ndice de **F**réquence de **T**raitements phytosanitaires (IFT) (évalué à 2,7 en 2017 ; 1,8 hors traitement de semences, Agreste, 2019). En comparaison, le colza a un IFT de 6,4 (5,6 hors traitement de semences) et le soja a un IFT de 1,8 (1,7 hors traitement de semences).

Le tournesol est un excellent précédent cultural (Lecomte et Nolot, 2011). Il laisse un sol bien structuré, sans risque de lixiviation de l'azote, et avec moins d'adventices et de champignons phytopathogènes. Au contraire, le soja laisse un sol très mal structuré (Lecomte et Nolot, 2011) et si le tournesol le suit, il a un risque accru d'être en compétition avec des adventices et de développer des maladies. Le colza augmente le risque de lixiviation et/ou de risque sanitaire pour d'autres cultures (tournesol, soja...).

Outre ses atouts agronomiques, il dispose aussi d'atouts environnementaux. En effet, il favorise la biodiversité des pollinisateurs au début de l'été (Delaplane et Mayer, 2000 ; Cerrutti et Pontet, 2016). Il est la première culture oléagineuse BIO en France (environ 33.000 ha sont cultivés en agriculture biologique en 2018, d'après l'Agence Bio). Son faible besoin en azote entraîne une faible émission de gaz à effets de serre (BIO IS, 2010). Le tournesol dispose donc de nombreux atouts agronomiques et environnementaux (Jouffret *et al.*, 2011 ; Debaeke *et al.*, 2017a).

Le tournesol est compatible avec de nombreux systèmes de culture (Lecomte et Nolot, 2011). Divers systèmes de culture incluant du tournesol sont rencontrés en France : par exemple, des systèmes de culture en rotation courte dans les coteaux du sud-ouest de la France, des systèmes en rotation longue (retour du tournesol sur plus de 2 an) en Poitou-Charentes, du tournesol irrigué en Rhône-Alpes et en système de culture biologique dans le sud en général. Le fait de pouvoir insérer le tournesol dans divers systèmes de culture constitue un avantage valorisant l'adaptabilité du tournesol. Il est donc important de pérenniser la culture du tournesol dans ces différents systèmes de culture.

2.4 ... et des freins limitants la compétitivité de la culture

La perte de compétitivité du tournesol peut avoir plusieurs origines. La fréquence du tournesol dans les rotations est stable dans le sud-ouest, en Poitou-Charentes et est en augmentation dans des

systèmes de cultures biologiques en rotation longue du sud de la France. En revanche la fréquence du tournesol dans les rotations longues est en forte baisse depuis plusieurs années dans le Centre, la Bourgogne et en Rhône-Alpes, où il a perdu plus de 50% de sa surface cultivée (Lecomte et Nolot, 2011). Du fait de sa rusticité, le tournesol est cultivé dans des parcelles où les conditions pédoclimatiques sont difficiles.

Sa perte de compétitivité est également liée aux attaques de ravageurs (insectes, mammifères, oiseaux, gastéropodes) qui se manifestent en début de cycle (Figure I.5), mais aussi aux maladies causées par des champignons qui se manifestent à partir de la phase végétative (Figure I.5).

Le changement climatique a un impact direct et indirect sur le tournesol car il agit sur sa physiologie et la reproduction de ses champignons phytopathogènes (Debaeke *et al.*, 2017b). Le changement climatique est caractérisé par une augmentation de la concentration atmosphérique en CO_2 (gaz à effet de serre), des périodes de précipitations qui sont défavorables au tournesol, des sécheresses et de hautes températures en été (Ainsworth *et al.*, 2008). Les épisodes pluvieux et l'augmentation de la concentration atmosphérique en CO_2 au cours de la phase végétative entraînent une augmentation de la production de biomasse et donc (du développement) du couvert végétal (Sims *et al.*, 1999 ; Cheng *et al.*, 2000 ; Ainsworth *et al.*, 2008). Cette augmentation de la surface foliaire facilite l'infection du tournesol par des champignons contaminant celui-ci par les feuilles. Au contraire, la sécheresse au cours de la phase végétative a un effet positif sur les champignons phytopathogènes. Cependant, elle peut avoir un effet négatif si cette sécheresse se manifeste à partir de la floraison puisqu'elle entraîne des diminutions du rendement de la culture (Connor et Hall, 1997 ; Chimenti *et al.*, 2001, 2002 ; Ahmad *et al.*, 2014) et de la concentration des graines en huile (Andrianasolo *et al.*, 2014, 2016).

Enfin, la monoculture de tournesol était possible avant 2005 (`https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000632052`) mais son retour fréquent a favorisé l'installation de champignons phytopathogènes (Mestries *et al.*, 2011).

2.4.1 Les bioagresseurs principaux

Quatre champignons phytopathogènes sont à l'origine des maladies les plus importantes (en fréquence et sévérité , Gulya *et al.*, 1997) sur le tournesol en France : le mildiou, le sclérotinia, le phomopsis et le phoma (Aubertot *et al.*, 2005). Leurs principales caractéristiques sont décrites Tableau I.1.

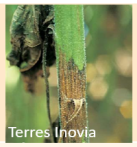
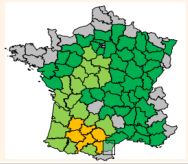
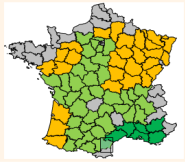
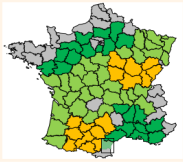
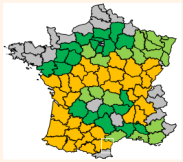
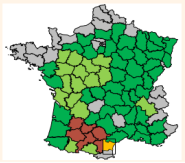
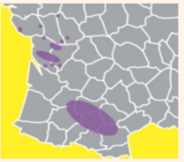
Nom		**Mildiou** *Plasmopara halstedii*	**Sclérotinia** *S. sclerotiorum*	**Phomopsis** *Diaporthe helianthi / P. helianthi*	**Phoma** *Leptosphaeria lindquistii / P. macdonaldii*	**Verticillium** *Verticillium dahliae*	**Orobanche** *Orobanche cumana*
symptôme							
Type de bioagresseur		Champignon biotrophe à diffusion limitée	Champignon nécrotrophe à diffusion large (sauf au collet)	Champignon nécrotrophe à diffusion large	Champignon nécrotrophe à diffusion large	Champignon nécrotrophe à diffusion limitée	Plante parasite à diffusion large
Conservation de l'inoculum	Forme	Oospores, mycélium	Sclérotes	Mycélium	Mycélium	Microslérotes	Graines
	Support	Résidus, sol, semences	Résidus, sol	Résidus	Résidus	Résidus, sol	Sol
	Durée de survie	10 ans	7 à 10 ans	1 an	Au moins 3 ans	14 ans	Au moins 10 ans
	Structure de dissémination	Zoospores	Mycélium, ascospores	Ascospores	Ascospores, pycniospores	Mycélium	Graines
Polycyclisme		Oui	Non	Non	?	Non	Non
Hôtes		Tournesol, adventices (Bidens *sp*, Xanthium *sp*, centaurées)	400	Tournesol, Adventices (Xanthium *sp*)	Tournesol	Nombreux (< 350)	Tournesol
Conditions favorables à l'installation du champignon		50 mm de pluie dans les 10 jours encadrant le semis (eau libre), température du sol > 10°C	42h d'eau libre à la surface des fleurons ; optimum 18-25°C	HR > 90 % pendant 36h ; optimum 25°C	Présence d'eau libre	20-27°C optimum 24°C	Présence d'eau libre Conditions encore méconnues
Occurrence Nuisibilité potentielle		Printemps pluvieux Forte	Tous les ans Forte	Tous les ans Forte	Tous les ans Forte	Emergent Forte	Emergent Forte
							

Tableau I.1 – Récapitulatif des caractéristiques des maladies provoquées par les principaux champignons phytopathogènes et de *Orobanche cumana* sur le tournesol (d'après Desanlis, 2013 ; Jestin, 2016)

Le mildiou (causé par *Plasmopara halstedii*) a été décrit pour la première fois en 1966 en France. Il est présent sur toutes les zones de production du tournesol. La maladie se caractérise par une destruction de la plante (pour une contamination à la levée ou la phase végétative), ou un nanisme avec des taches chlorotiques (Tableau I.1) et un feutrage blanc sur la face inférieure des feuilles pour les plantes survivantes. Lorsque les attaques sont précoces, les pertes de rendement sont proches du taux de plantes nanifiées (Meliala *et al.*, 1999). Lorsque les attaques sont tardives, les pertes s'élèvent à environ 0,5% du potentiel de rendement par pourcentage d'attaque (Meliala *et al.*, 2000). Son évolution en France est caractérisée par l'apparition rapide de nouvelles races présentant des mécanismes de virulence, contre lesquels il n'existe pas de lutte chimique curative (Tourvieille de Labrouhe *et al.*, 2000). La lutte contre le mildiou se fait en combinant l'utilisation de semences traitées et la mise en place de méthodes préventives non chimiques.

Le sclérotinia causé par *Sclerotinia sclerotiorum* (Lib. de Bary) contamine le tournesol depuis qu'il est cultivé en France. Il n'est pas spécifique au tournesol (Purdy, 1979) mais représente une de ses maladies majeures. Il est présent dans toutes les zones de cultures de la France et particulièrement dans les régions humides (Centre et Poitou-Charentes). Il peut attaquer tous les organes de la plante, sur lesquels il va provoquer un pourrissement blanc (Tableau I.1). Il se conserve sous la forme de sclérotes. Une perte de 98% du rendement est enregistrée si la culture de tournesol est attaquée à la floraison et moins de 12% si le tournesol est attaqué après la floraison (Masirevic et Gulya, 1992). La lutte génétique est essentielle contre le sclérotinia (Parts *et al.*, 2003).La lutte chimique est inefficace (Lamarque *et al.*, 1985). En revanche, la lutte biologique a fait ses preuves avec le produit Contans® W.G., une spécialité de biocontrôle à base de spores de champignons (Mestries *et al.*, 2011).

Le phoma (*Phoma macdonaldii* Boerema, forme téléomorphe *Leptospharia lindquistii*) a été étudié pour la première fois en 1976 en France. Il s'est étendu à toutes les zones de production du tournesol à partir de 1990 (Penaud *et al.*, 1994 ; Pérès et Lefol, 1996). Il peut toucher tous les organes de la plante. La maladie se caractérise par des lésions le long des nervures lui donnant un aspect de « patte d'oie », des taches noires sur la tige et sur le capitule et un dessèchement précoce des plantes suite aux attaques au collet. Les pertes pour les attaques sur tige peuvent atteindre 10 à 15% du rendement potentiel. Des pertes de 0,5 q.ha^{-1} ont été enregistrées pour 10% de plantes attaquées au collet (Penaud *et al.*, 1998), la forme la plus dommageable de la maladie (Gulya *et al.*, 1984 ; Donald *et al.*, 1987 ; Alignan *et al.*, 2006). Les pertes de rendement peuvent atteindre 13 q.ha^{-1} (Pérès *et al.*, 2000). Il n'y pas de traitement fongicide efficace et il n'existe pas de classement

de résistance variétale. La lutte agronomique est possible et sera exposée Partie IV.

Le phomopsis (*Phomopsis helianthi*, forme téléomorphe *Diaporthe helianthi*) a été observé pour la première fois en 1984 en France (Perny, 1992). Il s'est d'abord installé dans le sud-ouest avant de toucher toutes les zones de production du tournesol. Le champignon contamine le tournesol par les hydathodes puis progresse le long des veines, entraînant l'apparition de taches marrons. Il progresse jusqu'au pétiole puis atteint la tige où il fait une tache brune-rouge encerclante mortelle. Les pertes de rendement s'élèvent à 1 à 3 q.ha^{-1} et un point de teneur en huile lorsque 10% des plantes ont des taches encerclantes (Estragnat, 1992 ; Pinochet, 1994 ; Jouffret, 2005). À ce jour, il existe différentes méthodes de lutte (agronomique, génétique et chimique) qui permettent de maîtriser ce champignon (Gulya *et al.*, 1997). Elles seront exposées dans le Partie IV.

2.4.2 Les bioagresseurs secondaires

Cinq autres champignons sont responsables de maladies entraînant peu de dommages (Chollet *et al.*, 2002) : la verticilliose a été identifiée en 1971 dans le Sud-Ouest où elle s'est bien établie, mais son incidence est pour le moment limitée. La rouille blanche (causée par *Pustula helianthicola*), qui engendre des pertes de rendement restant faibles (Gulya *et al.*, 1997 ; entre 2 et 5 q.ha^{-1}) ; l'alternariose (*Alternaria helianthi* et *Alternaria helianthinficiens*), qui peut provoquer une diminution de la masse des graines et de la teneur en huile ; le botrytis (*Botrytis cinerea*) qui peut entraîner des pertes sur capitule n'excédant pas 1% ; et le macrophomina (*Macrophomina phaseolina*).

L'orobanche (*Orobanche cumana*) auparavant considérée comme secondaire, apparaît préoccupante depuis peu. L'orobanche (*Orobanche cumana*) est une plante parasitant exclusivement le tournesol. Connue depuis longtemps dans le pourtour méditerranéen, elle a été identifiée pour la première fois en 2007 en France (Jestin *et al.*, 2014) et elle commence à s'établir dans le Sud-Ouest (Tableau I.1, Jestin et Martin-Monjaret, 2016). Elle est hautement nuisible en cas d'infestation importante (pertes de rendement pouvant atteindre 100%, Dominguez, 1996 ; Jestin *et al.*, 2014). Cette plante est dépourvue de chlorophylle et vit aux dépends du tournesol. Elle le parasite par les racines, puis développe une hampe florale. Les graines, de très petite taille (200 à 300 μm), sont aisément disséminées par le vent, la pluie, les machines agricoles, les animaux et les semences de tournesol. Actuellement, seules les luttes génétique et chimique (herbicides) sont efficaces, mais non durables. Les méthodes de lutte agronomique sont encore à préciser.

2.5 Les bioagresseurs considérés

2.5.1 Choix des bioagresseurs étudiés

Parmi les principaux bioagresseurs présentés précédemment, trois ont été retenus pour cette thèse : le phomopsis, le phoma et l'orobanche. Le phomopsis et le phoma sont des maladies graves. Différentes connaissances ont été acquises sur ces maladies depuis leur apparition en France, notamment concernant les effets des pratiques agricoles sur leur développement. Des modèles peuvent être mis en place pour synthétiser ces connaissances. De plus, de nombreuses expérimentations ont été menées dans différentes régions et ont permis de collecter des jeux de données exploitables pour l'évaluation des modèles portant sur le phoma et le phomopsis pour la France. En revanche, trop peu d'observations étaient disponibles pour évaluer les autres principaux bioagresseurs (mildiou, sclérotinia). L'orobanche fait l'objet de recherches scientifiques et techniques car son aire de répartition est en expansion. Il est donc important de générer des connaissances à son sujet. Pour cela, une synthèse des connaissances sur les effets des pratiques agricoles et de la situation de production sur le développement de l'orobanche a été réalisée par Terres Inovia (Jestin, 2016) pour identifier les pratiques culturales et les variables environnementales dont il faut prioriser les études.

2.5.2 Cycle de vie des bioagresseurs étudiés

Les champignons

Diaporthe helianthi (responsable du phomopsis) et *Leptosphaeria lindquistii* (responsable du phoma) sont deux champignons phytopathogènes de la famille des Ascomycètes (embranchement des Mycètes caractérisé par la formation d'ascospores) (Tableau I.2). Leur cycle de vie est composé d'une phase asexuée et d'une phase sexuée. La reproduction asexuée est réalisée par les pycnides (structure de reproduction asexuée chez les ascomycètes qui libèrent des pycniospores (conidies). Les pycniospores de ces deux champignons sont connues pour ne propager la maladie qu'à de très faibles distances (Délos et Moinard, 1997). Cette phase est considérée comme négligeable dans l'épidémiologie de ces deux champignons (Mihaljcevic *et al.*, 1985 ; Délos et Moinard, 1997) et n'a pas été prise en compte au cours de cette thèse. La phase sexuée est réalisée par les pseudothèces (pour le phomopsis) et les pseudothèces (pour le phoma) qui libèrent des ascosporess. Les pseudothèces et pseudothèces sont des structures (Tableau I.2, Figures I.6 A et B) qui commencent leur développement sous forme de proto-pseudothèces/pseudothèces sur les cannes de tournesol à l'automne (Muntanola-Cvetkovic *et al.*, 1988). Leur développement s'interrompt en hiver et reprend à la fin de l'hiver lors de conditions climatiques (pluviométrie et température moyenne) favorables.

Les pseudothèces développent des asques pouvant contenir huit ascosporess. Ces dernières sont projetées dans l'air au cours d'évènements pluvieux de mai à août (Petrov *et al.*, 1981 ; Fayret et Assemat, 1987), puis sont disséminées par le vent et la pluie sur de longues distances (Maric *et al.*, 1988) et sont à l'origine de nouvelles infections (Figure I.6 A et B).

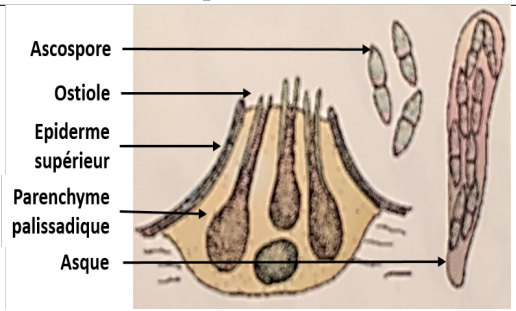
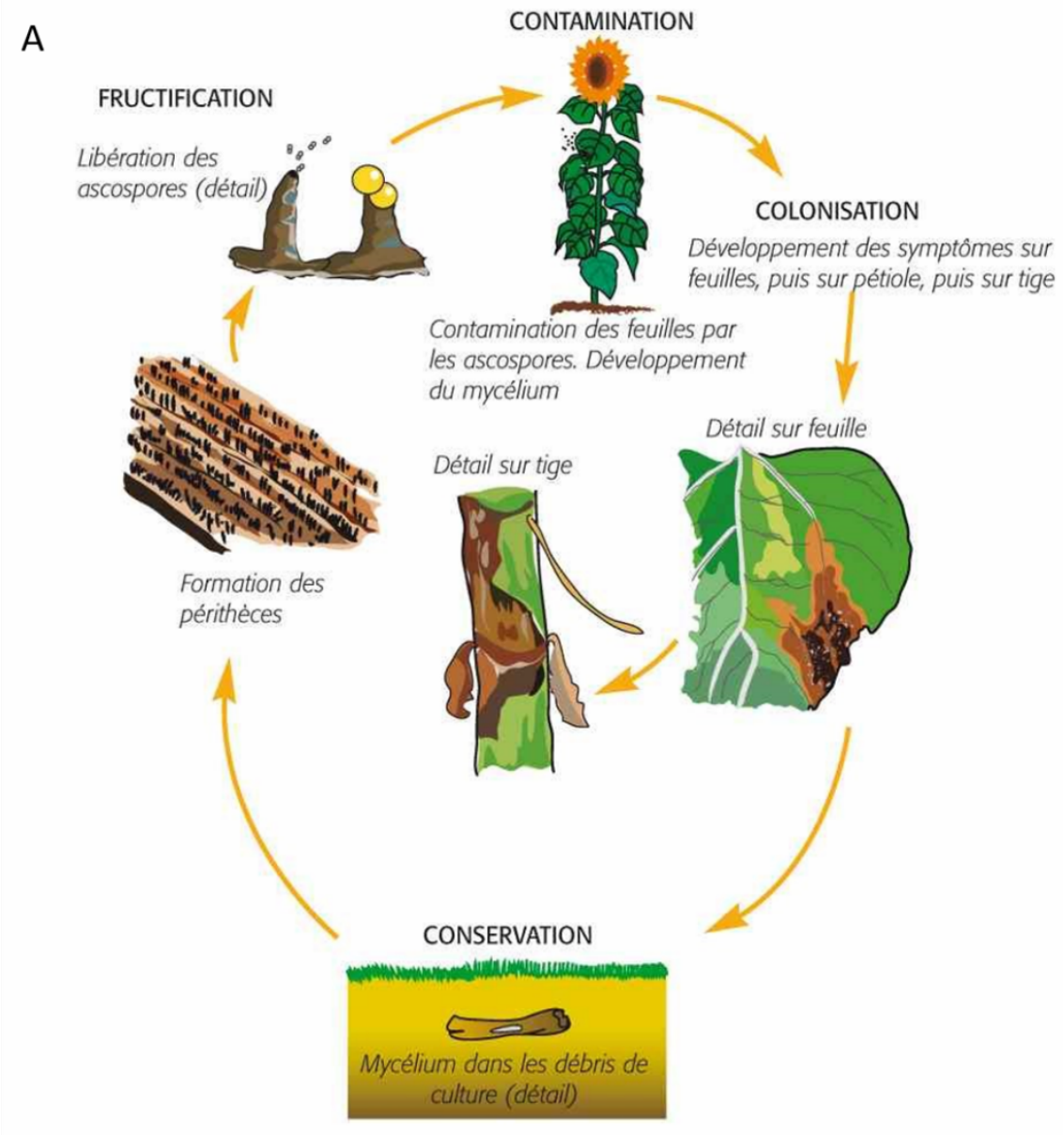
Forme sexuée	***Diaporthe helianthi*** **Munt-Cvet. (1981)**	***Leptosphaeria lindquistii*** **Frezzi (1968)**
Embranchement	Ascomycota	
Sous-embranchement	Pezizomycotina	
Classe	Sordariomycetes	Dothideomycetes
Ordre	Diaporthales	Pleosporales
Famille	Diaporthalaceae	Leptosphaeriaceae
Organe de fructification		

Tableau I.2 – Classification systémique de *D. helianthi* et *L. lindquistii* (d'après Agrios, 1936 ; Blackwell *et al.*, 2012)


A
CONTAMINATION
FRUCTIFICATION
Libération des ascospores (détail)
COLONISATION
Développement des symptômes sur feuilles, puis sur pétiole, puis sur tige
Contamination des feuilles par les ascospores. Développement du mycélium
Détail sur feuille
Détail sur tige
Formation des périthèces
CONSERVATION
Mycélium dans les débris de culture (détail)

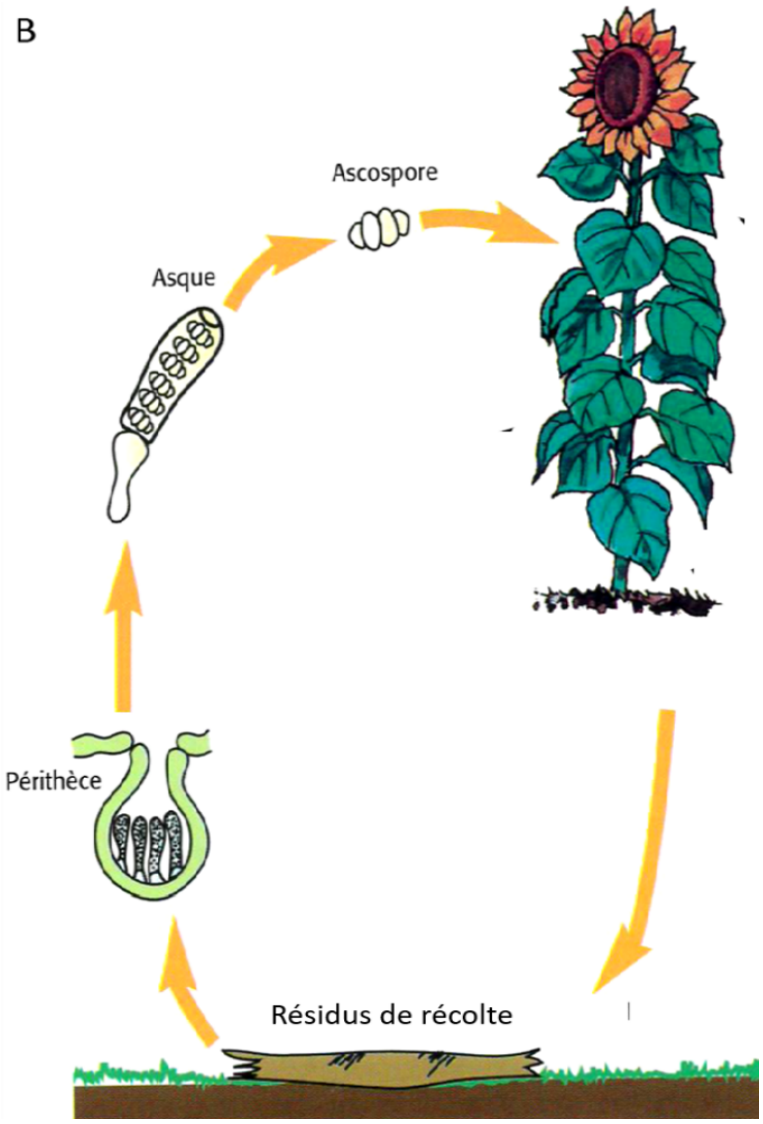


Figure I.6 – Cycle de vie de : *D. helianthi* (A) et *L. lindquistii* (B) (d'après Maric *et al.*, 1988 ; Chollet *et al.*, 2002)

La plante parasite

Orobanche cumana réalise sa reproduction sexuée en développant une hampe florale portant plusieurs capsules pouvant libérer chacune plusieurs centaines de graines qui seront aisément dispersées de manière biotique (par les outils de travaux agricoles, par des animaux traversant les parcelles) et abiotiques (par le vent et la pluie) (Tableau I.3 et Figure I.7).

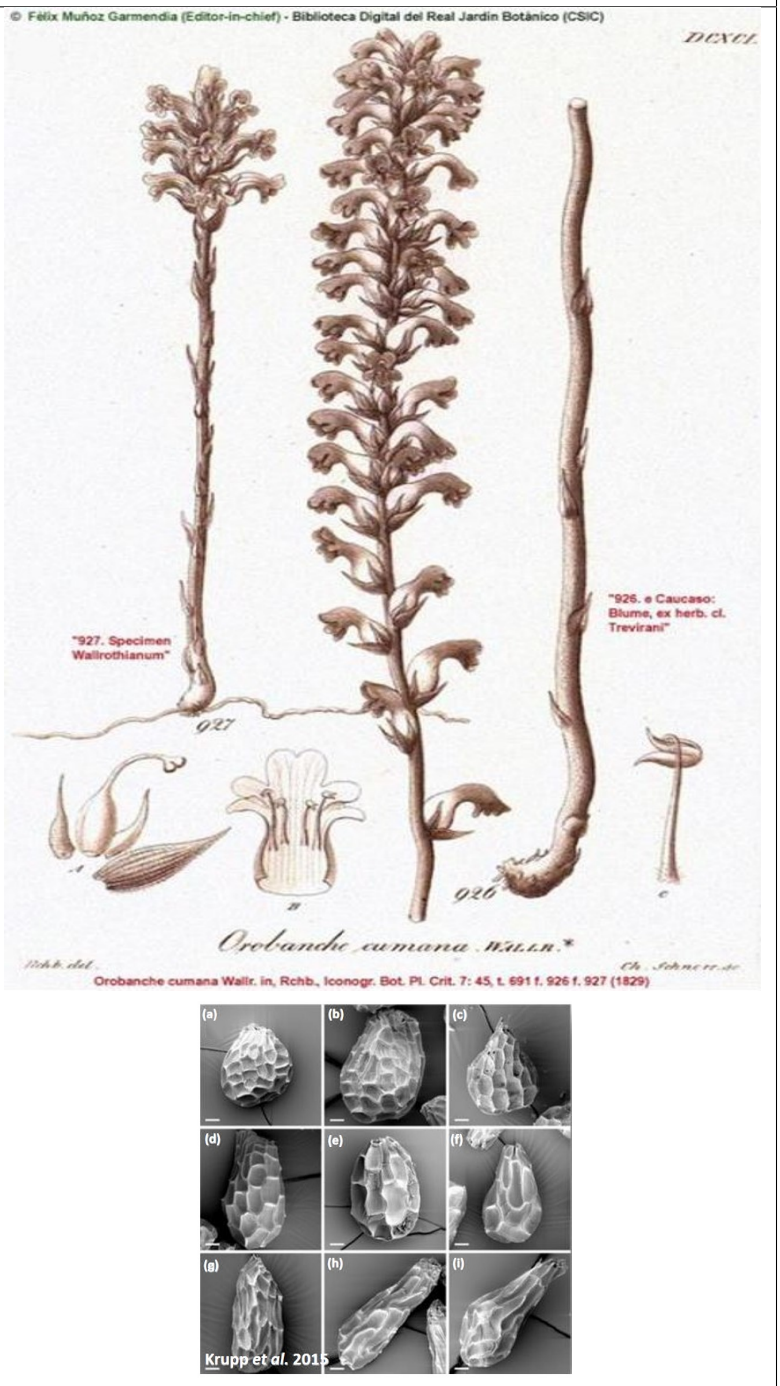
Forme sexuée	***Orobanche cumana*** **Wallr. (1825)**
Clade	Angiospermes, Eudicotylédones, Astéridées, Lamidées
Ordre	Lamiales
Tribu	Orobanchaceae
Plante et graine	

Tableau I.3 – Classification systémique de *O. cumana* (d'après Krupp *et al.*, 2015

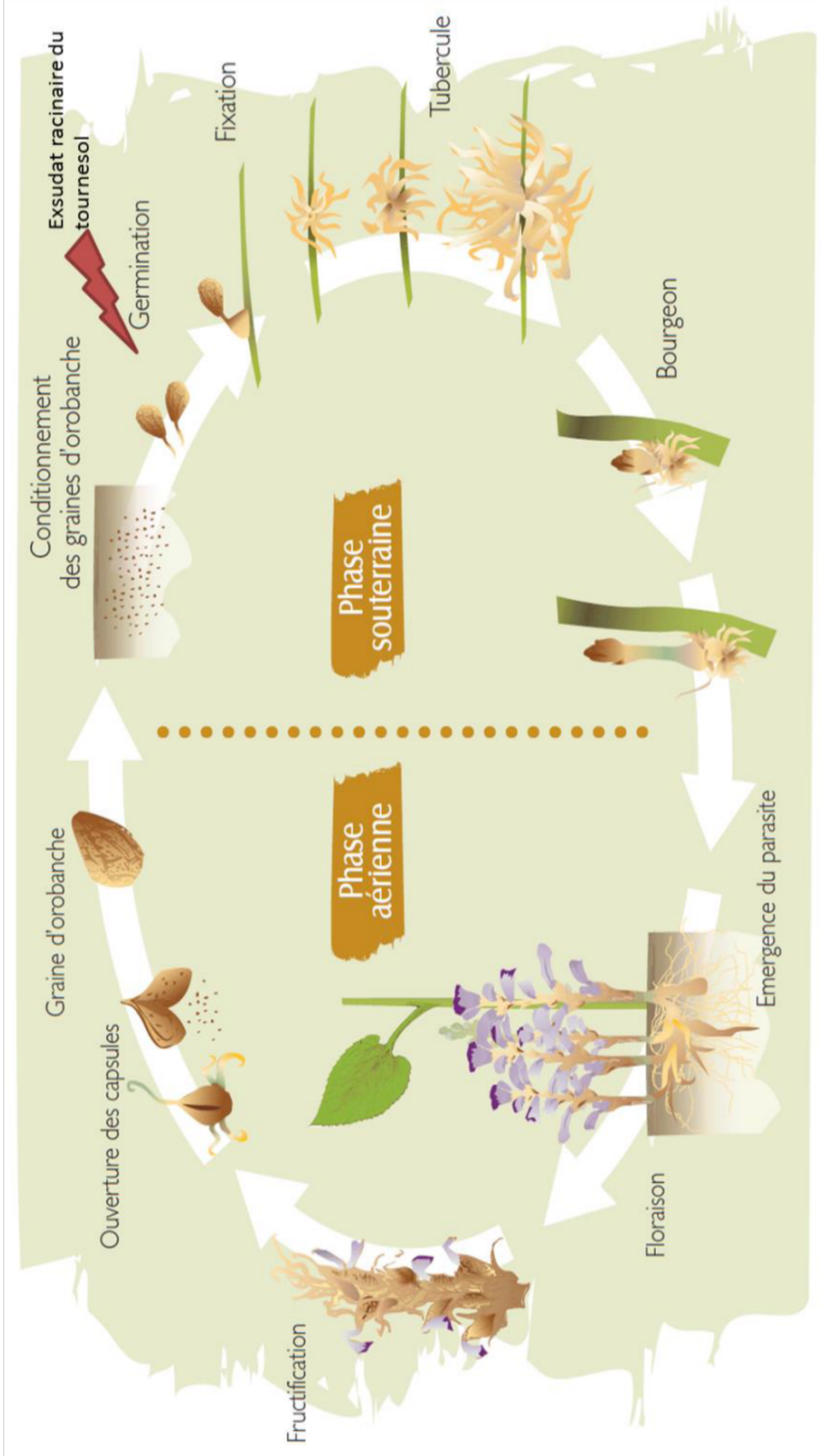


Figure I.7 – Cycle de vie de *O. cumana* (d'après Maric *et al.*, 1988 ; Chollet *et al.*, 2002, Terres Inovia)

La section suivante présente les différentes méthodes de protection (préventives et curatives) contre les bioagreseurs du tournesol.

3. Les méthodes de protection de cultures : la PIC et la PAEC...

3.1 La Protection Intégrée des Cultures

Les méthodes de protection des cultures ont évolué ces dernières décennies (Deguine *et al.*, 2016). Le concept de **P**rotection **I**ntégrée des **C**ultures (PIC) a joué un rôle essentiel dans ces évolutions. La réduction de l'utilisation de pesticides doit s'accompagner de l'utilisation de méthodes préventives pour protéger les cultures.

Face à un bioagresseur donné, l'application des principes de la PIC a pour but de perturber son cycle de quatre manières (Figure I.8) :

1. en réduisant les éléments infectieux de conservation des bioagresseurs (inoculum pour les champignons, stock grainier pour les adventices). La réduction des éléments infectieux de bioagresseurs peut être réalisée grâce à de nombreuses méthodes visant soit à accélérer la dégradation du support sur lequel le bioagresseur effectue une partie de son cycle de vie, à détruire directement le bioagresseur ou à diminuer sa viabilité. Cette réduction peut être réalisée par exemple par une rotation longue, le broyage de résidus, le travail du sol ou la gestion des repousses. Le contrôle de la qualité sanitaire des semences limite l'introduction d'éléments infectieux dans une parcelle.
2. en empêchant l'installation de la maladie (stratégie d'évitement).
3. en limitant la contamination en culture, la multiplication du bioagresseur et l'expression de la maladie, c'est-à-dire le développement des symptômes (dégâts) et des pertes de récolte (dommages). L'atténuation en culture peut se faire d'abord par un choix de variété résistante (3a sur la Figure I.8), qui ne subira pas de dégâts ou bien tolérante (3b sur la Figure I.8), c'est à dire que les dégâts n'entraîneront pas de dommage significatif. L'association de différentes

espèces ou de variétés peut agir par allélopathie sur la contamination et la multiplication du bioagresseur. Enfin, l'adaptation de la densité de semis, de la gestion de l'azote et de l'irrigation peut agir sur le micro-climat favorable aux maladies aériennes.

4. En mettant en place des opérations de rattrapage (4) pour limiter la multiplication du bioagresseur une fois celui-ci installé dans la parcelle. Il peut s'agir de lutte mécanique contre les adventices ou de lutte par biocontrôle ou de lutte chimique. Cette dernière est à utiliser en dernier recours, lorsque toutes les précédentes méthodes n'ont pas abouti.

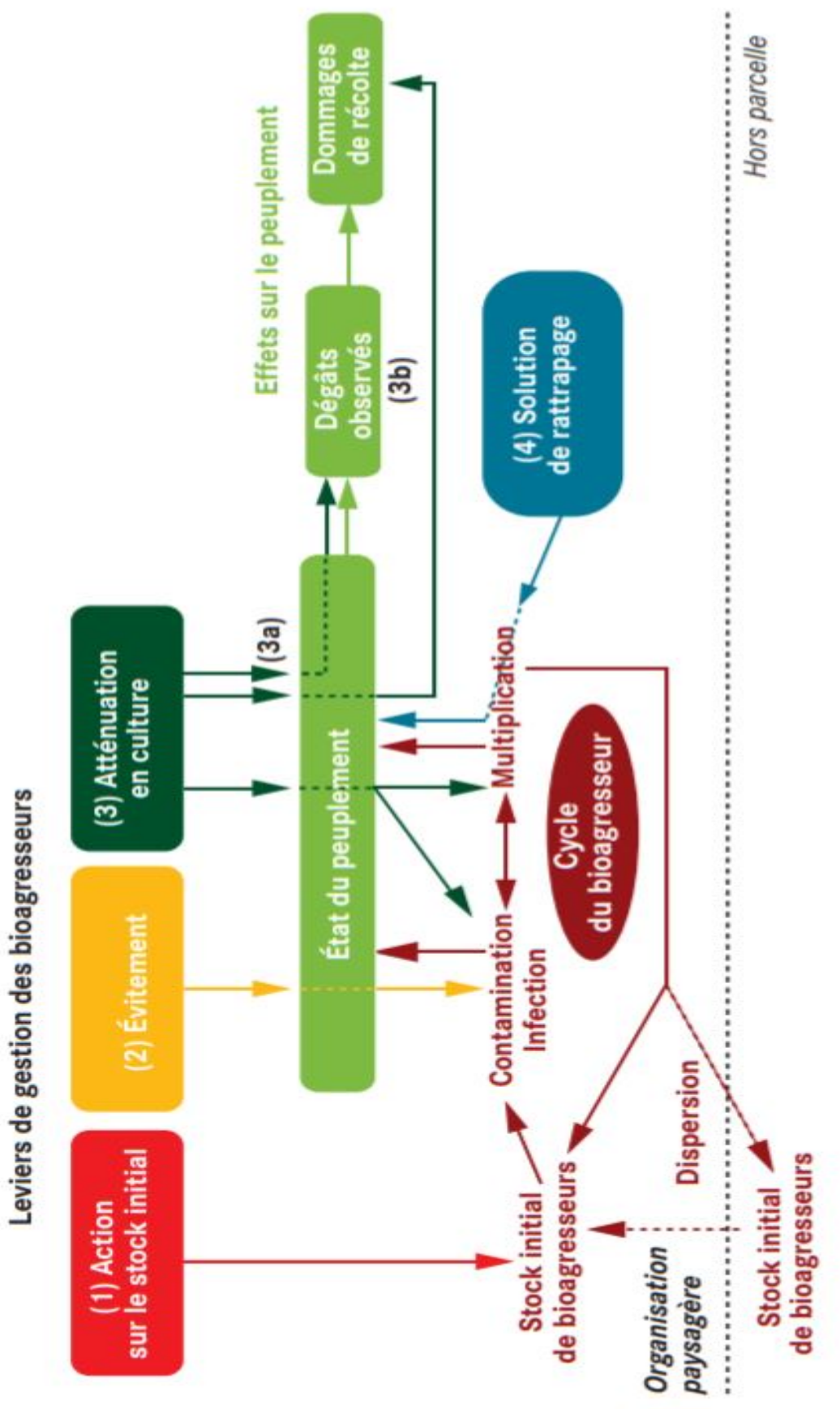


Figure I.8 – Schéma générique de la protection d'une culture face à un bioagresseur (Attoumani-Ronceux *et al.*, 2010)

3.2 La Protection AgroEcologique des Cultures

La PAEC (Figure I.9) correspond à la déclinaison des principes de l'agroécologie à la protection des cultures. Elle repose sur un raisonnement écologique en vue d'un fonctionnement durable des agroécosystèmes. La PAEC vise à favoriser le fonctionnement écologique des agroécosystèmes, en optimisant directement ou indirectement les interactions entre les communautés vivantes (végétales, animales et microbiennes). La santé du sol et la biodiversité des agroécosystèmes sont les deux axes principaux de la PAEC.

Elle met en avant les pratiques culturales qui favorisent l'établissement et l'équilibre de la faune et de la flore locales défavorables aux bioagresseurs. La PAEC vise à perturber le cycle de vie des bioagresseurs. et à établir le fonctionnement équilibré des peuplements cultivés et leur résilience vis-à-vis des stress biotiques, et à réduire les risques et l'intensité des infestations, infections, pullulations, épidémies et invasions de bioagresseurs.

Ces pratiques culturales doivent faire appel à des connaissances fines du fonctionnement des agroécosystèmes. Les connaissances mobilisées ne s'arrêtent pas au seul pathosystème considéré et doivent prendre en compte les différentes échelles spatio-temporelles dans lesquelles se trouve le pathosystème.

La PAEC touche donc à plusieurs domaines. L'agronomie rassemble un ensemble de disciplines étudiant les relations entre la culture, les pratiques culturales et le pédoclimat. L'écologie est une discipline englobant les études du fonctionnement des organismes, des populations et des écosystèmes. Ces deux domaines forment l'agroécologie. La protection agroécologique vise aussi à appliquer les concepts de la biologie de conservation. Cette dernière est un ensemble de disciplines étudiant les questions de perte, de maintien et de restauration de la biodiversité. D'un point de vue seulement biologique, elle recoupe donc les domaines de l'écologie en s'appuyant particulièrement sur trois disciplines, qui sont la biogéographie , l'écologie du paysage et la génétique des populations. La biologie de la conservation s'intéresse aussi aux questions économiques, sociales et politiques (non prises en compte dans cette thèse). Les combinaisons de techniques issues de la PIC et des savoirs issus des deux disciplines que sont l'agroécologie et la biologie de la conservation constitue la base de la PAEC. Ainsi la lutte chimique n'est utilisée qu'en cas d'extrême nécessité, après évaluation des risques grâce, entre autres, aux OAD. D'une manière générale, la mise en œuvre de la PAEC vise donc la durabilité des agrosystèmes (Deguine *et al.*, 2016).

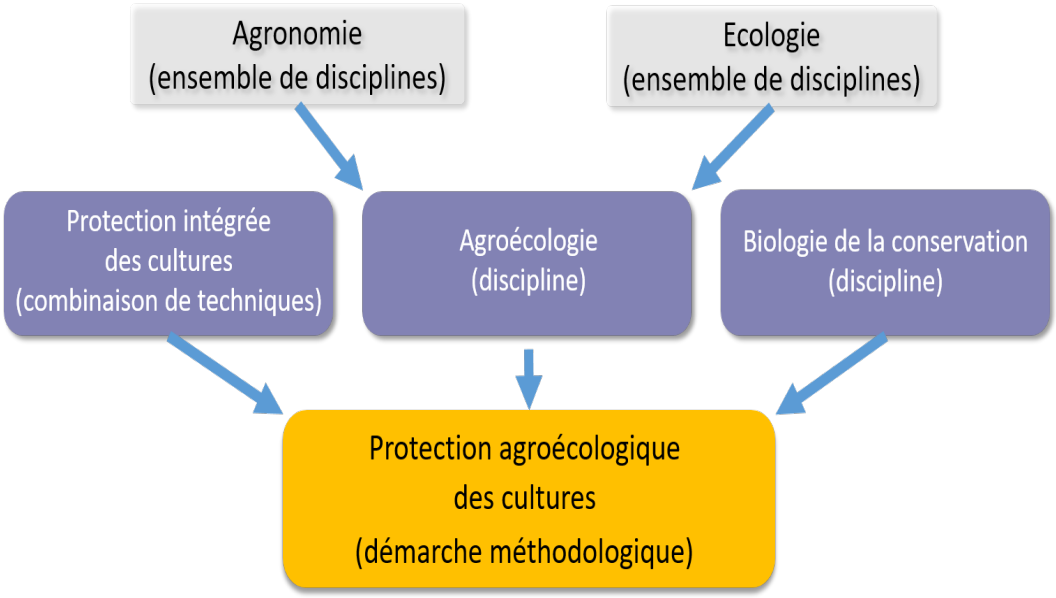


Figure I.9 – Origines de la PAEC (Deguine *et al.*, 2016)

3.3 Les méthodes de lutte contre les bioagresseurs d'intérêt

Les méthodes évoquées Figure I.8 pour la PIC et la PAEC doivent être adaptées à chaque bioagresseur (Deguine *et al.*, 2016). Les effets de certaines méthodes sont déjà identifiés et sont présentés Tableau I.4. Une description détaillée de ces méthodes pour chaque bioagresseur sera présentée Partie IV.

Composante du système	**Phoma**	**Phomopsis**	**Orobanche**
Génétique	+	+++	+++
Chimique	++	+++	+++
Physique	-	-	-
Biologique	-	-	-
Agronomique	++	++	+

Tableau I.4 – Évaluation de l'effet de composantes du système contre le phoma, le phomopsis et l'orobanche du tournesol (d'après Debaeke *et al.*, 2017b). + : Défavorable à la maladie, - : Favorable à la maladie

Toutes les méthodes de lutte n'ont pas la même efficacité contre ces trois bioagresseurs (Tableau I.4). La lutte génétique peut être employée contre le phomopsis et l'orobanche, par contre elle n'a pas permis le développement de variétés résistantes ou tolérantes contre le phoma. La lutte

chimique est efficace contre les trois bioagresseurs. La lutte physique inclue toutes les techniques dont le mode d'action primaire ne fait intervenir aucun processus biologique ou biochimique (Panneton *et al.*, 2000). Elle regroupe quatre types de lutte, les méthodes thermiques, électromagnétique, mécaniques et pneumatiques. En revanche, à l'heure actuelle, il n'existe pas d'agent de lutte biologique. La lutte agronomique a des effets moyennement marqués sur le phoma et le phomopsis. Par exemple, le broyage et l'enfouissement des résidus sont défavorables pour les deux maladies, tandis qu'un semis précoce et l'irrigation les favorisent. Pour l'orobanche, les effets des différentes pratiques culturales sont encore mal connus.

4. Rôle de la modélisation en protection des cultures

La mise en oeuvre des principes de l'agroécologie sur le tournesol se heurte à plusieurs difficultés d'ordres agronomique, technique et socio-économique. En effet, les conséquences d'un changement de pratique(s) culturale(s) sur le profil de dégâts (défini comme un ensemble de symptômes causés par un cortège de bioagresseurs) de la culture sont difficiles à prévoir (Savary *et al.*, 2006) puisqu'une technique peut défavoriser un bioagresseur et en favoriser un autre. Il est nécessaire d'appréhender la complexité du profil de dégâts dans son ensemble pour concevoir des itinéraires techniques et des systèmes de culture adaptés. Pour ce type d'approche complexe, il n'est pas possible de mettre en place des expérimentations pour étudier chaque interaction possible. Le suivi de parcelles au sein d'essais « systèmes » permet de générer des connaissances sur les interactions les plus courantes (par exemple, l'effet d'un niveau d'azote dans un contexte pédoclimatique sur un bioagresseur). La modélisation est un outil précieux pour synthétiser ces connaissances sur les maladies du tournesol, quantifier les effets des pratiques agricoles sur la sévérité de celles-ci et pour aider au développement d'outils d'aide à la conception de systèmes de culture.

Les modèles conçus pour les bioagresseurs sont pour la plupart des modèles épidémiologiques servant de base pour le conseil aux agriculteurs concernant le raisonnement de l'application de pesticides (Limasset et Darpoux, 1950 ; Aujas *et al.*, 2011). Ce sont des modèles pour la prévision de risques de maladies, de ravageurs ou d'adventices en fonction de variables climatiques. Peu de variables de pratiques culturales interviennent dans ces modèles. Ils servent à connaître les périodes de réceptivité de la culture et la période à laquelle l'inoculum peut créer des infections.

De manière générale, du fait du grand nombre de bioagresseurs auxquels une culture peut faire

face (champignons phytopathogènes, plantes adventices et ravageurs), peu de modèles abordent à la fois une culture et un profil de bioagresseurs (Aubertot *et al.*, 2005). En 2014, une nouvelle méthode, IPSIM (Aubertot et Robin, 2013 ; Robin *et al.*, 2013) a vu le jour pour répondre au besoin de modélisation des effets des systèmes de culture, du pédoclimat et du paysage sur le profil de dégâts. Ces modèles considèrent dans un premier temps les bioagresseurs de manière individuelle mais il est également possible de considérer plusieurs bioagresseurs en interaction pour une culture (Aubertot et Robin, 2013). Ces modèles ont pour vocation de mettre à disposition des acteurs du conseil agricole les connaissances disponibles sur les déterminants des bioagressions afin d'aider à la conception de stratégies de gestion durables des stress biotiques.

Les bioagresseurs présents dans une culture donnée dépendent du système de culture et de la situation de production. Un profil de bioagresseurs va créer un profil de dégâts. Si la sévérité de ces dégâts est trop élevée, des dommages se manifesteront (Figure I.10, se traduisant par une baisse de la quantité et/ou de la qualité des produits).

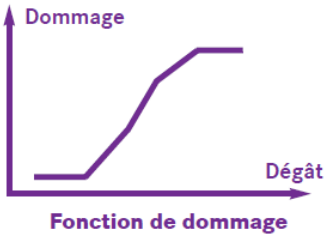


Figure I.10 – Représentation d'une fonction de dommage (Aubertot *et al.*, 2010)

Les pertes de rendements font aussi l'objet de modélisation. Ces modèles peuvent porter sur les pertes de rendements en ne considérant que les stress abiotiques dans une situation de production ou considérer les stress biotiques subis par la culture.

4.1 Exemples de modèles pour quelques cultures

4.1.1 Les modèles épidémiologiques

Les modèles épidémiologiques sur champignons servent à prédire les dynamiques des taux de maturation des fructifications et des ascosporess sexuées de champignons et les éjections d'ascosporess. Nous pouvons citer le modèle historique mis en place pour *Venturia inaequalis* (Gadoury et MacHardy, 1982). Il prédit la dynamique du taux de maturation des ascosporess en fonction de la somme de températures calculée à partir du 15 mars et n'a que deux paramètres. Ce modèle a été adapté aux conditions climatiques de plusieurs régions françaises par le service de protection des végétaux (Bouveris, 1980). Il a également été repris pour d'autres pathogènes

(Lovell *et al.*, 2004). Ce modèle ne prend pas en compte la pluviométrie qui est indispensable pour la reproduction sexuée de *D. helianthi* (Petrov *et al.*, 1981 ; Fayret et Assemat, 1987) et qui est supposée intervenir pour les Ascomycètes du genre *Leptosphaeria* (Rapilly, 1991).

La fusariose de l'épi a aussi fait l'objet d'un modèle (Corre *et al.*, 2015). Celui-ci calcule la maturation des périthèces et des ascosporess, ainsi que les taux d'éjections en fonction de la température et de la pluviométrie journalières. Il a 8 paramètres. Mais il définit un jour comme étant humide de manière approximative pour s'affranchir de l'humidité qui était difficile d'accès aux concepteurs du modèle.

Ces modèles peuvent sembler simples mais ils présentent des faiblesses pour nos cas d'étude. Par exemple, le besoin de données horaires peut rendre difficile l'utilisation d'un modèle car le nombre de stations donnant accès à ce type de données en France diminue. Un deuxième défaut est le recalcul de variables climatiques à partir d'autres variables, ce qui introduit des incertitudes. Enfin, les modèles sont composés de paramètres parfois dépourvus de sens biologique.

Le modèle SimMat (Aubertot *et al.*, 2006) ne présente pas ces faiblesses. Il a été conçu pour prédire les dynamiques des taux de maturation de pseudothèces et des éjections d'ascosporess de phoma du colza en fonction de la température moyenne journalière et de la pluviométrie journalière. Il ne mobilise que six paramètres ayant tous du sens du point de vue biologique. Il a été construit à l'origine pour prédire la reproduction sexuée du phoma du colza en Australie (Salam *et al.*, 2003). Ce modèle a été adapté aux conditions climatiques polonaises (Aubertot *et al.*, 2006), puis françaises (Aubertot et Mestries, communication personnelle). Il est facile à renseigner car il a été construit pour des zones australiennes où l'accès à des données climatiques fines (de type horaires) n'était pas possible.

4.1.2 Les modèles de dégâts et de dommages

La plateforme de modélisation IPSIM a permis le développement de différents modèles dont certains pour le blé (Robin *et al.*, 2013) ou encore la chayotte (publication en cours de préparation). Ces modèles intègrent les pratiques culturales, le contexte pedo-climatique et le contexte de la parcelle.

Enfin, parmi les modèles conçus pour analyser les pertes de rendement causées par plusieurs bioagresseurs, nous pouvons citer celui pour le riz (RICEPEST, Willocquet *et al.*, 2002), pour des agrosystèmes tropicaux (INFOCROP, Aggarwal *et al.*, 2006) et pour le blé (WHEATPEST, Willocquet *et al.*, 2008).

4.2 Les modèles pour le tournesol

4.2.1 Les modèles épidémiologiques

Un modèle de prédiction de risque a été conçu pour le phomopsis du tournesol. ASPHODEL (Délos et Moinard, 1997) prédit les dynamiques de maturation des pseudothèces de *D. helianthi*, les présences d'éjections et de contamination par le phomopsis. Il sert à déterminer l'utilité et le moment adéquat pour l'application de fongicides contre le phomopsis. Bien qu'il ait prouvé sa fiabilité (Desanlis *et al.*, 2013), il présente les défauts cités précédemment pour les modèles épidémiologiques d'autres cultures. Il est devenu difficile d'utilisation à cause de ses données horaires et nécessite d'être remplacé. Le modèle de remplacement doit prendre en compte la pluviométrie et la température moyenne, avoir le moins de paramètres possible et le plus grand nombre de ces paramètres doivent avoir un sens biologique. Pour cela, SimMat est un bon candidat. La fonction SimMat, le formalisme mathématique et informatique (sous R) d'ASPHODEL sont exposés dans les Annexes 4 et 5.

4.2.2 Les modèles de culture, les modèles de dégats et de nuisibilité

D'autres modèles existent pour le tournesol. STICS (Brisson *et al.*, 1998) prédit un nombre considérable de variables physiologiques et de variables du sol pour plusieurs cultures (dont le tournesol) en fonction du climat, du sol et des pratiques culturales. SUNFLO (Casadebaig *et al.*, 2011) est un modèle de culture dynamique qui prédit le rendement et la qualité des graines du tournesol en fonction de la variété de tournesol et des stress abiotiques qu'il subit. Mais ces modèles ne prennent pas en compte les stress biotiques.

Le tournesol n'a pas fait l'objet de modèle de dégâts évalué et validé. En ce qui concerne la modélisation de la nuisibilité, le modèle SUNFLOWER-PEST a été développé très récemment pour le tournesol à l'aide de la nouvelle plateforme X-PEST (`xpest.inra.fr`). Il s'agit d'une plateforme de modélisation générique permettant de développer des modèles dynamiques de nuisibilité qui prédisent le rendement d'une culture donnée en fonction des stress biotiques et abiotiques subis par la culture (Smits *et al.*, 2015). Bien que la plateforme X-Pest ait été mise au point pour permettre de mettre en place des modèles de dommages, ils n'ont pas été considérés dans cette thèse en raison du manque de connaissances permettant de modéliser les fonctions de dommages nécessaires pour faire le lien entre profil de dégâts et de dommages pour le tournesol.

5. Objectifs de la thèse

Cette thèse a porté sur plusieurs étapes du cycle épidémiologique des bioagresseurs. Les étapes de la production d'inoculum et de l'infection ont été considérées pour le phomopsis et le phoma, l'étape d'apparition des dégâts a été considérée pour les trois bioagresseurs.

Plusieurs objectifs sont visés dans cette thèse.

— La production de synthèses des connaissances disponibles dans la littérature scientifique et technique sur les effets d'un système de culture et de la situation de production sur chaque étape du cycle de vie des bioagresseurs d'intérêt. Ces synthèses devront servir de bases pour répondre aux objectifs suivants.

— L'adaptation du modèle générique épidémiologique SimMat aux deux champignons phytopathogènes donnant SimMat-*Diaporthe* et SimMat-*Leptosphaeria lindquistii*. Ces modèles visent à répondre à la question de la gestion de l'inoculum primaire.

— La construction et l'évaluation des modèles IPSIM - Sunflower - Phomopsis stem canker, IPSIM - Tournesol - Phoma et IPSIM - Tournesol - Orobanche pour prédire des dégâts causés par un bioagresseur en fonction du système de culture et de la situation de production. Ces modèles se basent sur les synthèses évoquées précédemment et un panel d'experts en phytopathologie, malherbologie, agronomie et modélisation. Les modèles IPSIM - Sunflower - Phomopsis stem canker, IPSIM - Tournesol - Phoma sont associés à des convertisseurs permettant de convertir des données facilement accessibles (nom de la variété semée, date de semis, par exemple). Le modèle IPSIM - Sunflower - Phomopsis stem canker a été évalué grâce à un jeu de données *ad hoc*. Ils permettront de répondre à terme à la question de la

gestion d'un ensemble de bioagresseurs de manière intégrée.

— Le développement de deux méthodes originales. La première permet le couplage de modèles quantitatif et qualitatifs, la deuxième permet l'optimisation de la qualité de prédiction de modèles mathématiques qualitatifs par hybridation de l'information contenue dans un jeu de données et des connaissances expertes. Ces méthodes sont appliquées aux modèles SimMat et IPSIM.

6. Organisation du manuscrit

La suite de ce manuscrit de thèse est organisée en quatre parties, présentées dans le schéma organisationnel ci dessous Figure I.11. La partie 4 est composée de 3 chapitres, pour chaque bioagresseur.

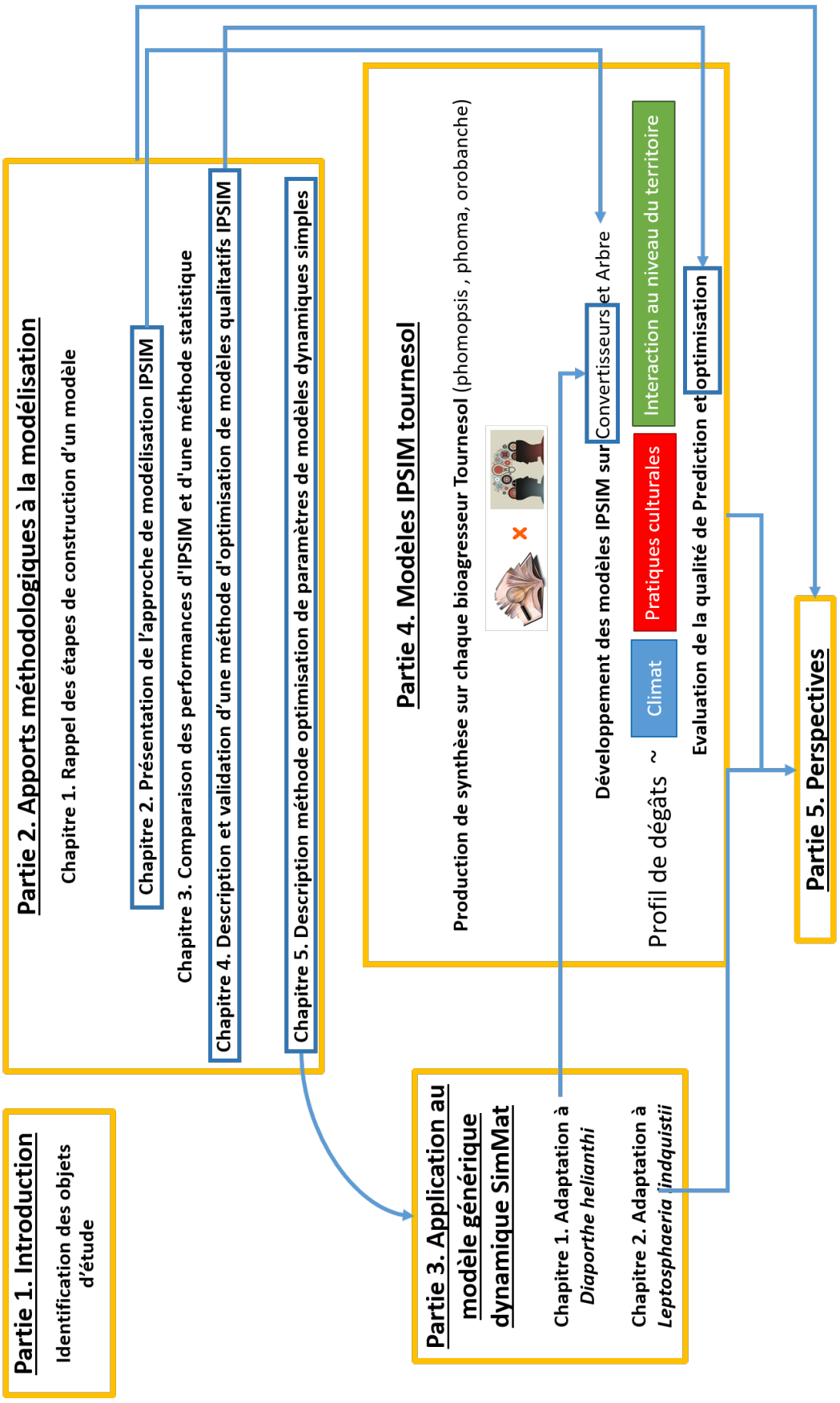


Figure I.11 – Schéma organisationnel du manuscrit

II Apports méthodologiques à la modélisation

1 Construction des modèles 71

1.1 Rappel des différentes étapes de construction d'un modèle

1.2 Les étapes de construction abordées au cours de la thèse

2 Présentation de l'approche de modélisation IPSIM 77

2.1 Le modèle générique IPSIM

2.2 Évaluation des modèles IPSIM dans la thèse

3 Comparaison des performances de l'approche IPSIM et d'une méthode statistique 83

3.1 Objectifs de l'étude

3.2 Matériel et méthodes

3.3 Résultats

3.4 Discussion

4 Méthode innovante améliorant la qualité prédictive des modèles IPSIM 91

4.1 Principe de la méthode

4.2 Validation informatique de l'algorithme améliorant la qualité prédictive des modèles IPSIM

5 Optimisation de paramètres de modèles dynamiques simples 101

1. Construction des modèles

1.1 Rappel des différentes étapes de construction d'un modèle

Qu'est-ce qu'un modèle ?

Le terme « modèle » peut faire référence à une large diversité d'objets. Un modèle se caractérise par sa nature (par exemple un modèle de type qualitatif ou quantitatif), mais aussi par le, ou les, processus qu'il représente (par exemple un modèle sur la dynamique entre une proie et son prédateur, un modèle de croissance de culture). Nous adoptons la définition, plus large, proposée par Edminster en 1978 qu'un modèle est une représentation simplifiée d'un système. Il s'agit ici d'utiliser des concepts, des méthodes, des théories mathématiques qui vont permettre de décrire, comprendre et prévoir l'évolution de phénomènes dans un domaine donné.

L'élaboration d'un modèle se fait selon les étapes (Figure II.1) décrites ci-dessous.

Point 1 : Quelle est l'utilité du modèle ? Qui s'en servira ?
Point 2 : Conceptualisation du fonctionnement du système considéré
Point 3 : Traduction du modèle conceptuel en un modèle de simulation
Point 4 : Evaluation du modèle d'évaluation
Point 5 : Analyse d'incertitude et analyse de sensibilité
Point 6 : Identification des paramètres à ré-estimer (si nécessaire)
Point 7 : Utilisation du modèle
Point 8 : Diffusion et accompagnement à l'utilisation du modèle

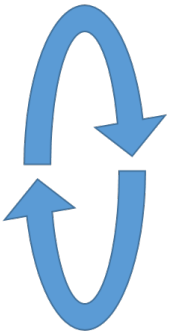


Figure II.1 – Étapes de construction d'un modèle

1. Quelle est l'utilité du modèle ? Qui s'en servira ?

Dans le domaine agricole, certains modèles sont des outils destinés à la recherche (Boote *et al.*, 1996 ; Cox, 1996). D'autres sont développés pour être utilisés par des conseillers agricoles, voire directement par des agriculteurs.

En agronomie, les modèles peuvent être utilisés *ex ante* pour aider à la conception de systèmes de culture ; *ex post* pour l'évaluation de performances de ces systèmes ; ou encore avec un horizon de simulation court pour des décisions tactiques (*e.g.* décision de traitement contre un bioagresseur donné). Les modèles de culture simulent l'impact des pratiques culturales sur les performances de systèmes composés du climat, du sol et de la culture (Loyce et Wéry, 2006). Le couplage de ce type de modèle à des outils permettant de prendre en compte la décision des utilisateurs de modèles a conduit à la mise en place d'OADs (Gaunt *et al.*, 1997 ; Aubry *et al.*, 1998 ; Girard et Hubert, 1999 ; Bergez *et al.*, 2001 ; Jones *et al.*, 2003). La mise en place d'OADs permet l'utilisation de modèles par un public plus large que les modélisateurs, en particulier pour les conseillers agricoles, voire les agriculteurs. L'utilisation d'OADs par des non-modélisateurs s'est heurtée à différents obstacles (Lynch *et al.*, 2000). En particulier, les sorties d'un modèle peuvent être en décalage avec l'expertise des utilisateurs ou bien les variables d'entrée du modèle peuvent être difficiles à renseigner. Il est donc particulièrement important d'identifier le futur utilisateur avant toute construction du modèle afin de mieux cerner les besoins et les contraintes. Le futur utilisateur doit être inclus dans le processus de construction de l'OAD.

2. Conceptualisation du fonctionnement du système considéré

Une fois que les attentes de l'utilisateur ont été identifiées, elles doivent contribuer à la formalisation du système étudié. Cette étape de conceptualisation consiste à formaliser une représentation du système à piloter. Cette représentation dépend de la vision que le modélisateur a du système, et de l'échelle temporelle et spatiale dans laquelle il se place (Voinov et Bousquet, 2010). Cette conceptualisation dépend étroitement du modèle mental de l'utilisateur. Ce modèle mental est construit sur l'expérience, les théories et les hypothèses que l'utilisateur pose (Johnson-Laird, 1983 ; Heemskerk *et al.*, 2003).

3. Traduction du modèle conceptuel en un modèle de simulation

Il existe de nombreuses manières de formaliser des modèles. Nous proposons les deux catégories suivantes :

— les modèles statiques (qualitatifs ou quantitatifs) qui n'intègrent pas explicitement la notion de temps. Ils décrivent l'état d'un système en un temps donné, et s'intéressent aux structures

de ce système (Fannader et Leroux, 1999).

— les modèles dynamiques qui intègrent explicitement la notion de temps. Ils sont basés sur la compréhension du comportement des éléments du système et s'intéressent à l'évolution des variables choisies. Nous identifions deux types de modèle dynamique : les modèles statistiques, qui sont construits à partir d'observations et d'hypothèses mathématiques, et les modèles mécanistes, qui sont basés sur des représentations de processus formalisés sous forme mathématique. Cette catégorisation est toutefois quelque peu théorique car les paramètres de formalismes mécanistes sont souvent ajustés à l'aide de méthodes statistiques.

Les deux types de modèles (statiques et dynamiques) ont été abordés au cours de cette thèse. Les modèles statiques qui ont été construits sont qualitatifs.

Les modèles sont composés de variables et de paramètres. Les variables sont des ensembles de valeurs que l'on observe et prédit. Les variables peuvent représenter des mesures de phénomène (par exemple, une mesure de rendement). Nous définissons les paramètres comme des réels constants dont la valeur doit être estimée. Les valeurs des paramètres peuvent être estimées grâce à la littérature et/ou par expertise.

4. Évaluation du modèle d'évaluation

Après le développement d'un modèle, celui-ci doit être évalué. L'évaluation consiste à comparer des données observées à des données simulées par le modèle. Un jeu de données *ad hoc* et des critères statistiques sont donc nécessaires. Si de nombreux outils statistiques sont disponibles pour évaluer un modèle quantitatif, les outils disponibles pour les modèles qualitatifs sont beaucoup moins fréquemment utilisés, notamment pour ce qui concerne les variables ordinales.

L'obtention de résultats insatisfaisants lors d'une évaluation peut indiquer que le modèle n'est pas adapté aux besoins de l'utilisateur.

5. Analyse d'incertitude et analyse de sensibilité

L'analyse d'incertitude donne des informations sur l'incertitude associée aux prédictions d'un modèle. Outre les incertitudes liées aux connaissances, les modèles ont trois autres sources d'incertitudes. Les variables d'entrée sont la première source, lorsque celles-ci sont générées par un instrument de mesure (les stations dans le cas de variables météorologiques par exemple) ou des enquêtes de pratiques culturales non homogènes (lorsqu'elles proviennent de différents enquêteurs). Ces erreurs sont difficilement rectifiables. La deuxième source est le modèle mathématique choisi. Cette source peut être limitée en faisant un choix de modèle adapté au préalable. La troisième source est la valeur des paramètres choisis. Elle se base sur l'expertise et les mesures identifiées

dans la littérature scientifique et technique du domaine concerné. Les valeurs adoptées pour les différents paramètres sont donc biaisées par l'avis de l'expert, de la manière dont l'enquête a été réalisée. C'est sur cette dernière source (les paramètres) que le modèle peut être modifié. Cette étape est l'analyse de sensibilité. Il s'agit d'une étape très importante car un même modèle avec des valeurs de paramètres différentes peut donner des résultats très différents. L'erreur peut être faite de remettre en question le formalisme mathématique d'un modèle cohérent vis à vis d'un système à cause de valeurs de paramètres incorrectes. La valeur de certains paramètres doit être modifiée pour que le modèle soit adapté au système étudié. Il s'agit d'une étape de calibration pendant laquelle les paramètres ne sont plus des constantes mais des variables dont il faut déterminer la valeur pour adapter le modèle au système étudié. Une analyse de sensibilité est alors habituellement réalisée en amont, notamment lorsque le nombre de paramètres est grand (comme pour le modèle conçu par Del Grosso *et al.* en 2011 avec plus de 1000 paramètres par exemple). Le principe est de faire varier la valeur de paramètres sur un intervalle de valeurs et d'étudier la réponse de la variable de sortie du modèle (Saltelli *et al.*, 2000). Il existe différentes méthodes permettant de réaliser une analyse de sensibilité (Wallach *et al.*, 2006 ; Saltelli *et al.*, 2008) qui ne feront pas l'objet de notre étude. En effet, cette thèse se concentre sur des modèles qualitatifs ou des modèles quantitatifs simples (faible nombre de paramètres et de variables d'entrée) pour lesquels une analyse de sensibilité n'est pas indispensable.

La méthode proposée a pour objectif d'identifier la combinaison de paramètres à ajuster permettant de maximiser la qualité de prédiction du modèle considéré. Pour chacune des combinaisons possibles, l'algorithme va estimer un critère statistique caractérisant la qualité de prédiction. Pour ce faire, une technique de validation croisée « Leave-One-Out » (Moore, 2001) a été mise en place. Afin d'être rigoureux et d'éviter toute corrélation spatio-temporelle potentielle (Wallach *et al.*, 2001), lorsqu'une situation est utilisée pour tester la qualité de prédiction du modèle, les données associées au même site ou à la même année sont exclues du jeu d'apprentissage de l'algorithme. Une boucle parcourt l'ensemble des sites-années de manière à utiliser toute l'information disponible. Ainsi les apprentissages et les tests du modèle sont réalisés sur l'ensemble du jeu de données tout en conservant l'indépendance entre données utilisées pour construire le modèle et données utilisées pour le tester (Figure II.2). Il s'agit de l'évaluation de la qualité de prédiction

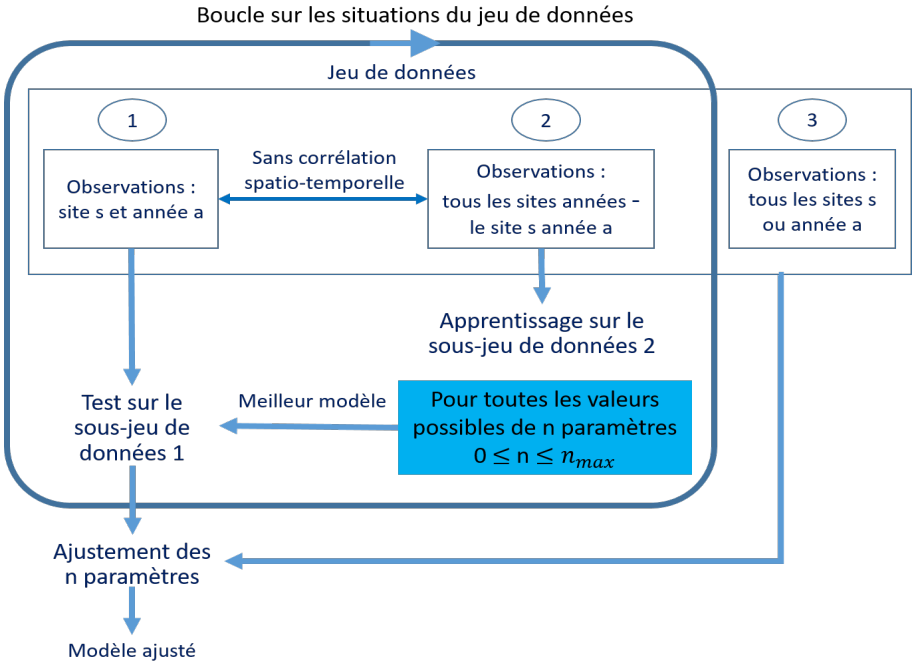


Figure II.2 – Principe de la validation croisée sans corrélation spatio-temporelle et de l'ajustement final. n_{max} est le nombre de paramètres du modèle et n est chaque entier entre 0 et n_{max}

Toutes les combinaisons de paramètre sont testées, soit

$$\sum_{k=1}^{k=n} \binom{n}{p} = 2^n \tag{1.1}$$

avec p un nombre de paramètres parmi n, le nombre de paramètres du modèle.

Chacune des 2^n combinaisons des n paramètres du modèle d'intérêt sont testées en validation croisée en utilisant le critère d'évaluation adapté à la nature du modèle. Les résultats obtenus pour chaque combinaison de paramètres sont synthétisés à l'échelle du jeu de données complet.

La combinaison de paramètres retenue est celle qui minimise les erreurs du modèle.

6. Identification des paramètres à ré-estimer (si nécessaire)

Cette étape vise à modifier les paramètres identifiés au cours de l'étape précédente en utilisant le jeu de donnée complet. Les valeurs des paramètres identifiées sont modifiées pour obtenir la meilleure qualité d'ajustement possible du modèle Figure II.2.

7. Utilisation du modèle

Dans le domaine agricole, les modèles de culture conçus ont généralement pour but de simuler la croissance d'une culture et le rendement associé en conditions de stress abiotiques, sous l'effet de pratiques agricoles. Les modèles focalisés sur les bioagresseurs ont pour vocation de décrire les

dynamiques épidémiques *sensu lato*, les dégâts associés, voire les dommages induits. Certains de ces modèles sont destinés aux **B**ulletin de **S**anté du **V**égétal (BSV). Ces modèles sont des OAD. Ils décrivent un système à l'instant présent et/ou à venir, qui évalue un risque et qui repose sur des règles de décision pour aider l'utilisateur à prendre des mesures adaptées par le biais de seuils ou d'algorithmes de prise de décision (Poisson, 2015). Les OAD sont de type tactiques ou stratégiques (Dubrulle *et al.*, 2014). Les OAD tactiques s'intéressent aux questions posées sur le court terme, pour atteindre un objectif de rendement par exemple. Les OAD stratégiques permettent de répondre à des questions posées le plus souvent sur le long terme pour évaluer la conduite d'un système sur un certain nombre de critères.

Les OAD peuvent donc être utilisés de deux manières différentes, de manière *ex ante* pour prévoir les effets de pratiques agricoles dans une situation de production donnée afin d'aider à la conception de systèmes de culture agroécologiques (Anonyme, 2014), et de manière *ex-post* pour réaliser des diagnostics agronomiques.

Un « bon » OAD est caractérisé par les sept critères suivants : il est utile, robuste, il nécessite des données d'entrée et renvoie des sorties faciles à utiliser, il est efficace, générique, interactif et enfin transparent (Cordier, communication personnelle).

8. Diffusion et accompagnement à l'utilisation du modèle

Une fois le modèle calibré et évalué, il doit être rendu accessible aux utilisateurs (notamment ceux qui ont contribué à son développement). Le modèle doit être diffusé, présenté et expliqué. Des formations peuvent être mises en place pour le diffuser et apprendre à l'utiliser. Des outils de diffusion comme une plateforme internet peuvent être mis en place pour promouvoir son utilisation.

1.2 Les étapes de construction abordées au cours de la thèse

Nous avons travaillé sur les étapes 1, 2, 3, 4, 6, et 7 de la procédure de construction d'un modèle (Figure II.1).

En particulier, cette thèse a contribué à la mise en place de deux méthodes permettant l'identification des paramètres à modifier pour des modèles quantitatifs et qualitatifs. Ces méthodes résultent d'une adaptation originale de la première méthode d'estimation de paramètres pour les modèles quantitatifs (Wallach *et al.*, 2006) aux modèles qualitatifs IPSIM. Ces méthodes sont décrites chapitre 4 et 5 de cette partie.

Les outils suivants ont été utilisés au cours de cette thèse pour l'évaluation de modèles qualitatifs et quantitatifs.

2. Présentation de l'approche de modélisation IPSIM

2.1 Le modèle générique IPSIM

Nous allons ici présenter la méthode de modélisation IPSIM. Les modèles IPSIM (Injury Profile SIMulator) sont construits à l'aide de la méthode DEX (Decision EXpert), implémentée dans le logiciel DEXi (disponible via le lien : `http://kt.ijs.si/MarkoBohanec/DEXi.html`, Bohanec *et al.*, 2013). Les modèles développés avec la plateforme IPSIM (Aubertot et Robin, 2013) ont pour but de prédire les dégâts causés par un cortège de bioagresseurs sur une culture en fonction de la situation de production et des pratiques agricoles. Les modèles conçus avec IPSIM sont des réseaux hiérarchiques déterministes qualitatifs (Aubertot et Robin, 2013 ; Robin *et al.*, 2013). Il s'agit de structures hiérarchiques (Figure II.3), développées grâce à des dires d'experts et d'une analyse de la littérature scientifique et technique (guides de culture, points techniques et enquêtes de Terres-Inovia par exemple) sur les effets du système de culture, du pédoclimat et de l'environnement sur les trois bioagresseurs d'intérêt.

Une structure générique a été proposée par Aubertot et Robin en 2013 (Figure II.3). Toutes les variables de ce type de modèles sont ordinales ou nominales. Chaque attribut dispose d'un nom, d'une description et d'une échelle pouvant prendre plusieurs niveaux. Les attributs de base représentent les variables d'entrée du modèle à renseigner par l'utilisateur. Dans l'arbre générique (Figure II.3), les attributs de base ne sont pas affichés en gras. Les attributs agrégés apparaissent en gras. La valeur de chaque attribut agrégé (en gras dans la Figure II.3) est définie à l'aide d'une table d'agrégation. Un de ces attributs est particulier. C'est celui qui se trouve à la racine de l'arbre puisqu'il représente la principale variable de sortie du modèle, et que sa valeur dépend des valeurs des attributs d'entrée, de la structure du modèle et de tables d'agrégation des variables « Cropping

Practices », « Sol and climate » et « Interactions at the territory level » (les interactions à l'échelle du paysage). Les différents attributs agrégés sont caractérisés par des tables d'agrégation. Les tables sont composées de lignes représentant toutes les combinaisons possibles des différentes valeurs des attributs agrégés.

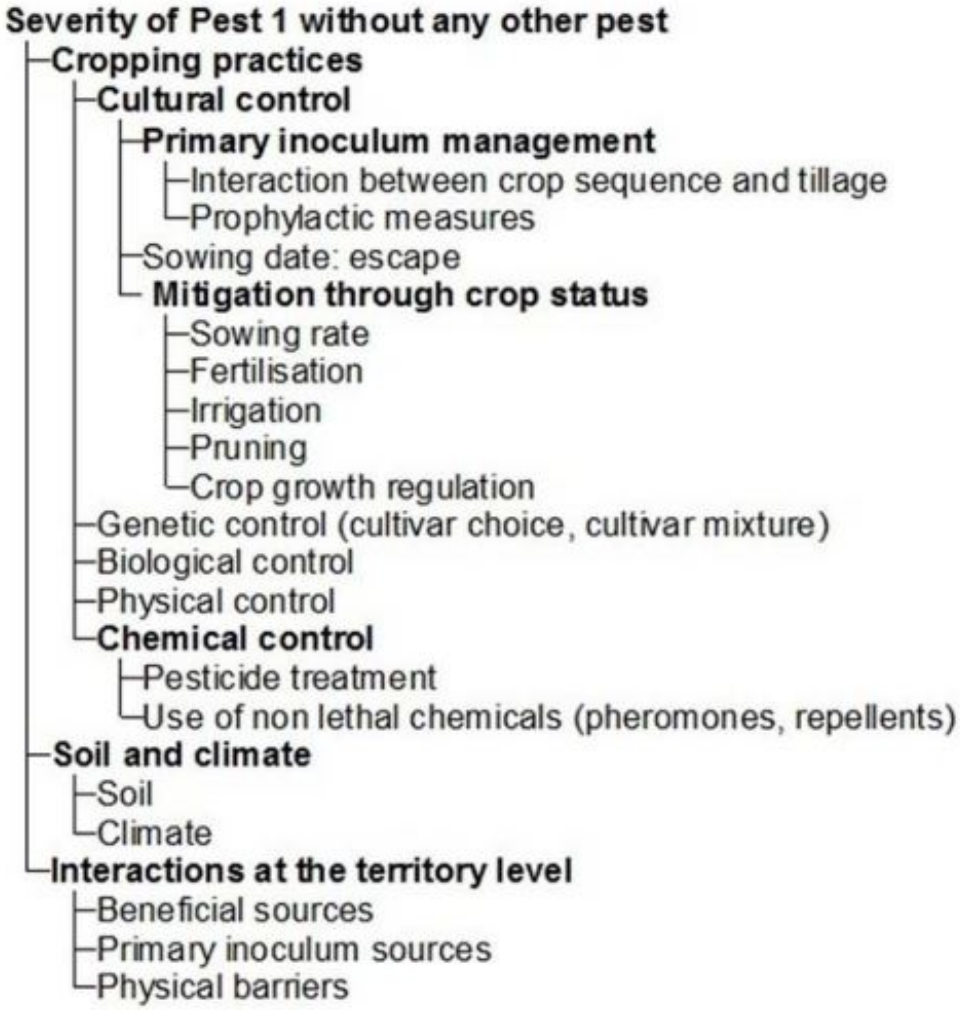


Figure II.3 – Structure générique d'IPSIM telle que proposée par Aubertot et Robin (2013)

Habituellement, des poids sont utilisés dans les processus d'analyse décisionnelle pour accorder une importance à une variable. Ils permettent de définir une fonction d'utilité à maximiser. Cette dernière s'écrit de la manière suivante (Bohanec, 2015) :

$$f(X_1, X_i, ..., X_n) = w_1.X_1 + w_i.X_i + ... + w_n.X_n \tag{2.1}$$

Où X_i représente des variables et w_i les poids associés à ces variables. Ces poids rendent compte de l'influence de chaque variable d'entrée X sur la variable de sortie $f(X_1, X_i, ..., X_n)$.

Dans le modèle DEXi, la fonction d'utilité n'est pas définie de cette manière mais par des règles de décision. Le principe qui régit le calcul des poids est illustré par l'exemple suivant :

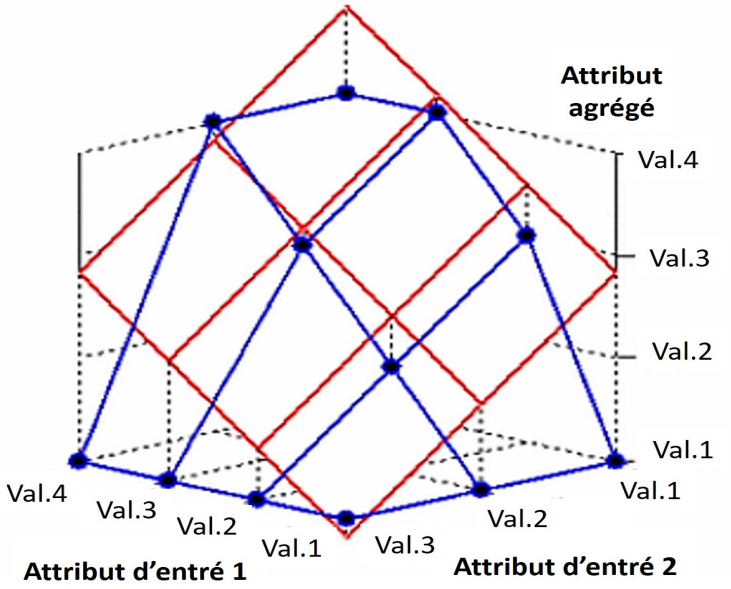


Figure II.4 – Illustration du calcul des poids dans un modèle DEXi composé de deux attributs agrégés composés respectivement de 3 et 4 valeurs. Les points bleus reliés représentent la fonction d'utilité définis par les règles de décision. En rouge, l'hyperplan créée par DEXi pour calculer les poids (d'après Bohanec, 2015). Val. = Valeur.

L'attribut agrégé dépend des valeurs de n attributs agrégés (dans l'exemple, Figure II.4, ce sont 2 attributs agrégés composés respectivement de 3 et 4 valeurs). Les règles de décision établies définissent un hyperespace à n + 1 dimension(s) (ici n + 1 = 3). Un hyperplan est défini par DEXi dans cet hyperespace de sorte qu'il soit le plus proche possible de l'hyperespace défini par les règles de décision (utilisation du critère des moindres carrés). Les poids sont approximés à partir des pentes de l'hyperplan. Plus une pente est importante pour un attribut, plus le poids relatif de cet attribut est élevé.

Chaque attribut est ainsi caractérisé par quatre poids : local, global, local normalisé et global normalisé. Le poids local caractérise l'influence d'un attribut X_i sur l'attribut agrégé ascendant. Le poids local normalisé caractérise l'influence d'un attribut X_i sur l'attribut agrégé ascendant en pondérant en fonction du nombre de valeurs adopté dans l'échelle de l'attribut X_i. Le poids global caractérise l'influence d'un attribut sur la variable de sortie du modèle. Le poids global normalisé caractérise l'influence d'un attribut X_i sur la variable de sortie du modèle en pondérant en fonction du nombre de valeurs de l'échelle de l'attribut X_i. On ne considérera donc que les poids normalisés au cours de ce travail.

Les modèles IPSIM ont pour vocation de servir d'OAD. Cette méthode prévoit la mise en place

de convertisseurs qui permettent la traduction des données de terrain (qui peuvent être de nature qualitative ou quantitative) en données qualitatives (principalement ordinales) pour le modèle. Les convertisseurs sont utilisés sous Excel et codés en **V**isual **B**asic **A**pplication (VBA). L'utilisateur y renseigne ses données d'entrée (comme par exemples le nom de la variété de tournesol, la date de semis, la quantité de fertilisation apportée). Le convertisseur renvoie les valeurs correspondantes aux variables d'entrée du modèle IPSIM.

2.2 Évaluation des modèles IPSIM dans la thèse

Elle repose sur l'utilisation de matrices de confusion et de deux critères statistiques. Les matrices de confusion sont des tables de contingence représentant la distribution des valeurs observées et simulées de deux variables ordinales ou nominales (Witten *et al.*, 2011). Un exemple avec des variables à trois niveaux est présenté Tableau II.1.

		Variable observée		
		Valeur 1	Valeur x	Valeur N
Variable simulée	Valeur 1	X	X	X
	Valeur x	X	X	X
	Valeur N	X	X	X

Tableau II.1 – Matrice de confusion entre une variable catégorielle observée et une variable catégorielle simulée. N : niveau maximal de la variable. x : valeur intermédiaire de la variable. X : un nombre de situation quelconque. En rouge, les situations sont les plus mal prédites ; en orange, les erreurs sont intermédiaires ; en vert, les situations sont bien prédites (cas d'une variable ordinale).

Ces matrices permettent de visualiser les erreurs du modèle, notamment les tendances du modèle à sur ou sous-estimer les observations.

Les matrices de confusion peuvent être résumées par de nombreux critères statistiques. Nous avons choisi les deux critères statistiques suivants :

Definition 2.1 — La précision. Il s'agit de la proportion de situations correctement prédites par le modèle. Elle renseigne sur la qualité générale du modèle. La précision varie de 0 (aucune prédiction n'est correcte) à 1 (toutes les prédictions sont correctes).

Definition 2.2 — kappa de Cohen pondéré quadratiquement. Ce critère (Agresti et Kateri, 2010) pénalise chaque erreur avec un poids égal à la différence du nombre de classes entre valeurs observées et prédites, élevé au carré. Ce critère a été choisi car il peut être interprété comme le pourcentage de variabilité des données expliquées par le modèle (Fleiss et Cohen, 1973) et parce qu'il est adapté aux variables ordinales. Le κ varie de -1 (désaccord total) à 1 (accord parfait). Un κ pondéré nul correspond à une indépendance entre les deux variables. L'interprétation des valeurs de κ a été réalisée à l'aide d'une grille proposée par Landis et Koch en 1977.

$$\kappa = \frac{P_o(w) - P_e(w)}{1 - P_e(w)} \tag{2.2}$$

Avec

$$P_o(w) = \sum_{i=1}^{i=r} \sum_{j=1}^{j=r} w_{ij} p_{ij} \tag{2.3}$$

qui est l'accord observé du modèle considéré,

et

$$P_e(w) = \sum_{i=1}^{i=r} \sum_{j=1}^{j=r} w_{ij} p_{i.} p_{.j} \tag{2.4}$$

qui est l'accord d'un modèle aléatoire

$$w_{ij} = \frac{1 - ((i-j)^2}{(r-1)^2)} \; ; \; p_{ij} = \frac{n_{ij}}{n} \; ; \; p_{i.} = \frac{n_{i.}}{n} \; ; \; p_{.j} = \frac{n_{.j}}{n} \tag{2.5}$$

et i est la i^{ieme} colonne de la matrice de confusion, j est la j^{ieme} ligne de la matrice de confusion. r est le nombre de classes, n est le nombre total d'observations.

3. Comparaison des performances de l'approche IPSIM et d'une méthode statistique

3.1 Objectifs de l'étude

Nous allons ici comparer la qualité de prédiction de modèles IPSIM à celle de modèles purement statistiques (construit uniquement à l'aide d'un jeu de données). Nous avons choisi la méthode CART, compte tenu de ses performances et de sa facilité de mise en œuvre. Les qualités de prédiction respectives de l'approche CART et de l'approche IPSIM ont été évaluées en fonction de la taille du jeu de données.

3.2 Matériel et méthodes

3.2.1 Description du principe de *Classification And Regression Tree*

Les modèles CART sont des arbres hiérarchiques (Breiman *et al.*, 1984 ; Morgan, 2014). Il s'agit de techniques d'apprentissages supervisées, car l'apprentissage permettant la construction du modèle se fait sur un jeu de données comportant les valeurs de la variable à prédire. Nous considérons ici des arbres de classification dont la variable de sortie est catégorielle. L'arbre sert donc à prédire la classe de la variable de sortie en fonction des variables explicatives. L'arbre consiste en une succession de tests numériques binaires. Cette arborescence est un partitionnement de l'espace p-dimensionnel dans lequel se trouve la variable de sortie. Chaque variable explicative peut être considérée comme une dimension dans cet espace p-dimensionnel. L'arbre de décision le divise en rectangles (lorsque p = 2 variables d'entrée), en cubes (lorsque p = 3 variables d'entrée) ou en hyper-rectangles (lorsque p > 3 variables d'entrée). Chaque partitionnement est testé en utilisant une fonction de coût. Chaque partition sélectionnée dans le modèle est celle qui a le coût le plus bas. Ce coût rend compte de la pureté du partitionnement, c'est-à-dire de la proportion des différentes

classes de sortie à chaque nœud. Ce coût repose sur le critère de Gini (Équation 3.1).

Definition 3.1 — Indice de Gini. Cet indice rend compte de la répartition d'une variable au sein d'une population (Gini, 1921). Sa gamme de valeur est [0 ; 1[. 0 correspond à un partitionnement est parfait et 1 correspond à une inégalité parfaite. Une valeur de 0,5 signifie que toutes les classes sont présentes en même proportion.

$$I_G(f) = 1 - \sum_{i=1}^{i=m} f_i^2 \tag{3.1}$$

Avec i l'indice des classes, m le nombre de classes et f_i, la proportion d'éléments de l'ensemble avec la classe i.

La construction de ce modèle se fait selon le principe de parcimonie. Le modèle doit être le plus simple possible tout en restant performant. Il y a deux manières d'arrêter le partitionnement (Salzberg, 1994) : soit pendant la construction de l'arbre, lorsque les feuilles ont atteint un effectif ou une homogénéité jugée suffisante, soit après construction de l'arbre dont les feuilles sont les plus homogènes possibles. Le jeu d'apprentissage est divisé en deux échantillons : l'échantillon d'expansion sur lequel est construit l'arbre et l'échantillon d'élagage sur lequel l'arbre est limité. L'arbre retenu est le plus petit arbre dont l'erreur n'excède pas $1 - \sigma$ avec σ, l'écart-type de l'erreur optimale obtenue sur l'échantillon d'élagage. σ est estimé de la manière suivante :

$$\sigma = \sqrt{\frac{\varepsilon * (1 - \varepsilon)}{n}} \tag{3.2}$$

Avec ε, l'erreur optimale et n la taille de l'échantillon

L'erreur seuil qui ne doit pas être dépassée est :

$$\varepsilon_{seuil} = \varepsilon + \theta * \sigma \tag{3.3}$$

Avec ε_{seuil} l'erreur optimale, θ = 1 et σ, l'écart-type de l'erreur optimale

3.2.2 Méthode de comparaison

Pour chaque approche, la relation entre le kappa de Cohen pondéré quadratiquement (κ) et la taille globale du jeu de données a été établie. Pour le modèle IPSIM, l'estimation du kappa de

Cohen pondéré quadratiquement a été effectuée en réalisant une série de n tirages aléatoires parmi les N situations disponibles (n variant de 1 à N). Pour chaque valeur de n, x tirages avec remise ont été réalisés. Pour le modèle statistique, l'estimation du Kappa de Cohen pondéré quadratiquement a été effectuée en considérant des tailles globales du jeu de données m = N_{min}, jusqu'à m = N avec N_{min}, le nombre de niveaux de la variable de sortie multiplié par 3. Nous multiplions par 3 pour pouvoir obtenir chaque niveau d'incidence dans les échantillons d'expansion, d'élagage et dans le jeu de test sélectionnés pendant la validation croisée (de type « Leave One Out » (Moore, 2001) éliminant les éventuelles corrélations spatiales ou temporelles entre les données utilisées) pour construire puis évaluer le modèle. Les modèles CART ont été construits en utilisant la fonction C5.0 de la librairie C50 sous R. Cette fonction permet l'utilisation de variables explicatives qualitatives.

3.2.3 Modèles existants utilisés

Deux modèles IPSIM ont été considérés : IPSIM - Chayotte - Mouche (Deguine *et al.*, en prep) et IPSIM - Blé - Rouille brune (Robin *et al.*, 2013) (Figures de l'Annexe 1 et 2). Ces modèles sont basés sur le même principe de prédiction de la sévérité des dégâts causés par un bioagresseur (un insecte pour IPSIM - chayotte - Mouche, et un champignon pour IPSIM - Blé - Rouille brune) sur une culture (la chayotte et le blé). La qualité de prédiction initiale du modèle d'IPSIM - Chayotte - Mouche est caractérisée par un kappa de Cohen pondéré quadratiquement de κ = 0,432. Le jeu de données est décrit dans la Tableau de l'Annexe 1. Il est composé de 47 observations et dispose de 15 variables explicatives. En raison du temps de calcul conséquent, de la taille des jeux de données et du nombre de sites-années de chaque jeu, les tailles de jeu de données suivantes ont été considérées pour le système chayotte-mouche : 20, 25, 30, 35, 40, 45 et 47. Pour chaque taille de jeu de données, x = 500 tirages ont été réalisés.

La qualité de prédiction initiale d'IPSIM-blé-rouille brune est caractérisée par un kappa de Cohen pondéré quadratiquement de κ = 0,68. Le jeu de données est décrit dans le Tableau de l'Annexe 2. Le jeu de données est composé de 1788 observations et dispose de 8 variables explicatives. Les tailles de jeu de données suivantes ont été considérées pour le système blé : 100, 300, 500, 700, 900, 1100, 1300, 1500, 1700 et 1788. Pour chaque taille de jeu de données, x = 1000 tirages ont été réalisés. En raison de leurs tailles respectives, le jeu de données lié à la chayotte est décrit par parcelle tandis que celui lié au blé est décrit par département.

3.3 Résultats

Les résultats obtenus pour la comparaison des approches CART et IPSIM sur les systèmes de la chayotte et du blé sont présentés Figures II.5 et II.6. Les Figures II.5A et II.6A présentent les résultats pour les modèles IPSIM et les Figures II.5B et II.6B présentent les résultats pour les modèles CART.

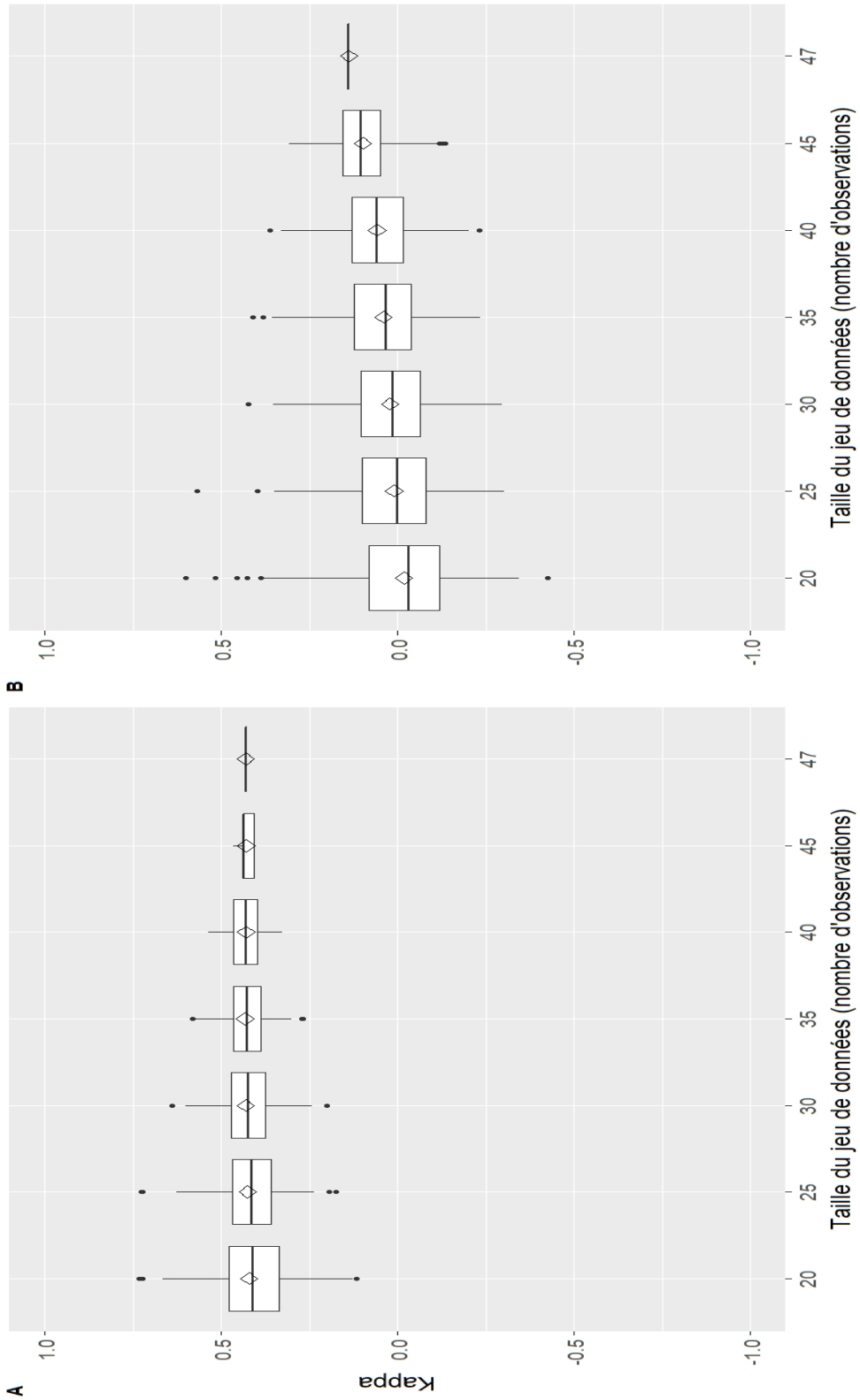


Figure II.5 – Boites à moustaches composées des 500 κ pondérés quadratiquement obtenues pour différentes tailles du jeu de données du système chayotte avec A : la méthode IPSIM. B : la méthode CART

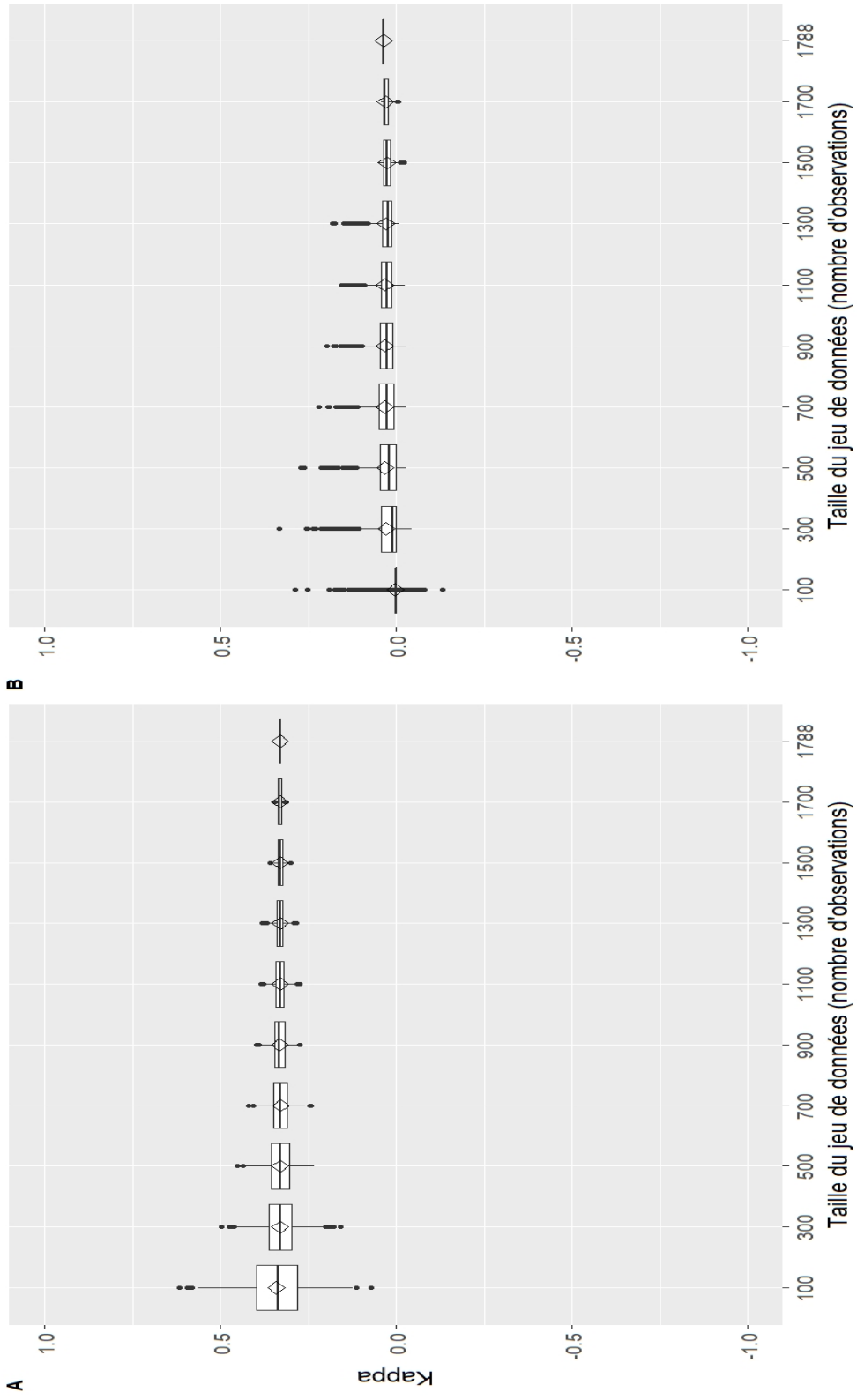


Figure II.6 – Boites à moustaches composées des 1000 κ pondérés quadratiquement obtenues pour différentes tailles du jeu de données pour le système blé avec A : la méthode IPSIM. B : la méthode CART

Toutes les boîtes à moustaches (hormis la dernière pour la méthode CART pour laquelle l'ensemble des données disponibles a été utilisé) résultent d'évaluations de qualités prédictives (*i.e.* de simulations obtenues par des modèles dont les paramètres ont été estimés sur un jeu de données indépendant de celui utilisé pour leur évaluation). Plus le jeu de données est grand, plus les boîtes à moustaches sont étroites. Ceci est normal puisque la grande taille des jeux de données augmente leur représentativité de l'ensemble du jeu et entraîne une moindre variabilité des κ. La dernière boîte à moustaches correspond, pour CART, à la qualité d'ajustement (apprentissage sur le jeu de données total). Pour IPSIM, qui ne nécessite pas d'apprentissage, la dernière boîte à moustaches correspond à la qualité de prédiction sur l'ensemble du jeu de données, ce qui explique l'absence de variabilité pour ces boîtes à moustaches Figures II.5 et II.6.

Les boîtes à moustaches obtenues avec la méthode IPSIM sont centrées sur la qualité de prédiction obtenue en utilisant tout le jeu de données ($\kappa = 0{,}432$ pour la chayotte ; et $\kappa = 0{,}68$ pour le blé). Les résultats obtenus avec la méthode CART sont moins bons : pour le système chayotte, ils sont centrés sur 0,0 et croissent lorsque la taille du jeu de données augmente, pour atteindre $\kappa = 0{,}14$. Pour le blé pour lequel nous disposions d'un plus grand jeu de données, les boîtes à moustaches sont légèrement positives et se stabilisent à $\kappa = 0{,}04$. Nous pouvons expliquer ces différences de performances entre IPSIM et CART notamment par le fait que, pour un même jeu de données le modèle IPSIM est évalué sur tout le jeu de données disponible, la méthode CART nécessite de diviser ce même jeu de données en plusieurs parties pour d'une part, l'apprentissage, et d'autre part son évaluation. Néanmoins, même lorsque l'on considère la qualité d'ajustement de la méthode CART, le κ demeure inférieur à celui de la méthode IPSIM.

3.4 Discussion

La méthode IPSIM obtient, pour toutes les tailles du jeu de données utilisé, de meilleures qualités de prédiction que la méthode CART. Ceci peut s'expliquer par le fait que l'expertise introduite par la méthode IPSIM permet d'intégrer une information bien plus large que pour la méthode CART qui ne dispose que du seul jeu de données pour élaborer tout l'arbre. Il faut néanmoins souligner que ces résultats ne concernent que les deux modèles étudiés et qu'il ne s'agit nullement d'une conclusion « universelle ». Les performances relatives des deux méthodes dépendront en effet de l'importance des données disponibles, et de la quantité et la qualité des informations disponibles dans la littérature scientifique et technique et chez les experts.

Néanmoins, ces résultats mettent en évidence l'importance des approches de modélisation reposant sur les dires d'experts pour les problèmes complexes ou les problèmes pour lesquels

les données sont rares. De plus, les modèles CART sont construits sur des jeux de données et ne peuvent être améliorés par la suite. Ce n'est pas le cas pour les modèles IPSIM. Ces derniers ne reposant que sur des dires experts et la littérature scientifique et technique pour leur construction, leurs qualités de prédiction peuvent encore être améliorées en utilisant un jeu de données. Ceci fait l'objet de la partie 4.

La méthode IPSIM a d'autres avantages. Les modèles IPSIM, malgré leur apparente simplicité liée à leur nature qualitative, permettent de prendre en compte une grande quantité de connaissances disponibles dans la littérature (Robin *et al.*, 2013). Cette méthode prévoie l'intégration de la dimension verticale de la protection des cultures (ensemble des méthodes de contrôle). Elle peut s'appliquer à un seul bioagresseur ou bien intégrer la dimension horizontale de la protection des cultures (gestion d'un cortège de bioagresseurs). Elle permet également de prendre en compte plusieurs échelles de temps, puisqu'on peut s'intéresser à l'échelle journalière (pour le climat), au cycle de vie du bioagresseur, au cycle cultural, à l'itinéraire technique ou plus largement à l'échelle de la succession culturale et on peut considérer toutes ces échelles au sein du modèle. Enfin, ces modèles peuvent considérer plusieurs échelles d'espace, depuis la parcelle jusqu'au paysage (le département par exemple).

4. Méthode innovante améliorant la qualité prédictive des modèles IPSIM

4.1 Principe de la méthode

La méthode de développement des modèles IPSIM est la même pour tous les modèles. On formule l'hypothèse selon laquelle la structure des arbres est équivalente à des équations pour les modèles quantitatifs, et que chaque ligne de chaque table d'agrégation est équivalente à un paramètre numérique dans ces modèles quantitatifs. Sans modifier la structure des modèles IPSIM, il est alors possible d'améliorer leurs qualités de prédiction en optimisant les règles d'agrégation définies initialement par expertise. On a proposé de mettre en place une méthode algorithmique originale en adaptant une méthode d'estimation des paramètres pour modèles quantitatifs (Wallach *et al.*, 2001).

4.1.1 Démarche algorithmique

Le principe de la méthode consiste à identifier la ou les règle(s) d'agrégation(s) à modifier pour améliorer la qualité de prédiction globale du modèle. Pour ce faire, il faut disposer d'un modèle IPSIM et d'un jeu de données indépendant comportant plusieurs situations pour lesquelles sont renseignés le lieu, l'année, la valeur de chaque variable d'entrée du modèle IPSIM considéré, et de la variable de sortie.

L'algorithme va considérer l'ensemble des combinaisons de règles d'agrégation (RdA) pouvant être optimisées. Si un modèle comporte n RdA, il y a 2^n combinaisons possibles à considérer. Pour chacune de ces combinaisons, l'algorithme va estimer un critère statistique caractérisant la qualité de prédiction (le kappa de Cohen pondéré quadratiquement). Pour ce faire, une technique de validation croisée « Leave-One-Out » (Moore, 2001) a été mise en place. Afin d'être rigoureux

et pour éviter toute corrélation spatio-temporelle potentielle (Wallach *et al.*, 2001), lorsqu'une situation est utilisée pour tester la qualité de prédiction du modèle, les données associées au même site ou à la même année sont exclues du jeu d'apprentissage de l'algorithme. Une boucle parcourt l'ensemble des sites-années de manière à utiliser toute l'information disponible. En comparant les différentes valeurs de kappa obtenues, on peut identifier la combinaison de RdA dont l'ajustement permet d'obtenir la meilleure qualité de prédiction possible. Une fois cette combinaison de RdA identifiée, on l'ajuste sur l'ensemble des situations, de manière à obtenir le meilleur modèle possible. Le kappa associé à ce dernier ajustement correspond alors à une qualité d'ajustement et non à une qualité de prédiction.

Pour un modèle qualitatif, chaque combinaison de paramètres est optimisée sur un jeu d'apprentissage de sorte à avoir le kappa le plus proche de 1. Puis le jeu test sert à réaliser des prédictions avec les paramètres obtenus. Le kappa est mesuré sur les prédictions réalisées sur tous les jeux tests. La combinaison de paramètres retenue est celle qui maximise le kappa.

4.2 Validation informatique de l'algorithme améliorant la qualité prédictive des modèles IPSIM

L'algorithme d'amélioration de la qualité prédictive des modèles IPSIM a été appliqué au modèle préliminaire d'IPSIM - Tournesol - Phoma tige. L'objectif de ce travail est la validation réciproque de deux codes informatiques développés de manière indépendante par deux programmeurs dans deux langages différents. J'ai codé l'algorithme et le modèle IPSIM - Phoma sur tige sous R, et un code générique a été développé sous C++ par l'équipe de **M**athématiques et **I**nformatique **A**ppliquées de **T**oulouse (MIAT) (G. Quesnel) et appliqué à ce modèle.

L'optimisation du modèle préliminaire IPSIM - Phoma sur tige s'est réalisée en plusieurs étapes sous R :

1. Les agrégations du modèle ont été synthétisées sous la forme de tables d'agrégation.
2. Une fonction qui mobilise le modèle et les tables a été créée pour analyser une ou plusieurs situations pour lesquelles les valeurs des variables d'entrée sont renseignées. Cette fonction renvoie les résultats pour chaque nœud interne et l'incidence finale. Cette fonction a été utilisée pour l'optimisation du modèle.
3. Une fonction a été codée afin de séparer les situations utilisées pour l'évaluation des situations utilisées pour l'optimisation des règles d'agrégation (principe de la validation-croisée « Leave One Out » sans corrélation spatio-temporelle).
4. Une boucle a été mise en place pour permettre l'optimisation de n règles d'agrégations parmi

les nmax règles (52)

5. Les simulations de cette fonction ont été comparées, pour validation, pour n=0, n=1, et n=2.

Une matrice de confusion a été construite et la qualité de prédiction a été caractérisée avec le kappa de Cohen (κ) pondéré quadratiquement. Une fois validé, le modèle a subi un ajustement d'une ligne, puis de deux lignes.

4.2.1 Description d'IPSIM - Tournesol - Phoma sur tige et du jeu de données

Le modèle préliminaire IPSIM - Phoma sur tige est présenté Figure II.7. Il prédit l'incidence des taches sur tige en fonction des pratiques culturales et de l'environnement. Il dispose de 9 attributs d'entrée, de 5 attributs agrégés (dont l'attribut final) et de 52 règles d'agrégation (Tableau II.2 à II.6). Les valeurs adoptées par chaque attribut sont exposées Figure II.7. Le tableau de données utilisé Tableau II.7 pour l'optimisation est composé de 156 situations. Chaque situation est caractérisée par un département, une année, dispose d'une incidence observée et des 9 attributs d'entrée. Les situations y sont regroupées en fonction de leur appartenance à un département et à une année. Pour un département et à une année, il y a un nombre de parcelles et un nombre d'incidences et leurs niveaux associés.

Figure II.7 – Structure préliminaire de l'arbre IPSIM - Phoma sur tige représentée sous DEXi à gauche. Les attributs agrégés sont représentés en gras. Valeurs adoptées par chaque attribut à droite. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

	Pratiques culturales	Climat pour la réussite des infections	Interaction au niveau du territoire	**Incidence du Phoma sur tige**
1	Défavorable	Défavorable	Défavorable	**Très élevée**
2	Défavorable	Défavorable	Favorable	**Elevée**
3	Défavorable	Favorable	Défavorable	**Elevée**
4	Défavorable	Favorable	Favorable	**Nul à faible**
5	Moyennement favorable	Défavorable	Défavorable	**Très élevée**
6	Moyennement favorable	Défavorable	Favorable	**Elevée**
7	Moyennement favorable	Favorable	Défavorable	**Nul à faible**
8	Moyennement favorable	Favorable	Favorable	**Nul à faible**
9	Très favorable	Défavorable	Défavorable	**Elevée**
10	Très favorable	Défavorable	Favorable	**Elevée**
11	Très favorable	Favorable	Défavorable	**Nul à faible**
12	Très favorable	Favorable	Favorable	**Nul à faible**

Tableau II.2 – Table d'agrégation de l'incidence du phoma sur tige composée des 12 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

	Contôle cultural	Contrôle chimique	**Pratiques culturales**
1	Défavorable	Non traité	**Défavorable**
2	Défavorable	Traité	**Moyennement favorable**
3	Moyennement favorable	Non traité	**Moyennement favorable**
4	Moyennement favorable	Traité	**Très favorable**
5	Très favorable	Non traité	**Moyennement favorable**
6	Très favorable	Traité	**Très favorable**

Tableau II.3 – Table d'agrégation des pratiques culturales composée des 6 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

	L'atténuation par l'état du couvert	Gestion de l'inoculum primaire	Esquive (date de semis)	**Contôle cultural**
1	Défavorable	Défavorable	NonTardive	**Défavorable**
2	Défavorable	Défavorable	Tardive	**Défavorable**
3	Défavorable	Favorable	NonTardive	**Défavorable**
4	Défavorable	Favorable	Tardive	**Défavorable**
5	Défavorable	Très favorable	NonTardive	**Moyennement favorable**
6	Défavorable	Très favorable	Tardive	**Moyennement favorable**
7	Favorable	Défavorable	NonTardive	**Moyennement favorable**
8	Favorable	Défavorable	Tardive	**Moyennement favorable**
9	Favorable	Favorable	NonTardive	**Très favorable**
10	Favorable	Favorable	Tardive	**Très favorable**
11	Favorable	Très favorable	NonTardive	**Très favorable**
12	Favorable	Très favorable	Tardive	**Très favorable**

Tableau II.4 – Table d'agrégation du contrôle cultural composée des 12 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

	Etat du couvert au stade F1	Etat du couvert au stade E1	**L'atténuation par l'état du couvert**
1	Défavorable	Défavorable	**Défavorable**
2	Défavorable	Favorable	**Défavorable**
3	Favorable	Défavorable	**Défavorable**
4	Favorable	Favorable	**Favorable**

Tableau II.5 – Table d'agrégation de l'atténuation par l'état du couvert composée des 4 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

	Rotation des cultures	Travail du sol après l'antéprécédent	travail du sol après le précédent	Gestion de l'inoculum primaire
1	retour tout les 2 ans	Semis direct	Semis direct	**Défavorable**
2	retour tout les 2 ans	Semis direct	Travail du sol sans labour	**Défavorable**
3	retour tout les 2 ans	Semis direct	Labour	**Très favorable**
4	retour tout les 2 ans	Travail du sol sans labour	Semis direct	**Défavorable**
5	retour tout les 2 ans	Travail du sol sans labour	Travail du sol sans labour	**Défavorable**
6	retour tout les 2 ans	Travail du sol sans labour	Labour	**Très favorable**
7	retour tout les 2 ans	Labour	Semis direct	**Très favorable**
8	retour tout les 2 ans	Labour	Travail du sol sans labour	**Très favorable**
9	retour tout les 2 ans	Labour	Labour	**Favorable**
10	retour sur plus de 2 ans	Semis direct	Semis direct	**Favorable**
11	retour sur plus de 2 ans	Semis direct	Travail du sol sans labour	**Très favorable**
12	retour sur plus de 2 ans	Semis direct	Labour	**Très favorable**
13	retour sur plus de 2 ans	Travail du sol sans labour	Semis direct	**Très favorable**
14	retour sur plus de 2 ans	Travail du sol sans labour	Travail du sol sans labour	**Très favorable**
15	retour sur plus de 2 ans	Travail du sol sans labour	Labour	**Très favorable**
16	retour sur plus de 2 ans	Labour	Semis direct	**Très favorable**
17	retour sur plus de 2 ans	Labour	Travail du sol sans labour	**Très favorable**
18	retour sur plus de 2 ans	Labour	Labour	**Très favorable**

Tableau II.6 – Table d'agrégation de la gestion de l'inoculum primaire composée des 18 règles de décision du modèle préliminaire IPSIM - Phoma sur tige telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma sur tige ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma sur tige.

Département	Année	Nombre de parcelles observées	incidence observées nombre d'observations : niveau(x)
16	2002	2	« Élevée » : 4
17	2002	1	« Nulle à faible » : 5
17	2006	2	« Nulle à faible » : 2
17	2007	2	« Élevée » : 5 ; « Très élevée » : 1
18	2010	1	« Élevée » : 6
18	2011	1	« Élevée » : 1
31	1999	4	« Nulle à faible » : 3 ; « Élevée » : 19
31	2000	3	« Élevée » : 20
31	2001	3	« Nulle à faible » : 11 ; « Élevée » : 6
31	2006	1	« Très élevée » : 1
31	2008	1	« Élevée » : 5
31	2009	1	« Élevée » : 1
31	2010	8	« Élevée » : 13 ; « Très élevée » : 11
31	2011	3	« Très élevée » : 20
31	2012	1	« Très élevée » : 4
32	2011	1	« Élevée » : 3 ; « Très élevée » : 1
79	2002	1	« Nulle à faible » : 1 ; « Élevée » : 2
85	2002	1	« Nulle à faible » : 2 ; « Élevée » : 1
86	2002	5	« Nulle à faible » : 5 ; « Élevée » : 3

Tableau II.7 – Description du jeu de données utilisé pour optimiser le modèle IPSIM - Phoma sur tige préliminaire. Numéro de département, année, nombre de parcelles observées et nombre et niveau d'incidences observées.

4.2.2 Résultats

L'algorithme réalisé sous C++ montre une évolution du kappa en fonction du nombre de règles d'agrégation ajustées (Figure II.8). Cette évolution présente des fluctuations jusqu'à 3 règles d'agrégation ajustées puis une chute du kappa, cette dernière correspondant à une sur-paramétrisation. Ici, le nombre optimal de règles d'agrégation ajustées s'élève à 2 et 3 (en rouge sur la Figure II.8).

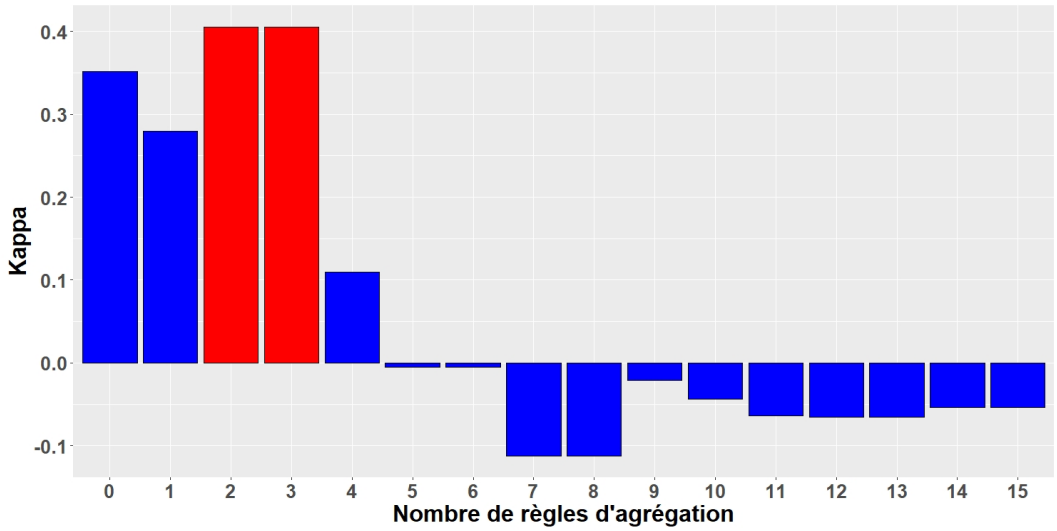


Figure II.8 – Évolution du kappa de Cohen pondéré quadratiquement avec le nombre de règles d'agrégation ajustées simultanément. En rouge le kappa maximum obtenu en validation-croisée. En bleu tous les autres kappas

Les mêmes kappas globaux (issus de la validation croisée) ont été observés entre les deux algorithmes issus des deux langages (R et C++), sans ajustement de règles d'agrégation, puis pour l'ajustement d'une règle, puis de deux règles d'agrégation simultanées (Tableau II.8).

Algorithme	**Sans ajustement**	**1 règle ajustée**	**2 règles ajustées**
Algorithme sous C	0,3517060367	0,2797297297	0,4055929782
Algorithme sous R	0,351706	0,2797297	0,405593

Tableau II.8 – Valeurs des kappas de Cohen pondérés quadratiquement obtenus avec l'algorithme codé sous C++ et codé sous R sans ajustement, puis avec l'ajustement d'une, puis de deux règles d'agrégation.

Le kappa obtenu pour la qualité de prédiction initiale est $\kappa = 0{,}352$ (Tableau II.8, Figure II.8), ce qui signifie que l'accord entre incidences observées et simulées est modéré (« fair », Landis et Koch, 1977).

Optimisation d'une règle d'agrégation

Pour l'ensemble des 156 situations du jeu de données, l'algorithme a retenu la modification d'une règle d'agrégation conduisant à la valeur la plus élevée du Kappa de Cohen pondéré quadratiquement. La meilleure modification proposée par l'algorithme concerne la 5eme règle de la table de l'incidence, avec le remplacement de la valeur initiale « Très élevée » par « Élevée ».

Lorsque les tables d'agrégations ont été modifiées avec la règle d'agrégation optimisée, c'est-à-dire lorsque la valeur « Très élevée » de la 5eme ligne de l'attribut agrégée « Incidence du Phoma

sur tige » a été remplacée par « Élevée », le κ mesuré sur l'ensemble du jeu de données s'est élevé de 0,352 à 0,612.

Optimisation de deux règles d'agrégation simultanées

La meilleure proposition de l'algorithme pour la modification de deux règles a porté sur deux tables différentes. La première modification proposée par l'algorithme est celle qui a été présentée au point précédent. La deuxième modification a porté sur la cinquième règle de la table de la gestion de l'inoculum primaire, avec le remplacement de la valeur initiale « Défavorable » par « Très favorable ». La valeur du κ est passée de 0,352 (accord modéré) à 0,639 (accord très bon).

L'algorithme non générique réalisé sous R a permis de valider les résultats de l'algorithme générique réalisé sous C++.

5. Optimisation de paramètres de modèles dynamiques simples

Cette section présente une méthode d'optimisation des paramètres pour un modèle quantitatif dynamique simple. Elle est basée sur la même stratégie que celle présentée dans la section précédente et repose sur l'approche proposée par Wallach et al. (2001).

Ils sont évalués à l'aide de trois critères statistiques.

Definition 5.1 — Mean Squared Error (MSE). Appelée aussi erreur des carrés moyenne (critère de la *MSE* Wallach *et al.* (2006), c'est une mesure de la déviation entre prédictions et observations. Sa gamme de variations est $[0\,;+\infty[$. Une *MSE* proche de 0 signifie que le modèle explique bien la variabilité observée :

$$MSE = \frac{1}{n}\sum_{i=1}^{i=n}(Y_i^{obs} - Y_i^{sim})^2 \tag{5.1}$$

Avec n: le nombre d'observations ; Y_i^{obs} la i^{ieme} observation; Y_i^{sim} la i^{ieme} prédiction.

Ce critère sert à caractériser la qualité de prédiction (MSE_p) des modèles quantitatifs au cours de validation-croisée et la qualité d'ajustement. La qualité de prédiction est la capacité du modèle à correctement prédire des observations qui n'ont pas servies à sa construction. La qualité d'ajustement est la capacité du modèle à correctement prédire des observations qui ont servies à sa construction.

Definition 5.2 — Biais. Il s'agit d'un indicateur des erreurs non-aléatoires. Il varie sur $]-\infty\,;+\infty[$. Un biais de 0 indique que les prédictions du modèle ne sont pas systématiquement sous ou

sur-estimées :

$$Bias = \frac{1}{n}\sum_{i=1}^{i=n}(Y_i^{obs} - Y_i^{sim}) \tag{5.2}$$

Avec n: le nombre d'observations ; Y_i^{obs} la i^{ime} observation; Y_i^{sim} the i^{ime} prédiction.

Definition 5.3 — Efficience. Cet indicateur compare la valeur de MSE du modèle avec la valeur de MSE obtenue avec un prédicteur naïf pris comme valeur moyenne des données observées. Il varie sur $]-\infty; 1]$. Un modèle réalisant des prédictions parfaites obtiendrait une efficience de 1. Une efficience de 0 indique que les prédictions sont égales à la moyenne des valeurs observées, c'est à dire que le modèle n'explique pas plus de la variabilité des mesures que la moyenne des mesures.

$$Efficiency = 1 - \frac{\sum_{i=1}^{i=n}(Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{i=n}(Y_i^{obs} - \bar{Y})^2} \tag{5.3}$$

Avec $\bar{Y}$ la moyenne des valeurs observées.

La relation entre valeurs simulées et prédites peut être représentée de la manière suivante (Équation 5.4) :

$$y_i = f(x_i;\theta) + \varepsilon_i \tag{5.4}$$

Où y_i est la valeur observée pour un élément i, $f(x_i;\theta)$ représente la prédiction correspondante avec f le modèle, x_i la valeur des variables explicatives pour l'élément i et θ, le vecteur de paramètres de f ; ε_i représente les résidus du modèle (la différence entre la simulation et l'observation de l'élément i).

Les paramètres θ peuvent être considérés comme des variables dont la valeur est constante, et qui dépendent du système étudié. Ces paramètres peuvent prendre différentes valeurs au sein d'intervalles. L'optimisation des paramètres d'un modèle f par rapport à un jeu de données peut-être réalisée par la méthode des moindres carrés ordinaires (**O**rdinary **L**east **S**quares (OLS), Wallach

et al., 2006). Cette méthode consiste à donner une valeur aux paramètres telle que (Équation 5.5) :

$$\widehat{\theta}_{OLS} = \arg\min_{p} \left\{ \sum_{i=1}^{i=n} [y_i - f(x_i;\theta)]^2 \right\} \tag{5.5}$$

Où n représente le nombre d'éléments et $\widehat{\theta}_{OLS}$ est le vecteur des valeurs des paramètres du modèle. Ce vecteur est la combinaison de paramètres. Il doit minimiser la *SCE* (*SSE* en anglais).

Definition 5.4 — Somme des Carrés des Écarts. Ce critère est un estimateur de la dispersion des valeurs autour de la moyenne.

$$SCE = \sum_{j=1}^{j=m} MSE_j * n_j \tag{5.6}$$

Avec j : une situation ; m : le nombre de situations ; MSE_j : la *MSE* de la situation j et n_j le nombre d'observations de la situation j.

Lorsque le modèle est basé sur une fonction non linéaire (*i.e.* différent de la forme $f(x) = ax + b$ avec a et b deux constantes $\in \mathbb{R}$ et $x \in \mathbb{R}$), un algorithme est alors nécessaire pour obtenir $\widehat{\theta}_{OLS}$. Deux grandes familles d'algorithme existent : celui de Gauss-Newton et celui de Nelder-Mead (Mead et Nelder, 1965). Le type d'algorithme utilisé lors de cette étude a été celui de Nelder-Mead car il peut minimiser n'importe quelle fonction, contrairement à celui de Gauss-Newton qui minimise spécifiquement la *SCE*. De plus il est simple à mettre en place. L'algorithme Nelder-Mead a besoin de valeurs de paramètres initiales pour disposer d'un point de départ x_0 puis créer un simplexe. Un simplexe est une figure à $p+1$ sommet(s) se situant dans un espace à p-dimension(s) avec p le nombre de paramètre(s) que l'on souhaite optimiser. Le point de départ est le premier sommet et les sommets suivants sont construits de sorte à définir un angle droit sur le point de départ et que :

$$x_j := x_0 + h_j e_j \tag{5.7}$$

où $j = 1, ..., p$ et h_j est un pas dans la direction du vecteur unitaire $e_j \in \mathbb{R}^n$.
Le principe général est le suivant : l'algorithme modifie ce simplexe dans l'espace à p-dimension(s) pour évaluer et proposer de nouvelles valeurs de paramètre(s). La Figure II.9 présente une illustration de ces processus pour deux paramètres.. Les axes représentent les paramètres. Chaque point

représente une valeur de f.

Dès l'obtention d'un simplexe l'algorithme procède en trois étapes (Mead et Nelder, 1965) :

1. la fonction à minimiser est calculée pour chaque sommet, qui sont ensuite ordonnés de la pire valeur à la meilleure valeur. Dans l'exemple présenté, trois sommets sont ordonnés dans l'ordre suivant : la pire valeur $f_p = max_j f_j$, $f_s = max_{j \neq p} f_j$, la seconde valeur et $f_m = min_{j \neq p} f_j$, la meilleure valeur (Figure II.9A).

2. le centroïde de la figure opposé au sommet pour lequel $f_p = max_j f_j$, est calculé de la manière suivante :
   $c := \frac{1}{n} \sum_{j \neq p} x_j$ (Figure II.9A).

3. la transformation qui consiste à remplacer le sommet avec f_p par un sommet ayant une meilleure valeur par réflexion (Figure II.9B), expansion (Figure II.9C) et contraction (Figure II.9D et E) ou rétrécissement (Figure II.9F) sur la droite qui relit le centroïde au mauvais sommet (Wright, 1996).

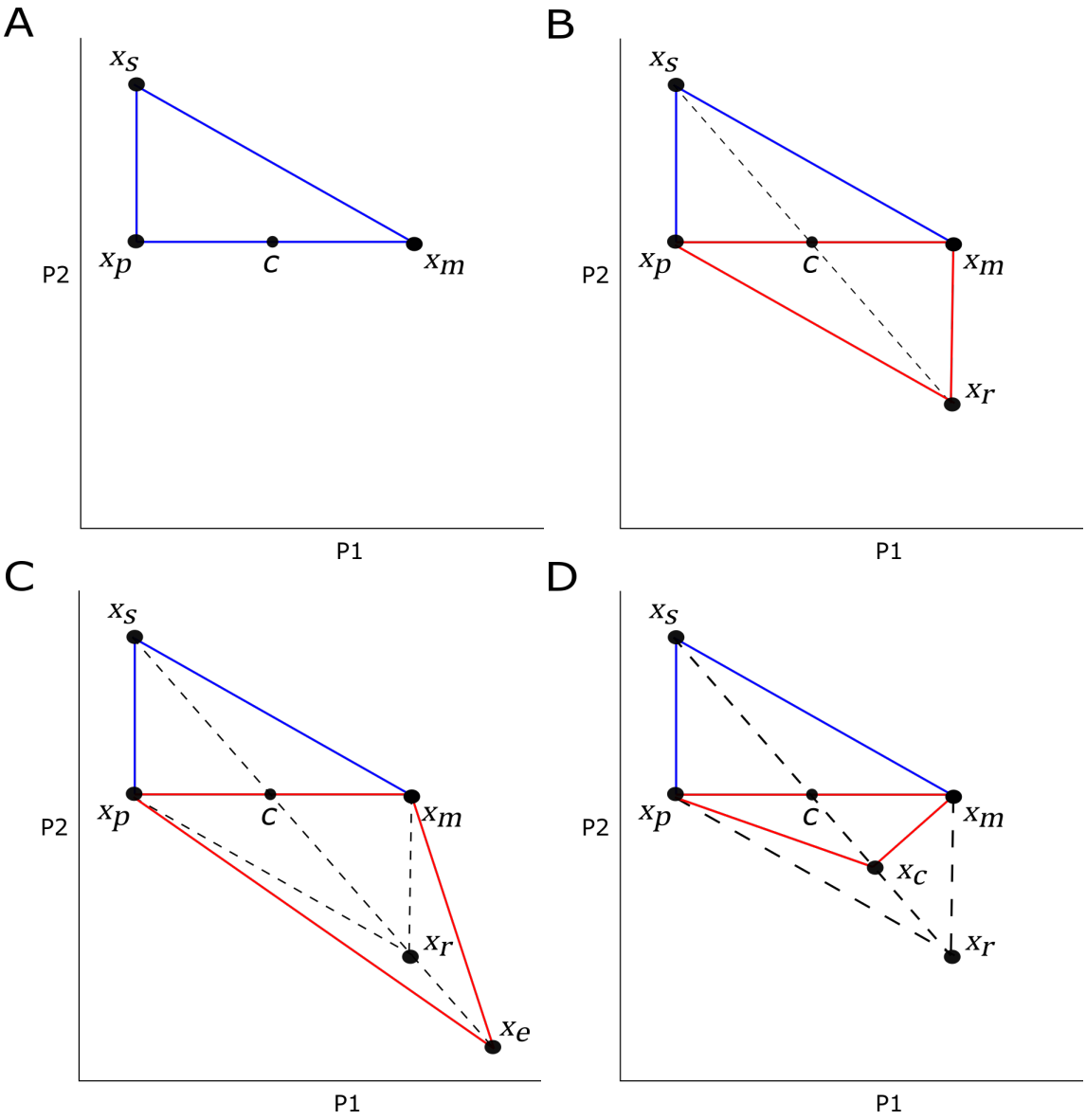
L'algorithme s'arrête si une des conditions suivantes est remplie :

— le nombre maximal d'itérations prédéfini dans l'algorithme est atteint.

— k évaluations successives de la fonction diffèrent de moins d'un seuil prédéfini (k étant un paramètre de l'algorithme).

— le simplexe est d'une taille inférieure à un seuil prédéfini, c'est à dire que les points du simplexe sont suffisamment proches. L'algorithme a alors convergé.

Un défaut de cet algorithme est sa tendance à la convergence sur un minimum local. La réinitialisation du simplexe initial est une manière de contourner ce problème (Törn, 1978 ; Voudouris, 1998 ; Huyer et Neumaier, 1999). Elle est réalisée en partant d'un nouveau point de départ, qui peut être aléatoire ou biaisé de sorte que l'exploration ne commence pas au même endroit que précédemment. Au cours de cette étude, cinq points de départ ont été utilisés dans les processus de validation croisée et trois points de départ ont été utilisés pour les ajustements finaux. Les valeurs de départ des paramètres sont étalées de manière régulière au sein de leurs intervalles.


A
x_s
x_p
c
x_m
P2
P1
B
x_s
x_p
c
x_m
x_r
P2
P1
C
x_s
x_p
c
x_m
x_r
x_e
P2
P1
D
x_s
x_p
c
x_m
x_c
x_r
P2
P1

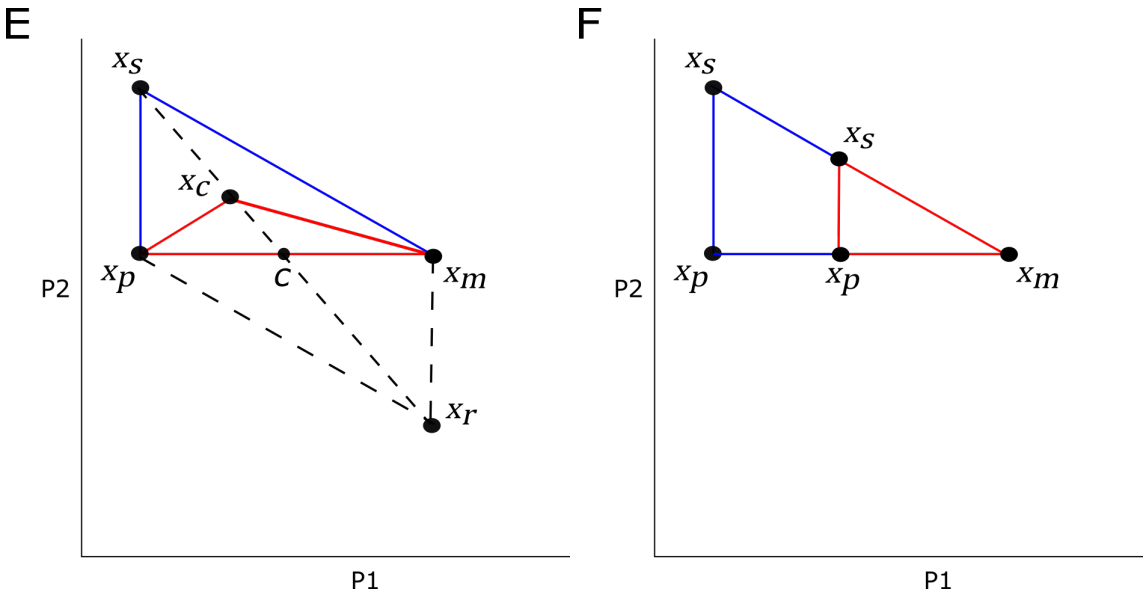


Figure II.9 – Illustration de la méthode algorithmique de Nelder-Mead sur un espace à 2-dimensions avec A : la mise en place d'un simplexe composé de trois points x_p (la pire valeur), x_s (la seconde valeur) et x_m (la meilleure valeur) et du centroïde du segment créé par x_p et x_m ; B : l'étape de réflection ; C : l'étape d'expansion ; D : l'étape de rétraction à l'extérieur du triangle ; E : l'étape de rétraction à l'intérieur du triangle et F : le rétrécissement.
Légende : • représente les points de simplexe et le centroïde ; en bleu, le simplexe présent ; en rouge, le nouveau simplexe ; en pointillés : les modifications en cours ; P1 et P2 les axes des paramètres.

Cet algorithme est implémenté sous différents langages. Deux protocoles ont servi à répondre aux objectifs de cet axe. Le premier protocole (développé par G. Quesnel du MIAT) repose sur la fonction nlopt de C++. Le deuxième protocole (développé par l'auteure de ce manuscrit) repose sur la fonction optim de R. Seule l'utilisation d'optim sera détaillée ici.

La fonction optim de la librairie « stats » a été conçue pour minimiser n'importe quelle fonction. Elle est codée de la manière suivante sous R :

```
optim(valeur-initiale, fonction-à-minimiser, arguments-pour-la fonction, méthode).
```

L'algorithme est mobilisé par optim. Entre parenthèses se trouvent les arguments d'entrée : « valeur-initiale » correspond aux valeurs initiales des paramètres à modifier. Ces valeurs sont à la base du premier simplexe. L'argument « fonction-à-minimiser » représente le critère statistique mesuré entre les observations et les prédictions. Cette fonction peut aussi nécessiter plusieurs arguments d'entrée nécessaires pour minimiser la *SCE*. La méthode utilisée dans ces études est la méthode « L-BGFS-B » qui permet de poser des bornes aux paramètres afin qu'ils gardent une cohérence biologique (lorsque le paramètre est observable).

III Modèles SimMat tournesol

1 Application à SimMat et *D. helianthi* . 109

2 Application à SimMat et *Leptosphaeria lindquistii* 155

2.1 Introduction

2.2 Matériel et méthodes

2.3 Résultats

2.4 Discussion

1. Application à SimMat et *D. helianthi*

Ce chapitre présente le développement et l'évaluation d'un nouveau modèle épidémiologique, facile d'utilisation, pour prévoir la dynamique de production d'inoculum primaire de *Diaporthe helianthi*, l'agent causal du phomopsis du tournesol. D'une manière générale, la connaissance de la temporalité de la production d'inoculum primaire d'agents pathogènes permet de mieux positionner les traitements fongicides, et/ou d'adapter les dates de semis dans une stratégie d'esquive. C'est le modèle générique SimMat, à la base conçu pour le phoma du colza, qui a été retenu pour développer un nouveau modèle adapté à *Diaporthe helianthi*. Ce choix est justifié par le fait que SimMat repose sur un formalisme mécaniste très simple, avec seulement sept paramètres, pour représenter la maturation d'organes de maturation, et la production de spores. Les variables d'entrée sont très simples à renseigner puisqu'il s'agit de la température moyenne journalière et la pluviométrie journalière. Une base de données comportant des observations provenant de dix sites-années (répartis sur six sites de deux départements d'Occitanie de 1989 à 1995, et 1997) a été mobilisée. Ces données ont permis l'adaptation de SimMat à l'aide de deux procédures algorithmiques originales basées sur des validations croisées contrôlant toute corrélation spatio-temporelles potentielle. L'objectif de ces procédures est d'identifier la combinaison de paramètres à ajuster (parmi $2^7 = 128$ possibilités) pour maximiser la qualité de prédiction des dynamiques de maturation des périthèces (procédure 1), ou la classe de précocité de la dynamique de maturation (procédure 2). La qualité de prédiction de la classe de précocité de production d'inoculum a été jugée satisfaisante (précision de 90%, et explication de 92% des données observées). Le modèle créé pourra contribuer, après une phase de test et d'appropriation par les praticiens, à l'élaboration des Bulletins de Santé du Végétal (Terres Inovia).

Ce chapitre correspond à l'article : Vedy-Zecchini M.A., Quesnel G., Bret-Mestries E., Robin M.H., Barbetti M., Brun F., Aubertot J.N. SimMat-*Diaporthe*: a model to predict the dynamics of *Diaporthe helianthi* inoculum production. Soumis à la revue *Crop Protection* le 11/11/2019.

SimMat-*Diaporthe*: a model to predict the dynamics of *Diaporthe helianthi* inoculum production

Marie-Anne Vedy-Zecchini[a], Gauthier Quesnel[b], Emmanuelle Bret-Mestries[c], Marie-Hélène Robin[a], Martin Barbetti[d], François Brun[e], Jean-Noël Aubertot[a*]

[a]*INRAE-INPT-ENSAT-EI-Purpan, University of Toulouse, UMR 1248 AGIR, Centre de recherche Occitanie-Toulouse, Auzeville, France*

[b]*INRAE, University of Toulouse, UR 1248 BIAT, Centre de recherche Occitanie-Toulouse, Auzeville, France*

[c]*Terres Inovia, Centre de recherche Occitanie-Toulouse, Auzeville, France*

[d]*University of Western Australia, Faculty of Science. School of Agriculture and Environment. Perth, Australia*

[e]*ACTA, Centre de recherche Occitanie-Toulouse, Auzeville, France*

---

* I am the corresponding author

Email address: jean-noel.aubertot@inrae.fr

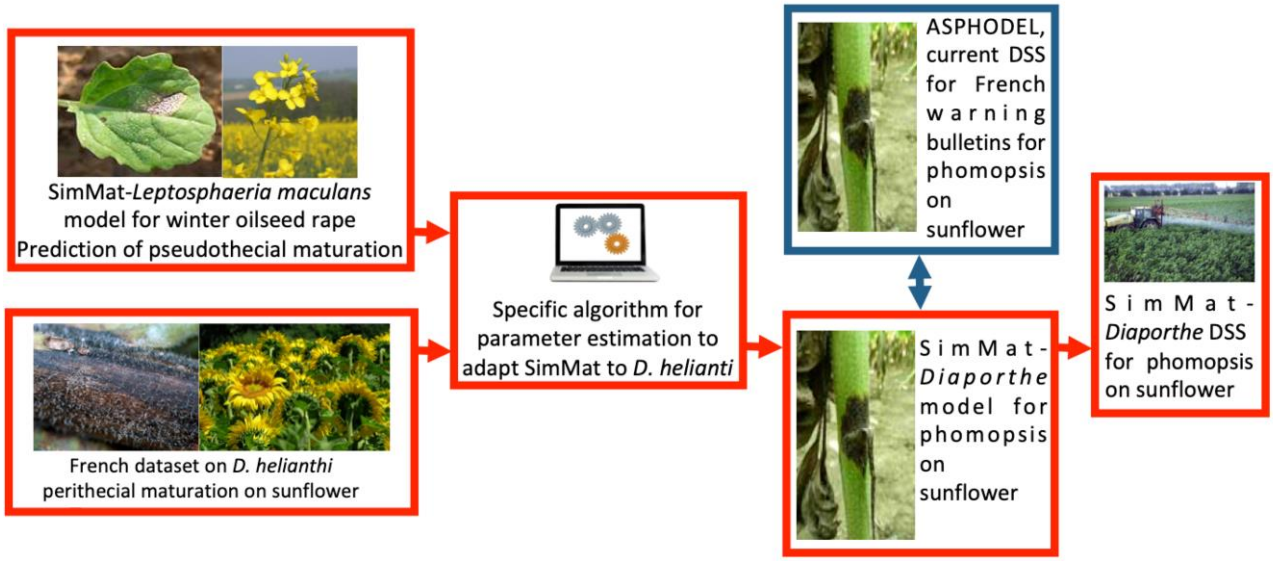
Manuscript submitted to Crop Protection November 11, 2019

Abstract

Phomopsis stem canker is one of the main pests on sunflower in France. In spring, *Diaporthe helianthi* develops pseudothecia on infected stubble that can release ascospores. In conjunction with host genetic resistance, fungicide applications can improve overall disease control. However, fungicide treatments need to be applied with a precise timing. In addition, the highest receptivity period for sunflower to ascospore showers could be shifted by altering sowing dates. A model is therefore essential to predict the dynamics of pseudothecial maturation. This paper presents SimMat-*Diaporthe*, a simple mechanistic dynamic model representing the *D. helianthi* primary inoculum production. The input variables are daily rainfall and temperature. The output variables are pseudothecial maturation dynamic and timing of ascospore release. It has seven parameters and is based on the assumption that pseudothecial maturation depends on the cumulated number of days favourable to maturation. A given day is considered as favourable to maturation if the mean daily temperature falls within a minimal and a maximal temperature range; and if the cumulated rainfall over a given number of days is higher than a given threshold. The proportion of mature pseudothecia that release their ascospores is a function of daily rainfall. A 10 sites-years French dataset was used to estimate parameters and assess its predictive quality. The initial parameter values were chosen using literature and expert knowledge. Original algorithms, using cross-validation preventing spatio-temporal correlations, permitted to identify the parameter combination to adjust (2^7=128 possibilities) to reach the best predictive quality either for the quantitative simulation of pseudothecial maturation dynamics (procedure 1), or the qualitative simulation of pseudothecial maturation precocity classes (procedure 2). The initial predictive quality of pseudothecial maturation dynamics by SimMat-*Diaporthe* was poor. Its Mean Squared Error of Prediction with initial parameters was $MSE_i^P = 0.05$; its bias was $Bias_i^P = -0.08$; its efficiency was $EF_i^P = -0.19$. Its initial quality of prediction for qualitative pseudothecial maturation precocity classes was much better, as described by the calculated precision ($P_{SM}^{P_i} = 0.70$) and quadratic weighted Cohen's kappa ($\kappa_{SM}^{P_i} = 0.68$). The optimization procedure 1 significantly improved the model, which has a limited quality of prediction for pseudothecial dynamics and performs much better for precocity classes of pseudothecial maturation ($MSE_{SM}^{P_{CV1}} = 0.036$; $Bias_{SM}^{P_{CV1}} = 0.024$;

$Efficiency_{SM}^{P_{CV1}} = 0.27$ and $P_{SM}^{A1} = 0.90$; $\kappa_{SM}^{A1} = 0.92$). Procedure 2 permitted to confirm the results obtained by procedure 1, with identical values for P_{SM}^{A1} and κ_{SM}^{A1}. The performances of SimMat were better than those of ASPHODEL, the historical model for the disease in France. A new model is now available to help decision making for sunflower protection in France. In addition, the new generic algorithms presented in this paper are now available to improve similar models.

Keywords: primary inoculum, parameter estimation, algorithm, simulation model, cross-validation, Decision Support System



Graphical abstract. Design of SimMat-*Diaporthe*, a simple mechanistic dynamic model representing *Diaporthe helianthi* primary ascospore inoculum production, to complement the historical ASPHODEL model in order to provide a new Decision Support System to control phomopsis stem canker on sunflower.

Highlights

- A new model is now available to predict *D. helianthi* primary inoculum production
- This model has a fair quality of prediction for pseudothecial maturation dynamics
- This model has a very good predictive quality of primary inoculum production precocity
- It is available for a new Decision Support System for French warning bulletins
- New algorithms are now available to implement the same approach for similar models

92 **1. Introduction**

93 *Diaporthe helianthi* is a phytopathogenic fungus of the Ascomycota phylum. It causes phomopsis stem
94 canker, one of the main fungal diseases on sunflower in the world (Thompson et al., 2011; Masirevic, 2016).
95 From 1 to 3 $q.ha^{-1}$ (10 to 30 $g.m^{-2}$) of yield are lost when 10% of plants of a given field have girdling spots
96 (Estragnat, 1992; Masirevic and Gulya, 1992). The fungus was identified for the first time in 1980 in
97 Yugoslavia (Maric and Masirevic, 1980). It was observed in France for the first time in 1984 in commercial
98 fields of Lauragais (Lamarque and Perny, 1980). Currently, in France, it is most prevalent and serious in the
99 South-West region (Perny, 1992). However, it is present in all French production areas of sunflower (Délos et
100 al., 1995).

101 After harvest, fungus mycelium survives on infected stubble left at soil surface until its decomposition.
102 Generally, this process takes ca. one year under French conditions (Slyusa et al., 1988). D. *helianthi*
103 accomplishes its sexual phase (teleomorph phase) and its asexual phase (anamorph phase) on this infected
104 stubble. The sexual reproduction (the only infectious phase of D. *helianthi* cycle; Mihaljcevic et al., 1985) is
105 the only phase considered in this study. The mycelium of *D. helianthi* performs its sexual reproduction by
106 producing pseudothecia on stubble (Muntanola-Cvetkovic et al., 1988). This production begins with the
107 formation of protopseudothecia in autumn (Muntanola-Cvetkovic et al., 1988) when temperatures are both
108 lower than 10°C and higher than 0°C and rainfall wets the infected stubble (Fayret and Assemat, 1987). The
109 maturation process stops in winter and restarts at early spring (Muntanola-Cvetkovic et al., 1988). Eight
110 ascospores develop inside each ascus during the maturation process in spring. Ascospores are released during
111 periods of rainfall from May to August and cause leaf infections during May and June (Petrov et al., 1981;
112 Fayret and Assemat, 1987). It is considered that the release of ascospores is significant when the percentage of
113 mature pseudothecia reaches 50% (Délos et al., 1995). The phenological period of sunflower that is most
114 receptive to ascospores is between flower bud and the beginning of flowering (Délos et al., 1995). Ascospores
115 can germinate during a period of ten days after their release. Fungus mycelium colonizes plant tissues
116 (Muntanola-Cvetkovic et al., 1988), and large necrotic symptoms appear on leaves. During 40 days, the fungus

progresses through main leaf veins, then reaches leaf petiole, and finally attacks the stem, generating girdling symptoms, and killing plants in most severe cases (Muntanola-Cvetkovic et al., 1988). Few preventive control methods exist against phomopsis stem canker. For example, actions on infected stubble (grinding and burying) increase their degradation, thus decreasing *D. helianthi* survival and primary inoculum production (Masirevic and Gulya, 1992). Crop sequences with rarer sunflowers imply lower primary inoculum production at the landscape level (Jinga et al., 1992). Crop management techniques, such as sowing rate, nitrogen fertilization and irrigation can affect the number of infected plants. A lower sowing rate (a decrease from seven to five plants.m^{-2}) decreases infection by 22% (Debaeke et al., 2003). A balanced nitrogen fertilization can limit the proportion of plants killed by phomopsis stem canker. On the contrary, irrigation applied at flowering increases the proportion of infected plants (Debaeke et al., 2003; Debaeke and Moinard, 2010; Desanlis et al., 2013). A disease escape strategy can be adopted. For example, in France, late sowing (May instead of April) can shift the receptive period of sunflower (at the flower bud development period) to *D. helianthi* ascospores (Debaeke et al., 2003). However, these cultural control methods have only limited and partial effects on the disease and require to be utilized in conjunction with resistant cultivars. Sunflower cultivars are classified into four levels of susceptibility: susceptible, slightly susceptible, very slightly susceptible and resistant (Terres Inovia, 2019). Fungicidal control is the only curative method available to farmers. For example, combinations of a strobilurin with a triazole or a carboxamide fungicide are officially registered against phomopsis stem canker in France. Their efficacy is considered sufficient as they can increase yield by 10 q.ha^{-1} (100 g.m^{-2}) on average, in a field where 50% of sunflower plant have lesions (Duroueix and Penaud, 2013). Their use depends on sunflower stage, cultivar susceptibility, and a risk analysis provided by the Bulletin de Santé du Végétal based on ASPHODEL simulations (Terres Inovia, 2019). To date, ASPHODEL is the only model available to help manage phomopsis stem canker on sunflower (Délos and Moinard, 1997). It has been used for decades for sunflower phomopsis risk analyses provided to French farmers via special bulletins called “Avertissements Agricoles” (www.pestobserver.eu). It is still currently used for the risk analyses provided by the Bulletin de Santé du Végétal since 2009 (Chambre d’Agriculture Occitanie, 2018). It is a deterministic and dynamic model predicting daily projections and contaminations as a function of five input variables (hourly temperature and

143 hygrometry; and daily minimum temperature, maximum temperature and rainfall). It has eleven parameters.
144 ASPHODEL suffers from two major drawbacks. The first one is that it needs hourly weather variables that can
145 be difficult to collect. In addition, the dataset used for its calibration is no longer available, which prevents
146 from estimating its quality of prediction. Other models have been developed in order to predict inoculum
147 production of Ascomycota. Among them, SimMat is a mechanistic, deterministic and dynamic model
148 (Aubertot et al., 2006). It is derived from the Blackleg Sporacle model developed in Australia in 2001 to
149 manage phoma stem canker on oilseed rape (Salam et al., 2003), caused by *Leptosphaeria maculans*. SimMat
150 has been developed for Polish (Aubertot et al., 2006) and French conditions (Lô-Pelzer et al., 2010). This
151 model predicts the daily rate of maturation of the fruiting bodies (pseudothecia) and the daily rate of maturation
152 of pseudothecia releasing ascospores. It has two input weather variables that are easy to provide (mean daily
153 temperature and daily rainfall) and a total of seven parameters. It is currently used in France to help manage
154 phoma stem canker on oilseed rape in Normandy (Bulletin de Santé du Végétal). We chose to adapt SimMat
155 to phomopsis on sunflower for three reasons. First, it has simple and easily obtainable input variables and is
156 therefore easier to use. Second, it has a simpler and more mechanistic formalism than ASPHODEL. Third, *D.*
157 *helianthi* shares many similar biological traits with *L. maculans*. They are both Ascomycota and their sexual
158 and asexual reproduction cycles are similar, with both teleomorph and anamorph phases. In addition, the
159 maturation of fruiting bodies (pseudothecia for *L. maculans* and *D. helianthi*) are similarly affected by weather
160 conditions (temperature and humidity). This study first presents an analysis of the respective predictive quality
161 of ASPHODEL and SimMat-*Diaporthe* models, with initial parameters taken from the literature. Then, the
162 adaptation of SimMat-*Leptosphaeria maculans* to *D. helianthi* is presented. Finally, we compare and discuss
163 the benefits and limitations of both modelling approaches to support the development of a new Decision
164 Support System to control Phomopsis on sunflower.

165

166

167

2. Materials and methods

2.1. Description of the generic structure of the SimMat model

SimMat is based on the concept of cumulative days favourable to maturation of sexual fruiting bodies over a given period of time. The two output variables of the model are: the daily rate of mature fruiting bodies and the daily rate of mature fruiting bodies releasing their ascospores. In this model, a given day *i* is considered as favourable to maturation if the mean daily temperature $\theta(i)$ is higher than a given threshold (θ_{min}^{SM}) and lower than another threshold (θ_{max}^{SM}); and if the mean daily cumulated precipitations over the last n_d days is higher than a threshold R_{min}^{SM}. Days favourable to maturation are then cumulated since the beginning of simulation:

$$CDFM(d) = \sum_{i=1}^{i=d} \delta_{rs}(i) \quad (1)$$

Where i is a daily summation index; d is the considered day; $\delta_{rs}(i)$ is the Kronecker symbol with:

$$\begin{cases} r = s \; if \; \theta(i) > \theta_{min}^{SM} \; and \; \theta(i) < \theta_{max}^{SM} \; and \; \sum_{j=0}^{j=n_d-1} R(i-j) > R_{min}^{SM} \\ and \\ r \neq s \; otherwise \end{cases} \quad (2)$$

where *j* is a cumulation index over the last n_d days.

Inoculum produced by stubble left at soil surface is supposed to be non-limiting. It is hypothesized that the probability that a fruiting body is mature follows a Gaussian distribution, like many other biological processes.

This distribution is characterized by two parameters: N_{FD}, the expected number of favourable days to maturation and σ_{FD}, the associated standard deviation. A dynamic binary variable "day favourable to maturation" describes whether a day is favourable or not to maturation. It is equal to 1 if the day is favourable to maturation, otherwise it is equal to 0. The proportion of mature fruiting bodies P_{MP} *(d)* at *d* day is given by Equation 3. R_{EP} *(d)* is the cumulated proportion of mature fruiting bodies that have released their ascospores at the day *d* (Equation 4). R_{MFP}*(d)* is the rate of mature fruiting bodies ready to release their ascospores. It is

189 calculated as the proportion of mature fruiting bodies divided by the proportion of non-empty fruiting bodies.
190 It is calculated using Equation 5. The variable *R(d)* is the daily rainfall at day *d*.

191

192 $$P_{MP}(d) \approx \frac{1}{\sqrt{2\pi\sigma_{FD}}} \int_0^{CDFM(d)} \exp\left(\frac{-(x-N_{FD})^2}{2\sigma_{FD}}\right) dx \quad (3)$$

193 $$R_{EP}(d) = R_{EP}(d-1) + (P_{MP}(d) - R_{EP}(d-1))(1 - \exp(-k * r(d))) \quad (4)$$

194 $$R_{MFP}(d) = \frac{P_{MP}(d) - R_{EP}(d)}{1 - R_{EP}(d)} \quad (5)$$

195 The seven parameters of SimMat are described in Table 1.

196

Parameter	Definition	Unit
θ_{min}^{SM}	Minimum daily temperature threshold enabling the fruiting bodies maturation	°C
θ_{max}^{SM}	Maximum daily temperature threshold enabling the fruiting bodies maturation	°C
R_{min}^{SM}	Minimal rainfall threshold enabling the fruiting bodies maturation	mm
N_d	Number of days for cumulating rainfall	day
N_{FD}	Expected number of favourable days	day
σ_{FD}	Standard deviation of favourable days	day
k	Parameter related to the proportion of mature pseudothecia that will release their ascospores as a function of daily rainfall	mm^{-1}

197

198 **Table 1. Parameters, definitions and units of SimMat's parameters.**

199 SimMat was initially calibrated for pseudothecia of *L. maculans* responsible for phoma stem canker
200 on oilseed rape. Then, simulations usually started by mid-July, as it is the harvesting period of oilseed

rape in France. For phomopsis on sunflower, a study performed in Yugoslavia showed that mature pseudothecia of *D. helianthi* could be observed in February (Muntanola-Cvetkovic et al., 1988). This date is the beginning date in ASPHODEL. We therefore chose to start simulations on February 1st for SimMat-*Diaporthe*.

2.2. Short description of the ASPHODEL model

ASPHODEL is composed of three sub-models (fungus, plants and fungicide) that predict epidemics of *D. helianthi* (Délos and Moinard, 1997). The fungus sub-model simulates pseudothecial maturation, ascospore releases and leaf infection, using February 1st as the first day of simulation. The input variables of the fungus sub-model are hourly or three hourly relative humidity and temperature, and daily rainfall. This sub-model has 11 parameters. The present study only considered pseudothecial maturation and ascospore releases. The pseudothecial maturation proportion is computed using hourly or three hourly relative humidity and temperature. The daily ascospore releases begin when pseudothecial maturation proportion reaches 0.50 and depend on daily rainfall.

2.3. Description of the dataset used to evaluate the adjustment quality of ASPHODEL and to estimate SimMat-Diaporthe parameters

2.3.1. Pseudothecial maturation

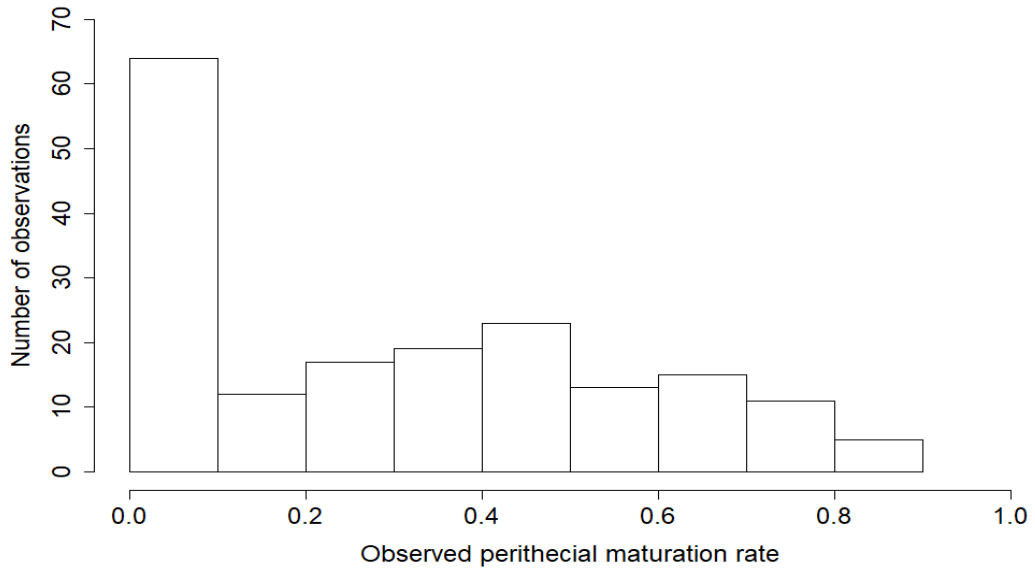
A dataset of ten sites-years was used to evaluate both models and estimate parameters of SimMat-*Diaporthe*. This dataset was created by the Service Régional de la Protection des Végétaux Midi-Pyrénées to analyse the impact of weather variables on pseudothecial maturation and ascospore release (Délos and Moinard, 1997). This dataset is only a small part of the dataset that had been used to estimate the parameters of ASPHODEL. Unfortunately, the rest of this dataset was unavailable. *D. helianthi* pseudothecial maturation and ascospore release were monitored on ten sunflower fields in Haute-Garonne

and Tarn, between 1989 and 1997 and was regularly sampled to dissect and observe pseudothecia under the microscope (5 infected stubble x 5 pseudothecia x 10 asci per sample; Fabregue, 1992) on 12 to 25 dates from March to August. In all, 179 observations were performed. The pseudothecial maturation rate ranged from 0 to 0.87. The dataset is described in Table 2. Maturation of *D. helianthi* fruiting bodies has been monitored in a few studies in Yugoslavia, Romania and France, respectively (Petrov et al., 1981; Jinga et al., 1987; Fabregue, 1992). Pseudothecial maturation occurs more or less rapidly, as a function of weather scenario. Using these references and our own expert knowledge, we proposed three ordinal classes to characterize pseudothecial maturation dynamic precocity of *D. helianthi*: early, intermediate and late. We considered that a dynamic was "early" if the threshold of 50% of mature pseudothecia was reached before the end of May, "intermediate" in June or July; and "late" in August or latter, respectively. There were five situations with an early pseudothecial maturation dynamic: the threshold of 0.50 of mature pseudothecia was reached in May due to a warm winter (Quint-Fonsegrives in 1990, Ondes in 1993, Balma in 1994 and 1995 and Auzeville in 1995). Three situations presented an intermediate precocity, the threshold of 0.50 of mature pseudothecia was reached in June (Cahuzac-sur-Vère in 1990, Quint-Fonsegrives in 1991 and Balma in 1992). Then, the two Revel situations in 1989 presented late dynamics, in spite of a warm winter. In these situations, maturation began at the end of May and ended in August. These late maturation dynamics were explained by low rainfall during spring.

Geographical provenance of infected stubble	Year	N_{obs}	Observation period	Range of pseudothecial maturation rate	Maturation precocity
Revel (31)	1989	24	21/04/1989 - 28/08/1989	0.0 - 0.76	Late
Revel (31)	1989	24	21/04/1989 - 28/08/1989	0.0 - 0.63	Late
Cahuzac-sur-Vère (81)	1990	23	30/03/1990 - 26/07/1990	0.0 - 0.73	Intermediate
Quint-Fonsegrives (31)	1990	14	15/03/1990 - 22/07/1990	0.0 - 0.84	Early
Quint-Fonsegrives (31)	1991	14	15/03/1991 - 22/07/1991	0.0 - 0.83	Intermediate
Balma (31)	1992	25	15/04/1992 - 27/07/1992	0.0 - 0.75	Intermediate
Ondes (31)	1993	17	15/04/1993 - 21/07/1993	0.0 - 0.87	Early
Balma (31)	1994	12	28/03/1994 - 26/05/1994	0.08 - 0.50	Early
Balma (31)	1995	14	20/03/1995 - 29/05/1995	0.12 - 0.50	Early
Auzeville (31)	1997	12	10/04/1997 - 27/06/1997	0.10 - 0.72	Early

Table 2. Presentation of the dataset (composed of 179 observations) used for modeling. Geographical provenance of infected stubble (31: Haute-Garonne, 81: Tarn); Maturation year; N_{obs}: number of observations; Observation period; Range of pseudothecial maturation rate (minimal and maximal observed values); and Maturation precocity.

The distribution of pseudothecial maturation rates is presented Figure 1. More than one-third of the rate observed was comprised between 0 and 0.1 (64 observations out of the 179).



257

258 **Figure 1. Distribution of pseudothecial maturation rates observed on the entire dataset (sample size = 179).**

259

260 *2.3.2. Weather variables*

261 The SAFRAN database (Météo France) was used to extract mean daily temperatures and rainfall,
262 along with hourly temperatures and relative humidities for each site presented in the previous section.
263 This national database uses a 8 x 8 km^2 square grid pattern. Each site was associated to one and only
264 one node, according to its position.

265

266 *2.4. Criteria used to evaluate quantitative and qualitative predictions*

267 *2.4.1. Evaluation of quantitative predictions*

268 Quantitative predictions were evaluated using three criteria. The Mean Squared Error (*MSE* criterion
269 (Wallach et al., 2006), presented in Equation 6) is a measure of deviation between predictions and
270 observations. Its ranges [0; +∞[. The closer to 0 the *MSE*, the better the model.

$$MSE = (1/n)\sum_{i=1}^{i=n}(Y_i^{obs} - Y_i^{sim})^2 \quad (6)$$

With *n*: observations number; Y_i^{obs} the ith observation; Y_i^{sim} the ith prediction.

The *Bias* (Equation 7) is an indicator of non-random errors. It ranges]−∞; +∞[. A bias close to zero indicates that model predictions do not suffer from a strong systematic deviation.

$$Bias = (1/n)\sum_{i=1}^{i=n}(Y_i^{obs} - Y_i^{sim}) \quad (7)$$

The *Efficiency* (Equation 8) compares the *MSE* value of the model with *MSE* value obtained with a naive predictor taken as the mean value of the observed data. It ranges]−∞; 1]. A model with perfect predictions has an efficiency of 1. A model with the mean observed values as predictor for each situation has an efficiency of zero.

$$Efficiency = 1 - [\sum_{i=1}^{i=n}\left(Y_i^{obs} - Y_i^{sim}\right)^2 / \sum_{i=1}^{i=n}\left(Y_i^{obs} - \hat{Y}\right)^2 \quad (8)$$

With $\hat{Y}$, the mean of observed values.

2.4.2. Evaluation of qualitative predictions

Confusion matrices are contingency tables that presents the distribution of observed and simulated ordinal variables (Witten et al, 2011). They were summarized by two criteria:

- the precision (*P)*, equal to the proportion of correctly predicted situations

- the squared weighted κ of Cohen (Equation 9; Agresti and Kateri, 2010). This latter criterion penalizes errors with weights equal to the squared value of the number of difference classes between observed and predicted values. This criterion was chosen because it can be interpreted as the proportion of data variability explained by the model (Fleiss and Cohen, 1973). T he κ ranges from -1 (total disagreement) to 1 (perfect agreement). A null weighted κ indicates a total independency between the considered

ordinal variables. The interpretations of κ values were performed using a reading grid proposed by Landis and Koch (1977).

$$\kappa = \frac{P_o(w) - P_e(w)}{1 - P_e(w)} \quad (9)$$

Where $P_o(w) = \sum_{i=1}^{i=n_c} \sum_{j=1}^{j=n_c} w_{ij} p_{ij}$ is the observed agreement of the considered model and

$P_o(w) = \sum_{i=1}^{i=n_c} \sum_{j=1}^{j=n_c} w_{ij} p_{i.} p_{.j}$ is the agreement of a random model, and $w_{ij} = 1 - ((i-j)^2/(n_c-1)^2)$; $p_{ij} = n_{ij}/n; p_{i.} = n_{i.}/n; p_{.j} = n_{.j}/n.$

i is the ith column of the confusion matrix, j is the jth row of the confusion matrix. n_c is the number of classes, n is the total number of observations.

2.5. Evaluation of ASPHODEL's adjustment quality

ASPHODEL does not predict pseudothecial maturation rates. It was therefore not possible to assess its quantitative quality of prediction, like what was done for SimMat-*Diaporthe*. ASPHODEL was evaluated on its capacity to predict qualitative classes of maturation precocity, using a confusion matrix and associated statistics: precision and squared weighted κ of Cohen, denoted $P^A_{ASPHODEL}$ and $\kappa^A_{ASPHODEL}$ respectively (see Table 3). It is important to highlight that we would not assess the quality of prediction of ASPHODEL since the situations of the presented dataset were already used to parameterize ASPHODEL. Due to the limited number of sites-years in the dataset (10), it was not possible to perform a cross-validation approach.

2.6. Parameter estimation of SimMat-Diaporthe using quantitative maturation dynamics

312 The estimation of SimMat-*Diaporthe*'s parameters to predict pseudothecial maturation rate (R_{MFP}) was
313 performed using the dataset previously presented, and the MSE criterion. We designed an innovative procedure
314 composed of three steps that will be denoted thereafter "procedure 1".

315 First, initial values of parameters were determined using literature and expertise. The values for temperature
316 thresholds were two published (Table 7). The value of the rainfall threshold for SimMat-*Diaporthe*, $R_{min}^{\mathrm{SM}-D.h.}$,
317 was unknown. We thus chose the same value as for SimMat-*Leptosphaeria maculans* (Aubertot et al., 2006):
318 $R_{min}^{\mathrm{SM}-D.h.} = R_{min}^{\mathrm{SM}-L.maculans}$. The other parameters were estimated by expertise (Table 7). The predictions
319 obtained with SimMat-*Diaporthe* with these initial values were then compared with observations. This
320 comparison provided an evaluation of the initial predictive quality of SimMat-*Diaporthe*.

321 Second, the best combination of parameters to be adjusted was identified using a cross-validation method
322 without spatio-temporal correlation. Each situation of the dataset was used to assess the quality of prediction
323 of the model after adjustment of its parameters using all other independent situations (i.e. situations with a
324 different year and a different location). In all, there were $\sum_{i=1}^{i=n} C_i^m = 2^n$ possible combinations of parameters
325 to adjust (including the case without any adjustment). Each of the 2^7 = 128 possibilities was considered and
326 the parameter combination with the lowest error of prediction was identified and selected (Table 3).

327 At last, the identified parameter combination was adjusted using the entire dataset to obtain the best possible
328 model. After this second estimation of parameters, SimMat-*Diaporthe* was evaluated on its capacity to predict
329 maturation dynamics observed in the dataset using the same statistical criteria as presented in section 2.4.1
330 (evaluation of the adjustment quality). They are denoted MSE_{SM}^{A1}, $Bias_{SM}^{A1}$ and $\mathrm{Efficiency}_{SM}^{A1}$. The quality of
331 adjustment was also obtained using precocities and was characterized by the squared weighted Cohen's κ_{SM}^{A1}
332 and the precision P_{SM}^{A1} (Table 3).

333

334

335

336 *2.7. Parameter estimation of SimMat-Diaporthe using qualitative precocities*

337 A second procedure was performed and was denoted "procedure 2". Each of the $2^7 = 128$ possibilities
338 of parameter combination was adjusted using the entire dataset to obtain the best κ criterion. The precision was
339 also computed. They are denoted κ_{SM}^{A2} and P_{SM}^{A2}. The parameter combination with the lowest error of prediction
340 was identified and selected. The quality of adjustment was also described according to the quantitative
341 pseudothecial maturation dynamics and was characterized by the MSE_{SM}^{A2}, the $Bias_{SM}^{A2}$ and Efficiency$_{SM}^{A2}$ (Table
342 3).

343 All the statistical criteria used in this study are synthesized in Table 3.

344

345 *2.8. Computer implementation*

346 SimMat-*Diaporthe* and numerical procedures 1 and 2 were implemented both in R (https://www.R-
347 project.org) and C++ (ISO/IEC 14882:2017) in order to validate computer codes. In R, the optim function of
348 the {stats} library was used to perform optimizations (Ordinary Least Squares), with the method "L-BFGS-
349 B" that constrains parameters between a lower and an upper bound (Byrd et al., 1995). The irr and vcd packages
350 were used to compute the squared weighted Cohen's κ.

351 In C++, the {NLopt} library (Johnson, 2009) was used to perform optimizations. {NLopt} is an open-source
352 library for nonlinear optimization, providing a common interface for a number of different free optimization
353 routines. With the SimMat-*Diaporthe* model, the best time constrained algorithm for optimization was the
354 Nelder-Mead simplex algorithm (Nelder and Mead, 1965), updated with an explicit support for bound
355 constraints (Box, 1965). The optimization algorithm from {NLopt} was run multiple times with different
356 starting points in order avoid local minima. Solution searching was stopped when the best value for the
357 considered criterion (i.e. *MSE* and squared weighted Cohen's κ) was found five times. The C++ code used the
358 C++17 thread system to perform parallelization. The overall optimization typically took two minutes on an

359 Ubuntu Linux 16.04 computer with an Intel (R) CPU Xeon (R) CPU E5-1620 v2 @ 3.70GHz and 32 Gb
360 memory.

361

Symbol	Criterion	Model	Procedure	Quality	Step
$P^{A}_{ASPHODEL}$	Precision	ASPHODEL	-	Adjustment	Final
$P^{P_i}_{SM}$	Precision	SimMat-*D.*	-	Prediction	Initial
P^{A1}_{SM}	Precision	SimMat-*D.*	1	Adjustment	Final
P^{A2}_{SM}	Precision	SimMat-*D.*	2	Adjustment	Final
$\kappa^{A}_{ASPHODEL}$	Squared weighted Cohen's Kappa	ASPHODEL	-	Adjustment	Final
$\kappa^{P_i}_{SM}$	Squared weighted Cohen's Kappa	SimMat-*D.*	-	Prediction	Initial
κ^{A1}_{SM}	Squared weighted Cohen's Kappa	SimMat-*D.*	1	Adjustment	Final
κ^{A2}_{SM}	Squared weighted Cohen's Kappa	SimMat-*D.*	2	Adjustment	Final
$MSE^{P_i}_{SM}$	Mean Squared Error	SimMat-*D.*	-	Prediction	Initial
$MSE^{P_{CV1}}_{SM}$	Mean Squared Error	SimMat-*D.*	1	Prediction	C.V.
MSE^{A1}_{SM}	Mean Squared Error	SimMat-*D.*	1	Adjustment	Final
MSE^{A2}_{SM}	Mean Squared Error	SimMat-*D.*	2	Adjustment	Final
$Bias^{P_i}_{SM}$	Bias	SimMat-*D.*	-	Prediction	Initial
$Bias^{P_{CV1}}_{SM}$	Bias	SimMat-*D.*	1	Prediction	C.V.
$Bias^{A1}_{SM}$	Bias	SimMat-*D.*	1	Adjustment	Final
$Bias^{A2}_{SM}$	Bias	SimMat-*D.*	2	Adjustment	Final
$\text{Efficiency}^{P_i}_{SM}$	Efficiency	SimMat-*D.*	-	Prediction	Initial
$\text{Efficiency}^{P_{CV1}}_{SM}$	Efficiency	SimMat-*D.*	1	Prediction	C.V.
$\text{Efficiency}^{A1}_{SM}$	Efficiency	SimMat-*D.*	1	Adjustment	Final
$\text{Efficiency}^{A2}_{SM}$	Efficiency	SimMat-*D.*	2	Adjustment	Final

362 **Table 3. Statistical criteria used in the study. SM: SimMat. CV: Cross-Validation.**

3. Results

3.1. Evaluation of ASPHODEL's adjustment quality

The confusion matrix built with observed precocities and ASPHODEL's predictions are presented in Table 4A.

The associated precision was $P^{A}_{ASPHODEL} = 0.60$ (Table 5). The squared weighted Cohen's Kappa computed for the predictive quality of ASPHODEL was $\kappa^{A}_{ASPHODEL}$= 0.057 (Table 5). This indicates a very low agreement between ASPHODEL's predictions and observations. Despite a fair precision, this poor agreement results mainly from the two late situations of the dataset of Revel 1989 (Table 4A). ASPHODEL correctly predicted the five early dynamics and one intermediate dynamic out of three (Balma 1992, Table 4A). The intermediate situations from Cahuzac-sur-Vère in 1990 and Quint-Fonsegrives in 1991 were incorrectly predicted as early (Table 4A).

A

		Predicted precocity			
		Early	Intermediate	Late	Total
Observed precocity	Early	5	0	0	5
	Intermediate	2	1	0	3
	Late	2	0	0	2
	Total	9	1	0	10

375

B

		Predicted precocity			
		Early	Intermediate	Late	Total
Observed precocity	Early	5	0	0	5
	Intermediate	1	2	0	3
	Late	0	2	0	2
	Total	6	4	0	10

376

C

		Predicted precocity			
		Early	Intermediate	Late	Total
Observed precocity	Early	5	0	0	5
	Intermediate	1	2	0	3
	Late	0	0	2	2
	Total	6	4	0	10

377 **Table 4. Confusion matrices with observed and predicted pseudothecial maturation precocity classes.**

378 **A, ASPHODEL; B, SimMat-*Diaporthe* before parameter estimation and C, SimMat-*Diaporthe* after**

379 **parameter estimation using procedure 1 or 2 (results were identical with the two procedures).**

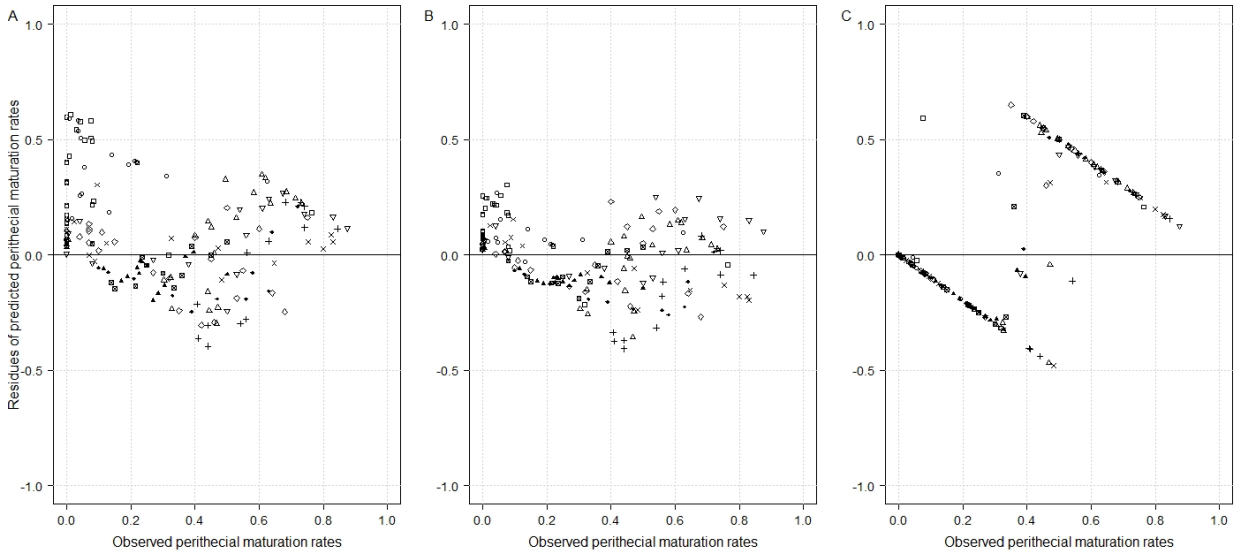
P	Value	κ	Value
$P^{A}_{ASPHODEL}$	0.60	$\kappa^{A}_{ASPHODEL}$	0.057
P^{Pi}_{SM}	0.70	κ^{Pi}_{SM}	0.68
P^{A1}_{SM}	0.90	κ^{A1}_{SM}	0.92
P^{A2}_{SM}	0.90	κ^{Pi}_{SM}	0.92

Table 5. Statistical criteria for predictions of maturation precocity classes. Values of the precision *P* and the squared weighted Cohen's Kappa computed for the initial predictive quality of SimMat-*Diaporthe* and the adjustment quality of ASPHODEL, and SimMat-*Diaporthe* after use of procedures 1 and 2. Criteria are described Table 3.

3.2. Parameter estimation of SimMat-Diaporthe using quantitative data (procedure 1)

3.2.1. Evaluation of the initial predictive quality of SimMat-Diaporthe on pseudothecial maturation dynamics

Figure 3A represents the comparison between observed pseudothecial maturation rates and values predicted with SimMat-*Diaporthe*, using initial parameter values. In general, the model over-estimated observed pseudothecial maturation rates: $Bias^{Pi}_{SM} = -0.08$ (Table 6). This is especially the case for observed pseudothecial maturation rates lower than 0.20 and higher than 0.60. $Efficiency^{Pi}_{SM}$ was low ($Efficiency^{Pi}_{SM} = 0.19$; Table 6), indicating a limited quality of prediction.



400

401 **Figure 3. Residues of SimMat-*Diaporthe* predictions for the entire dataset (179 observations) with: A, initial parameter values; B: final**
402 **parameter values with procedure 1 and C, final parameter values of procedure 2 (Table 7). □ Revel, 1989; ○ Revel, 1989; △ Cahuzac-**
403 **sur-Vère, 1990; – Quint-Fonsegrives, 1990; × Quint-Fonsegrives, 1991; ◇ Balma, 1992; ▽ Ondes, 1993; ⊠ Balma, 1994; ▲ Balma, 1995;**
404 **◆ Auzeville, 1997.**

MSE	**Value**	***Bias***	**Value**	***Efficiency***	**Value**
$MSE_{SM}^{P_i}$	0.054	$Bias_{SM}^{P_i}$	-0.083	$Efficiency_{SM}^{P_i}$	0.187
$MSE_{SM}^{P_{CV1}}$	0.036	$Bias_{SM}^{P_{CV1}}$	0.024	$Efficiency_{SM}^{P_{CV1}}$	0.269
MSE_{SM}^{A1}	0.021	$Bias_{SM}^{A1}$	0.008	$Efficiency_{SM}^{A1}$	0.683
MSE_{SM}^{A2}	0.082	$Bias_{SM}^{A2}$	-0.068	$Efficiency_{SM}^{A2}$	-0.222

Table 6. Statistical criteria for prediction of maturation dynamics. Values of *MSE*, *Bias* and *Efficiency* computed for the initial predictive quality of SimMat-*Diaporthe*, the predictive quality obtained by cross-validation in procedure 1, and the adjustment quality obtained using procedures 1 and 2.

3.2.2. Identification of the combination of SimMat-Diaporthe's parameters to adjust to minimize errors of prediction of pseudothecial maturation dynamics

The range of values tested, the initial and final values of parameters are presented Table 7. Procedure 1 selected three parameters to adjust to improve the quality of prediction: $\theta_{min}^{SM-D.}$, N_d and k. The parameters $\theta_{max}^{SM-D.}$, $R_{min}^{SM-D.}$, N_{FD} and σ_{FD} were not modified. Final values of $\theta_{min}^{SM-D.}$ and N_d were close to their initial values. $\theta_{min}^{SM-D.}$ slightly increased: 10.3 °C versus 10.0 °C before estimation. N_d and k decreased: 4 versus 7 days and 0.008 versus 0.046 mm^{-1} respectively (Table 7).

Parameter	Unit	Tested range	Initial value	Sources	Final value	
					Procedure 1	Procedure 2
$\theta_{min}^{SM-D.}$	°C	]0; 20[	10.0	Terekhov et al., 2004; Mathew, 2015	10.3	10.0
$\theta_{max}^{SM-D.}$	°C	]20; 40[	30.0	Gulya et al., 1997	30.0	30.0
$R_{min}^{SM-D.}$	mm	]0; 100[	4.0	Aubertot et al., 2006	4.0	11.9
N_d	day	]1; 30[	7	Expertise	4	5
N_{FD}	day	]1; 100[	30	Expertise	30	16.76
σ_{FD}	day	]1; 100[	15	Expertise	15.0	1.51
k	mm^{-1}	]0; 1[	0.046	Aubertot et al, 2006	0.008	0.084

Table 7. Results obtained with optimization procedures 1 and 2 (see sections 2.6 and 2.7) of the SimMat-*Diaporthe*'s parameters described in Table 1. Units, tested range of values, initial parameter values before estimation, sources for these initial values, parameter values after procedures 1 and 2 and the associated variations.

The modification of the three parameters during the cross validation of procedure 1 divided *MSE* by 1.5 ($MSE_{SM}^{P_{CV1}} = 0.036$ versus $MSE_{SM}^{P_i} = 0.054$, Table 6). The model did not over-estimate pseudothecial maturation rates anymore as described by the reduced Bias ($Bias_{SM}^{P_{CV1}} = 0.024$ versus $Bias_{SM}^{P_i} = -0.083$, Table 6). The Efficiency became positive ($\text{Efficiency}_{SM}^{P_{CV1}} = 0.27$ versus $\text{Efficiency}_{SM}^{P_i} = -0.187$, Table 6) and indicated a limited quality of prediction.

3.2.3. Optimization of SimMat-Diaporthe's parameters using the entire dataset

Figure 3B represents the residues of pseudothecial maturation rates predicted by SimMat-*Diaporthe*, after estimation of three parameters using the entire dataset at the end of the procedure 1. Comparisons between observed and simulated dynamics before and after optimization are presented Appendix A (central figures after optimization with procedure 1) for each situation. The observed pseudothecial maturation rates ≤ 0.20 and ≥ 0.60 were less over-estimated (Figure 3B and Appendix A), and the overall bias was improved ($Bias_{SM}^{A1} = 0.008$). The MSE_{SM}^{A1} was low ($MSE_{SM}^{A1} = 0.021$, Table 6) and the $\text{Efficiency}_{SM}^{A1}$ was high ($\text{Efficiency}_{SM}^{A1} = 0.68$, Table 6), indicating a good overall quality of adjustment.

3.2.4. Evaluation of the predictive quality of SimMat-Diaporthe on classes of pseudothecial maturation precocity after optimization using pseudothecial maturation dynamics

The confusion matrix built with observed pseudothecial maturation precocities and SimMat-*Diaporthe*'s predictions is presented Table 4C.

SimMat-*Diaporthe* obtained an excellent precision $P_{SM}^{A1} = 0.9$ (Table 5) and a high value for the squared weighted Kappa of Cohen $\kappa_{SM}^{A1} = 0.921$ (Table 5). More than 92% of the observed maturation classes' variability was explained by SimMat-*Diaporthe*, indicating a very good agreement between observations and predictions. Results for early dynamics were similar to those of ASPHODEL (Tables 4A and C). The dynamics

for Revel were now correctly predicted as late (Table 4C). After optimization, only the dynamic from Cahuzac-sur-Vère in 1990 was incorrectly classified (this intermediate precocity class was predicted as early).

3.3. Parameter estimation of SimMat-Diaporthe using qualitative data (procedure 2)

3.3.1. Evaluation of the initial predictive quality of SimMat-Diaporthe on classes of pseudothecial maturation precocity

The confusion matrix built with observed classes of maturation precocity and SimMat-*Diaporthe*'s predictions is presented Table 4C.

SimMat-*Diaporthe* had a good precision $P_{SM}^{P_i}$= 0.7 (Table 5) and a good squared weighted Kappa of Cohen $\kappa_{SM}^{P_i}$= 0.681 (Table 5), indicating a quite good agreement between observations and predictions. Results for the early dynamics were similar to those of ASPHODEL (Table 4A). The two late maturation classes from the "Revel" site were predicted as intermediate (Table 4B) and one intermediate maturation class was predicted as early.

3.3.2. Identification of the combination of SimMat-Diaporthe's parameters to adjust to minimize errors of prediction of classes of pseudothecial maturation precocity

Procedure 2 led to the same confusion matrix as procedure 1 (Table 4C). The range of values tested, the initial and final values of parameter are presented Table 7. The procedure 2 did not modify the two temperature thresholds and modified the other five parameters. The values of R_{min}^{SM-D} and *k* increased: 11.9 versus 4.0 mm and 0.084 versus 0.046 mm^{-1} before estimation, respectively. The values of N_d, N_{FD}, and σ_{FD} decreased: 5 versus 7 days; 16.756 versus 30 days; 1.51 versus 15 days, respectively.

3.3.3. Evaluation of the predictive quality of SimMat-Diaporthe on pseudothecial maturation dynamics after optimization using classes of pseudothecial maturation precocity

Figure 3C represents the residues of pseudothecial maturation rates predicted by SimMat-*Diaporthe*, after optimization with procedure 2. The predicted rates over-estimated observed pseudothecial maturation: $Bias_{SM}^{A2}$= -0.068 (Table 6). values. On the contrary, predicted rates under-estimated observed for pseudothecial maturation rates for 3 situations (Quint-Fonsegrives in 1990, Ondes in 1993 and Auzeville in 1997). The MSE_{SM}^{A2} was not too high as compared to other values(MSE_{SM}^{A2} = 0.082, Table 6), but the Efficiency$_{SM}^{A2}$ was negative (Efficiency$_{SM}^{A2}$ = -0.222, Table 6), indicating a poor quality of adjustment.

Comparisons between observed and simulated dynamics before and after optimization for each situation are presented in Appendix A (right figures after optimization with procedure 2).

4. Discussion

4.1. Outcomes of the evaluation and parameter optimization procedures

4.1.1. ASPHODEL

Despite a fair precision ($P_{ASPHODEL}^{A}$= 0.60, Table 5), ASPHODEL showed a poor adjustment quality ($\kappa_{ASPHODEL}^{A}$=0.057, Table 5). ASPHODEL predicted pseudothecial maturation earlier than what was observed. This result is somewhat surprising because ASPHODEL (Délos and Moinard, 1987) had been parameterized previously on a bigger dataset that included the dada presented in this paper. This result questions the robustness of ASPHODEL, but the size of the dataset used is not sufficient to ditch it. In addition, we identified that its formalism did not always rely on directly observable variables. Notably, it does not compute directly daily rates of pseudothecial maturation, but an associated index that cannot be observed. When this index reaches a given threshold, it is considered that pseudothecial maturation rate is 50%. This is why the quantitative dynamics of pseudothecial maturation could not be used with this model, and that the adjustment quality of ASPHODEL was assessed only with classes of pseudothecial maturation. In addition, the

ASPHODEL formalism is based on the assumption that pseudothecia refill after ascospore releases, although this phenomenon is not described in the scientific literature to our knowledge. It is the main difference with SimMat in terms of biological processes taken into account. Since ASPHODEL parameters were adjusted with a much bigger dataset (alas unavailable) than the one presented here, we did not implement a procedure to re-estimate these parameters.

4.1.2. SimMat-Diaporthe

Parameter optimization using Procedures 1 and 2 greatly improved the initial (quantitative and qualitative) qualities of prediction of SimMat-*Diaporthe* (Tables 5 and 6). The statistical criterion minimized in the cross-validation of procedure 1 was reduced ($MSE_{SM}^{P_{CV1}}$ = 0.036 versus $MSE_{SM}^{P_i}$ = 0.054), and the two other statistical criteria were slightly improved. However, these values indicated a limited quality of prediction. Procedure 2 also improved the model, with a satisfactory final quality of adjustment for qualitative predictions (Table 5), but a poor quality of adjustment for pseudothecial maturation dynamics (Table 6).

The procedure 1 modified three out of the seven parameters: $\theta_{min}^{SM-D.}$, N_d, and k. Generally, the optimized values were close or close enough to the initial values (Table 7). The procedure 2 modified five out of the seven parameters.

As expected, SimMat-*Diaporthe*'s predictive quality on pseudothecial maturation dynamic was higher using the set of parameters optimized with procedure 1. The two procedures led to the same quality of adjustment for qualitative classes of maturation precocity (Table) 5. We therefore recommend using SimMat-*Diaporthe* with the set of parameters optimized with procedure 1, since it will provide the most accurate prediction both for quantitative dynamics of pseudothecial maturation and qualitative classes of maturation precocity.

The choice of using SimMat was based on similarities between *D. helianthi* and *L. maculans* life cycles. Both pathogens survive on infected stubble and produce primary inoculum at soil surface when environmental

522 conditions are favourable. We chose the same value as for *L. maculans* for the initial value of rainfall threshold.
523 With procedure 2, the final value of $R_{min}^{SM-D.}$ was about three times higher than the initial value (i.e. 4.0 mm,
524 SimMat-*L. maculans*, Table 7). This initial value was appropriate for oilseed rape stubble, but it did not seem
525 high enough to moisten sufficiently stubble of sunflower in the model. Initial values for temperature thresholds
526 were taken from published studies. It is remarkable that these two parameters were not modified by procedure
527 2 and just very slightly modified by procedure 1. These values are also consistent with other studies: a
528 minimum values of 8°C was reported for pseudothecial maturation (Gulya et al., 1997) and Fabregue (1992)
529 reported that the maximum temperature threshold was higher than 20°C. The final values of N_d for *D. helianthi*
530 (4 days) remained close to the initial value (7 days, SimMat-*L. maculans*; Aubertot et al., 2006). In spite of
531 comparable life cycles, the fruiting bodies of these two fungi require specific environmental conditions to
532 develop. At last, final SimMat-*Diaporthe*'s temperature threshold values were close to those of ASPHODEL.

533

534 *4.2. Methodological considerations*

535 The methods used to estimate parameters described in this paper are innovative in two ways. First, the
536 algorithmic procedures considered both quantitative and qualitative predictions. Usually, optimization is
537 performed using quantitative data for quantitative models and using qualitative data for qualitative models.
538 Second, the implemented algorithmic procedures took into account all possible combination of paramaters to
539 adjust and did not proceed to a selection of parameters beforehand. Indeed, it is usually computationally
540 impossible to test each parameter combination to adjust. Generally, a sensitivity analysis is conducted before
541 parameter estimation to identify the most influential parameters (Wallach et al., 2006). The low number of
542 parameters (7) in SimMat permitted to design protocols based on a method proposed by (Wallach et al., 2001)
543 to test all the $2^7 = 128$ parameter combinations. It is important to highlight that the method presented in this
544 paper is generic and can be applied to any simple model with a limited number of parameters (i.e. lower than
545 10) with a reasonable size for the dataset used (i.e. lower than 100 situations). This limitation is not theoretical
546 and is provided for typical current computational power.

547 The R function *optim* from the {Stat} library, was used in the procedure 1 to minimize MSEs and in the
548 procedure 2 to maximize quadratic weighted Cohen's kappa parameters. It searched parameter values leading
549 to minimum or maximum values for statistical criteria (global optima). It is a common problem of *optim*
550 (Wallach et al., 2006) to get stucked in a local optimum. In order to limit this risk, a protection was
551 implemented: different initial seed values comprised between bounds were used in the algorithm (5 for
552 parameter estimation using the entire dataset and 3 for cross-validations).

553 The dataset used in this paper was not originally intended to help develop a simulation model. These data
554 were indeed collected to help writing warning bulletins for farmers in the French Occitanie region. These
555 observations require specific skills and are time consuming. The dataset used for parameter estimation was
556 therefore quite limited (179 observations over 10 sites-years). This prevented us from estimating the quality of
557 prediction of maturation precocity classes by SimMat-*Diaporthe* using a cross-validation approach, since only
558 few situations remained after taking into account potential spatio-temporal correlations. In addition, the dataset
559 used was unbalanced: 35.8% of observed pseudothecial maturation rates (64 in 179) were comprised between
560 of 0 and 0.1 (all dates confounded). A dataset with observations with values more evenly distributed between
561 0 and 1 would have provided a more useful information. However, pseudothecial maturation rates could not
562 be controled for the field observations reported in this study. In spite of its small size, its unbalanced nature,
563 and the limited number of provenances of infected stubble, the dataset used presented a range of pseudothecial
564 maturation dynamics: 5 situations corresponded to early maturations, 3 had intermediate, and 2 had late
565 maturation dynamics. In addition, the dataset used was collected in the 90's in the place of origin of phomopsis
566 stem canker in France. New species of *Diaporthe sp.* were observed in Australia (Thompson et al., 2011) and
567 in the US (Mathew et al., 2018). However, no identification of new *Diaporthe sp.* species was reported in
568 France. It would therefore be valuable to test the model with new data, in order to make sure that the main
569 characteristics of the pathogen's lifecycle have not evolved over time. This would require to implement a
570 network of pseudothecial maturation observations.

571

4.3. *Impact and perspectives*

The present study permitted to develop a new model predicting primary inoculum production of *Diaporthe helianthi*. Like pioneer models, it relies on the concept of days favourable to the development of the considered microorganism (Fabregue, 1991; Moinard and Eychenne, 1998). Also, it compares to many models that describe the effect of temperature integrated over time on fruiting body maturation of other fungi and represent the effect of rainfall on spore release (Nugent, 1950; Hyre, 1954; Gadoury and MacHardy, 1982; Rossi et al., 2000; Salam et al, 2007; Legler et al., 2012; Roubal and Nicot, 2016; Singh et al., 2016; Gonzalez-Dominguez et al., 2017). It was designed to be as parsimonious as possible in terms of number of parameters, and to have input variables as easy to provide as possible. As compared to the historical model, ASPHODEL, SimMat-*Diaporthe* does not require hourly weather variables and is therefore easier to use. It also performed better than ASPHODEL when predicting classes of precocity for pseudothecial maturation.

SimMat-*Diaporthe* is now available to help writing warning bulletins for France. Whenever possible, the authors recommend to use both the historical model and SimMat-*Diaporthe* in order to permit a transition as smooth as possible. Additional data would be helpful to improve the quality of prediction of both models. Optimization algorithms are now available in R to improve models as soon as additional data are available. These generic algorithms are also available to improve any similar model.

5. Acknowledgements

We thank Benoît Persyn from the Agroclim unit (INRAE, Avignon) and Météo France for providing the weather dataset used in this study. We are also thankful to Philippe Debaeke (INRAE, Toulouse), Luc Champolivier (Terres Inovia, Toulouse), Tom Gulya (United State Departement of Agriculture, retired, California), David Camilo Corrales (INRAE, Toulouse) for their useful advices. This work was supported by Terres Inovia and the Occitanie region through the doctoral research fellowship of the first author. We thank

595 INP-Toulouse for the travelling fellowship granted to Martin Barbetti to visit the other authors as part of the
596 program “2019 Visiting Professor Toulouse INP”.

597

598 **References**

599 Agresti A., Kateri M. 2010. Analysis of ordinal categorical data. Matched-pairs data with ordered categories.
600 John Wiley & Sons, New Jersey, 2nd edition.

601 Aubertot J.N., Salam M., Diggle A., Dakowska S., Jedryczka M. 2006. SimMat, a new dynamic module of
602 Blackleg Sporacle for the prediction of pseudothecial maturation of *L. maculans*/*L. biglobosa* species
603 complex. Parameterisation and evaluation under Polish conditions. Integrated Control in Oilseed Crops
604 IOBC/wprs Bulletin 29 (7): 277–285.

605 Box M. 1965. A new method of constrained optimization and a comparison with other methods. Computer
606 Journal 8 (1): 42–52.

607 Byrd R. H., Lu P., Nocedal J., Zhu C. 1995. A limited memory algorithm for bound constrained optimization.
608 Technical Report. Northwestern University Department of Electrical Engineering and Computer Science.

609 Chambre d’Agriculture Occitanie. 2018. Bulletin de Santé du Végétal Grandes cultures. Edition Ouest-
610 Occitanie. URL https//occitanie.chambre-agriculture.fr/fileadmin/userupload/Occitanie/512 Fichiers-
611 communs/documents/BSV/BSV ex-MP/Grandes cultures/bsvgrandes cultures N29 21062018.pdf.
612 Accessed 11.10.2019.

613 Debaeke P, Estragnat A., Reau R. 2003. Influence of crop management on sunflower stem canker (*Diaporthe*
614 *helianthi)*. Agronomie 23 (7): 581–592.

615 Debaeke P, Moinard J. 2010. Effect of crop management on epidemics of phomopsis stem canker (*Diaporthe*
616 *helianthi*) for susceptible and tolerant sunflower cultivars. Field Crops Research 115 (1): 50–60.

617 Délos M., Moinard J., Jacquin D. 1995. Etude et surveillance du phomopsis : des pièges au modèle, Phytoma
618 - La Défense des Végétaux 473: 25–27.

619 Délos M., Moinard J. 1997. ASPHODEL : Modèle de simulation des épidémies de phomopsis du tournesol
620 (*Diaporthe helianthi)*, in: ANPP (Ed.), Cinquième conférence internationale sur les maladies des plantes,
621 Tours, France.

622 Desanlis M, Aubertot J.N., Mestries E., Debaeke P. 2013. Analysis of the influence of a sunflower canopy on
623 *Phomopsis helianthi* epidemics as a function of cropping practices. Field Crops Research 149: 63–75.

624 Duroueix F., Penaud A. 2013. Les fongicides en renfort contre phoma et phomopsis, Perspectives Agricoles
625 400: 18–21.

626 Estragnat A. 1992. Phomopsis du tournesol. Oléoscope 8: 8–10.

627 Fabregue C. 1991. Phomopsis du tournesol. Rapport Général. URL
628 https://archive.org/stream/RN_GC_Phomopsis_du_tournesol_1991/Phomopsis%20du%20tournesol%201
629 991_djvu.txt. Accessed 11.10.2019.

630 Fabregue C. 1992. Phomopsis du tournesol. Oléoscope 8: 11.

631 Fayret J., Assemat P. 1987. Evolution du *Diaporthe helianthi* (*Phomopsis helianthi*) Munt-Cvet et al. et
632 différenciation des organes reproducteurs sur les plants de tournesol après la période de végétation,
633 *Informations Techniques CETIOM* 98: 2–11.

634 Fleiss J., Cohen J. 1973. The equivalence of weighted Kappa and the intraclass correlation coefficient as
635 measures of reliability. Educational and psychological measurement 33: 613–619.

636 Gadoury D. M., MacHardy W. E. 1982. A model to estimate the maturity of ascospores of *Venturia inaequalis*.
637 Phytopathology 72 (7): 901-904.

638 Gonzalez-Dominguez E., Armengol J., Rossi V. 2017. Biology and epidemiology of *Venturia* species affecting
639 fruit crops: a review. Frontiers in Plant Science 8. 1496.

640 Gulya T., Rashid K. Y., Masirevic S. M. 1997. Sunflower diseases, in: Sunflower Technology and Production
641 Agronomy Monograph n◦35. ASA, CSSA, SSSA, Madison, WI (USA): 263–379.

642 Hyre. R. A. 1954. Progress in forecasting late blight of potato and tomato. Plant Disease Reporter Supplement
643 38: 245–253.

644 Jinga V, Iliescu H., Rafaila C., Ciurea A., Ionita A., Csép N. 1987. Cercetari privind prognoza aparitiei atacului
645 de patare bruna si fringere a tulpinilor (*Diaporthe-Phomopsis*) de floarea-soarelui, Analele Institutul de
646 Cercetari pentru Protectia Plantelor 20: 57–68.

647 Jinga V., Iliescu H., Ionita A., Csép N., Iordache, E. 1992. Technological factors in the integrated control of
648 sunflower diseases, in: ISA (Ed.), Proceedings of 13th International Sunflower Conference, Pisa, Italy. 761–
649 766.

650 Johnson S.G. 2009. The NLopt nonlinear-optimization package. http://github.com/stevengj/nlopt

651 Lamarque C., Perny R. 1985. Nouvelle maladie du tournesol : le phomopsis, Cultivar 179: 57–58.

652 Landis J. R., Koch G. G. 1977. The measurement of observer agreement for categorical data. Biometrics 33:
653 159-174.

654 Legler S. E., Caffi T., Rossi V. 2012. A nonlinear model for temperature-dependent development of *Erysiphe*
655 *necator chasmothecia* on grapevine leaves. Plant Pathology 61 (1): 96–105.

656 Lô-Pelzer E., Bousset L., Jeuffroy M., Salam M.U., Pinochet X., Boillot M., Aubertot J.N. 2010. SIPPOM-
657 WOSR: A Simulator for Integrated Pathogen Population Management of phoma stem canker on Winter
658 OilSeed Rape: I. Description of the model. Field Crops Research 118 (1): 73–81.

659 Maric A., Masirevic S. M. 1980. Pojava sive pegavostistabla (*Phomopsis* sp.) do sada nepoznale bolesti
660 suncokreta, Zastita Bilja 12: 421–23.

661 Masirevic S. M., Gulya T. 1992. Sclerotinia and Phomopsis two devastating sunflower pathogens. Field Crops
662 Research 30 (3-4): 271–300.

663 Masirevic S. M. 2016. Sunflower diseases research progress and management, in: ISA (Ed.), Proceedings of
664 the 19th International Sunflower Conference. Edirne, Turkey. 61–70.

665 Mathew F. M., Alananbeh K. M., Jordahl J. G., Meyer S. M., Castlebury L. A., Gulya T., Markell S. G. 2015.
666 Phomopsis stem canker: a reemerging threat to sunflower (*Helianthus annuus*) in the United States,
667 Phytopathology 105 (7): 990–997.

668 Mathew F., Harveson R., Gulya T., Thompson S., Block C., Markell S. 2018. Phomopsis stem canker of
669 sunflower. Plant Health Instructor.

670 https://www.apsnet.org/edcenter/disandpath/fungalasco/pdlessons/Pages/PhomopsisStemCanker.aspx.
671 Accessed 11.10.2019.

672 Mihaljcevic M., Muntanola-Cvetkovic M., Vukojevic J., Petrov M. 1985. Source of infection of sunflower
673 plants by *Diaporthe helianthi* in Yougoslavia. Phytopathologische Zeitschrift 113: 334–342.

674 Moinard J., Eychenne N. 1998. Le Phomopsis. Guide Grandes Cultures. Dossier support des Avertissements
675 Agricoles routés et telecopiés 12.

676 Muntanola-Cvetkovic M, Vukojevic M, Mihaljcevic M. 1988. Pathohistology of sunflower stems attacked by
677 *Diaporthe helianthi*. Canadian Journal of Botany 67: 1119–1125.

678 Nelder J., Mead R. 1965. A simplex method for function minimization. Computional Journal 7: 308–313.

679 Nugent T. J. 1950. Three years' experience forecasting late blight in tidewater Virginia. Plant Disease Reporter
680 Supplement 190: 9–13.

681 Perny R. A. 1992. L'évolution du phomopsis en France, Oléoscope 8: 7.

682 Petrov M., Muntanola-Cvetkovic M, Mihaljcevic, M. 1981. New findings on the disease of sunflower plants
683 caused by *Phomopsis helianthi* Munt.-Cvet. et al. (*Diaporthe helianthi* Munt.-Cvet et al.). Archives des
684 Sciences Biologiques (Beograd) 33: 13–19.

685 Rossi V., Ponti I., Marinelli M., Giosuè S., Bugiani R. 2000. A new model estimating the seasonal pattern of
686 air-borne ascospores of *Venturia inaequalis* (Cooke) Wint. in relation to weather conditions, Journal of
687 Plant Pathology 82 (2): 111–118.

688 Roubal C., Nicot P. C. 2016. Apple scab: numerical optimization of a new thermal time scale and application
689 for modelling ascospore release in southern France. Plant Pathology 65 (1): 79–91.

690 Salam M., Khangura R., Diggle A., Barbetti M. 2003. Blackleg Sporacle: a model for predicting onset of
691 pseudothecia maturity and seasonal ascospore showers in relation to blackleg of canola, Phytopathology 93
692 (9): 1073–1081.

693 Salam M. U., Fitt B. D. L., Aubertot J. N., Diggle A. J., Huang Y. J., Barbetti M. J., Gladders P., Jedryczka M,
694 Khangura R. K., Wratten N., Fernando W. G. D., Penaud A., Pinochet X., Sivasithamparam K. 2007. Two
695 weather-based models for predicting the onset of seasonal release of ascospores of *Leptosphaeria maculans*
696 or *L. biglobosa*, *Plant Pathology* 56 (3): 412–423.

697 Singh K., Kumar J., Singh A., Prasad R., Singh R., Prasad D. 2016. Maturation, ascospores discharge pattern
698 and relevance of mills criteria for predicting apple scab infection period in India. Plant Pathology Journal
699 15 (3): 108–123.

700 Slyusa E., Antonova L., Bochkarev N. 1988. On the spread of Phomopsis infection on sunflower, Zashchita I
701 Karantin Rastenii 1: 33–35.

702 Terekhov V. I., Begunov I. I., Begunova A. G. 2004. Role of aerial infection in sunflower phomopsis
703 (*Phomopsis* [*Diaporthe*] *Helianthi* epiphytology, in: Proc. 16th International Sunflower Conference, Fargo,
704 ND (USA): 124–128.

705 Terres Inovia. 2019. Guide de culture tournesol. URL:
706 https://www.terresinovia.fr/documents/20126/453413/Guide tournesol Terres-Inovia2019.pdf/bfcbf68a-
707 091a-dade-97a1-2b8ee4222f62?t=1553962903039. Accessed 11.10.2019.

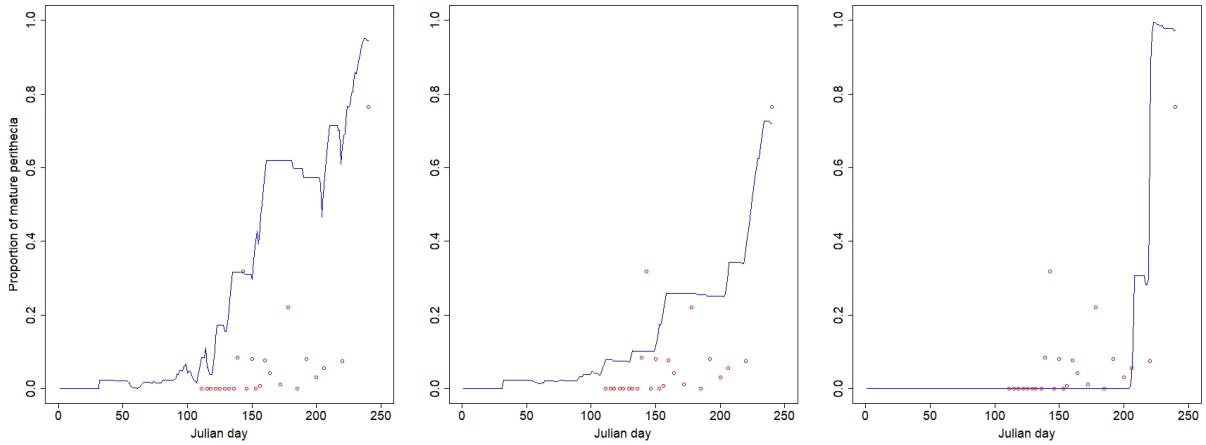
Thompson S., Tan Y., Young A., Neate S., Aitken E., Shivas R. 2011. Stem cankers on sunflower (*Helianthus annuus*) in Australia reveal a complex of pathogenic *Diaporthe* (*Phomopsis*) species. Peersonia 27: 80–89.

Wallach D., Goffinet B., Bergez J. E., Debaeke P., Leenhardt D., Aubertot J. N. 2001. Parameter estimation for crop models: a new approach and application to a corn model. Agronomy Journal 93:757–766.

Wallach D., Makowski D., Jones J. 2006. Working with dynamic crop models: evaluation, analysis, parameterization, and applications. Elsevier, 2nd edition.

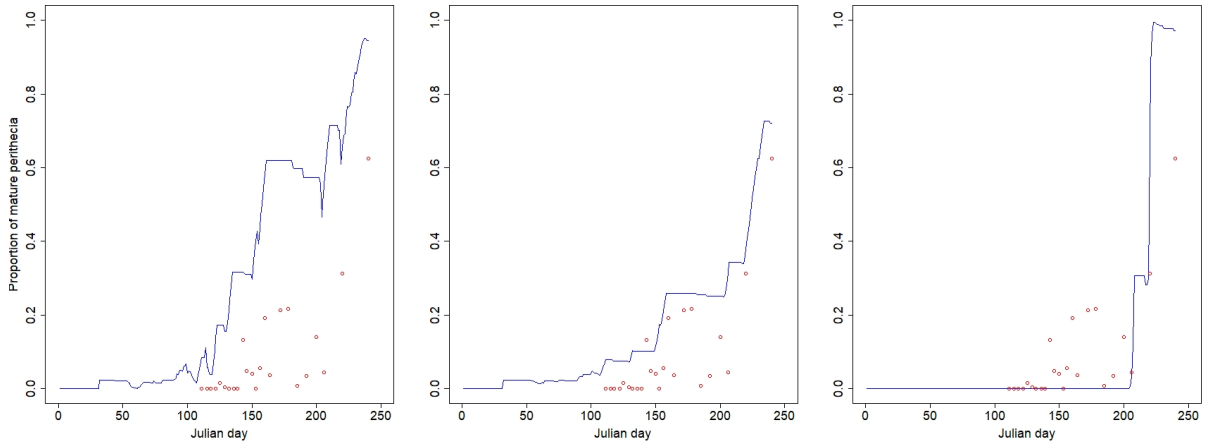
Witten I. H., Frank E., Hall M. A. 2011. Data mining practical machine learning tools and techniques. Elsevier, third edition.

730 **Appendix A. Comparisons between dynamics observed and simulated before and after both optimization**
731 **procedures for each situation**



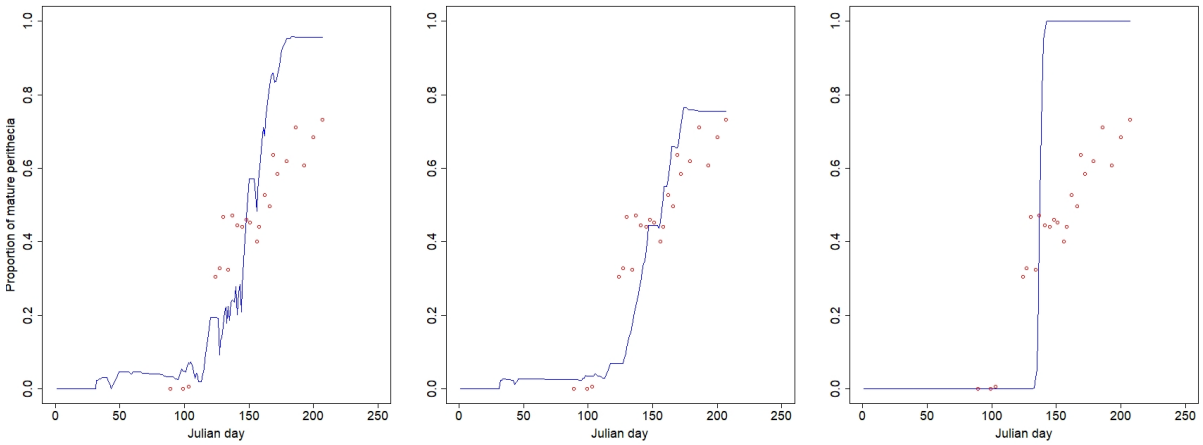
732

733 **Figure A.1. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
734 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
735 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
736 **2 (Table 7). Revel (1989).**



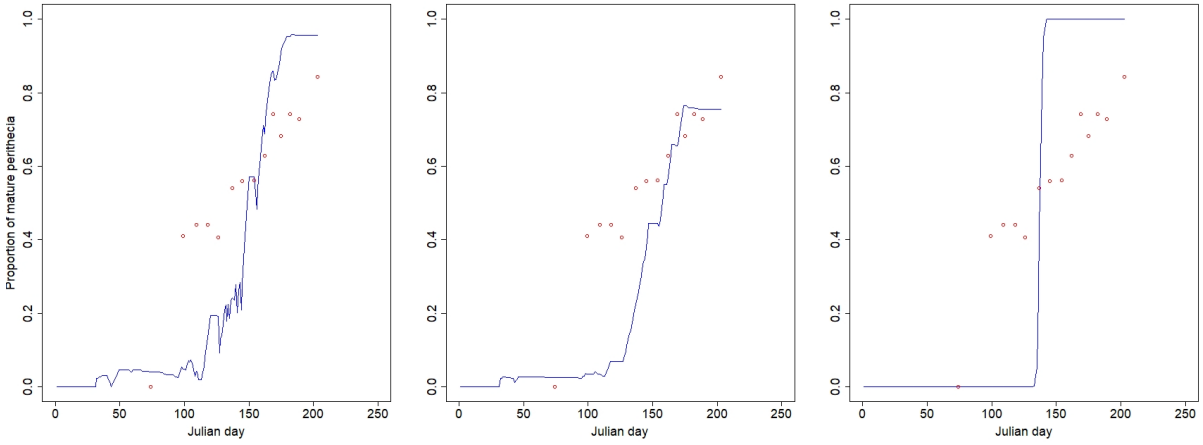
737

738 **Figure A.2. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
739 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
740 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
741 **2 (Table 7). Revel (1989).**



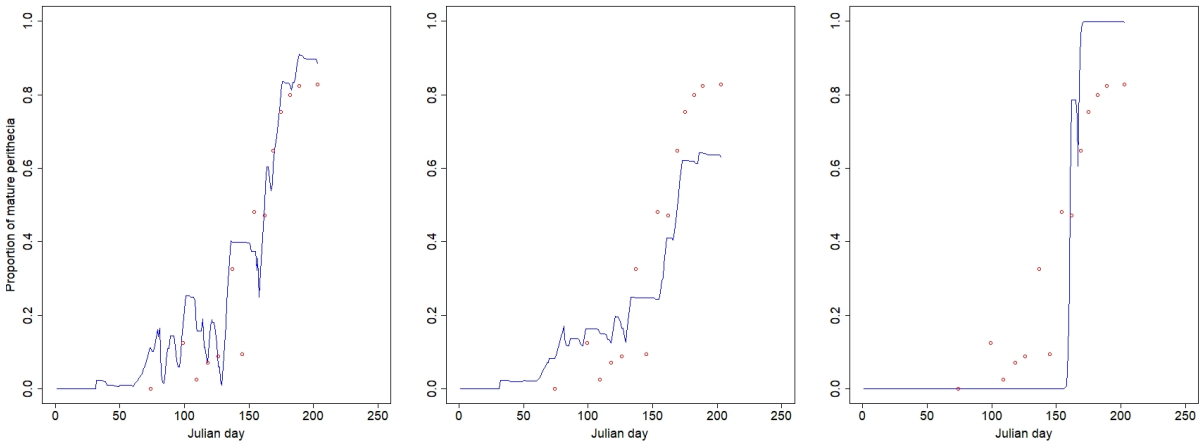
742

743 **Figure A.3. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
744 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
745 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
746 **2 (Table 7). Cahuzac-sur-Vère (1990).**



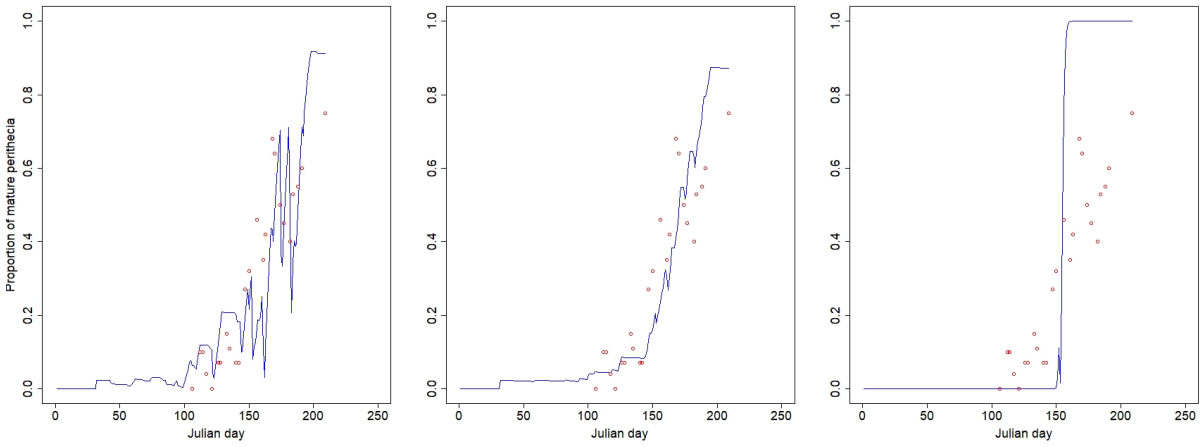
747

748 **Figure A.4. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
749 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
750 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
751 **2 (Table 7). Quint-Fonsegrives (1990).**



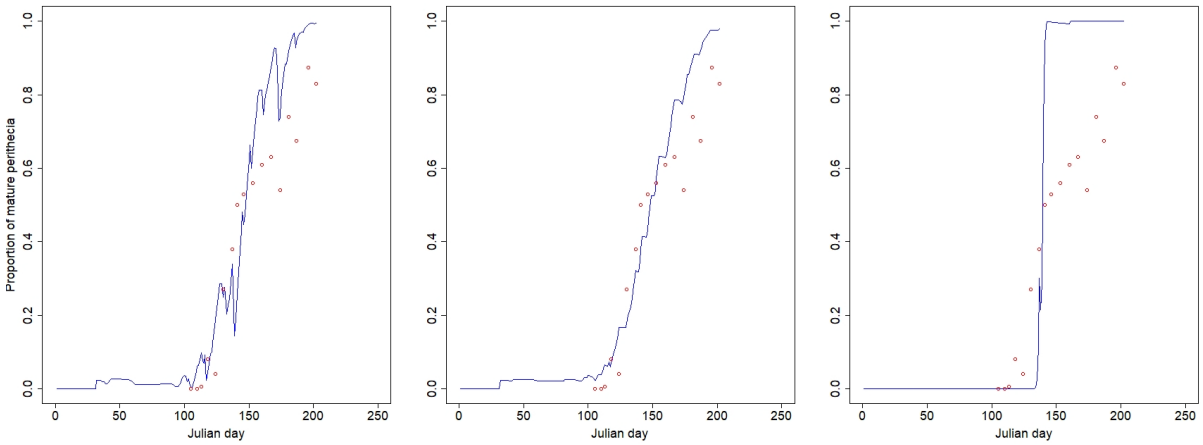
752

753 **Figure A.5. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
754 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
755 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
756 **2 (Table 7). Quint-Fonsegrives (1991).**



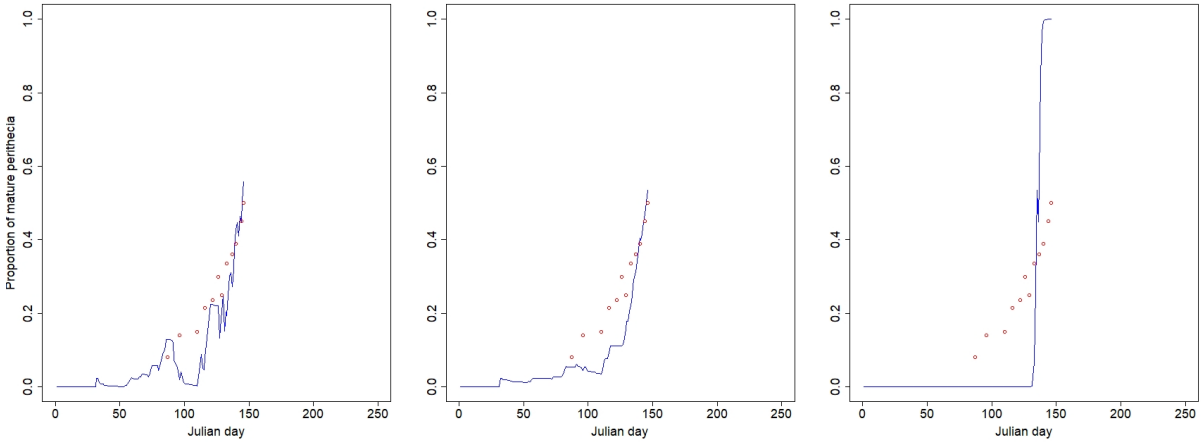
757

758 **Figure A.6. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
759 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
760 **parameter optimization using procedure 1 (Table 7). F: Balma (1992).**



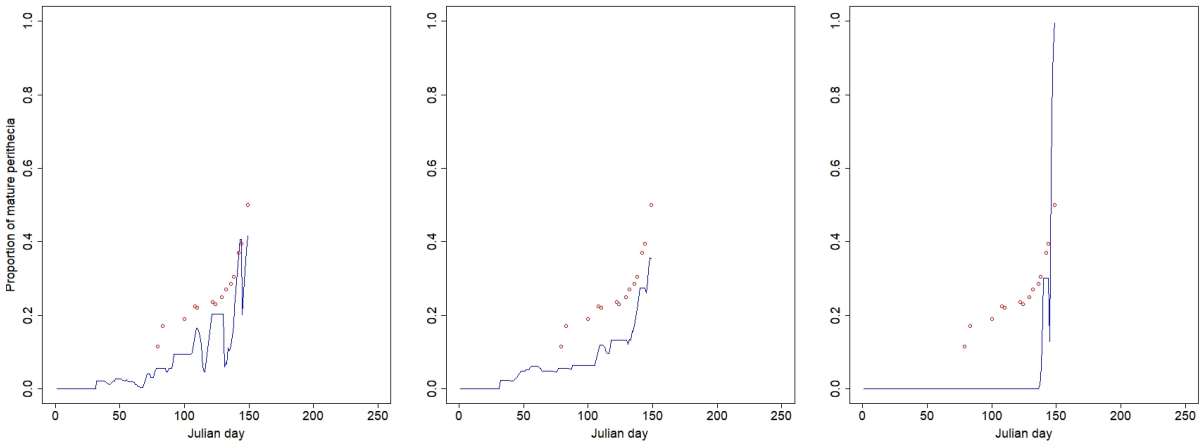
761

762 **Figure A.7. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
763 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
764 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
765 **2 (Table 7). Ondes (1993).**



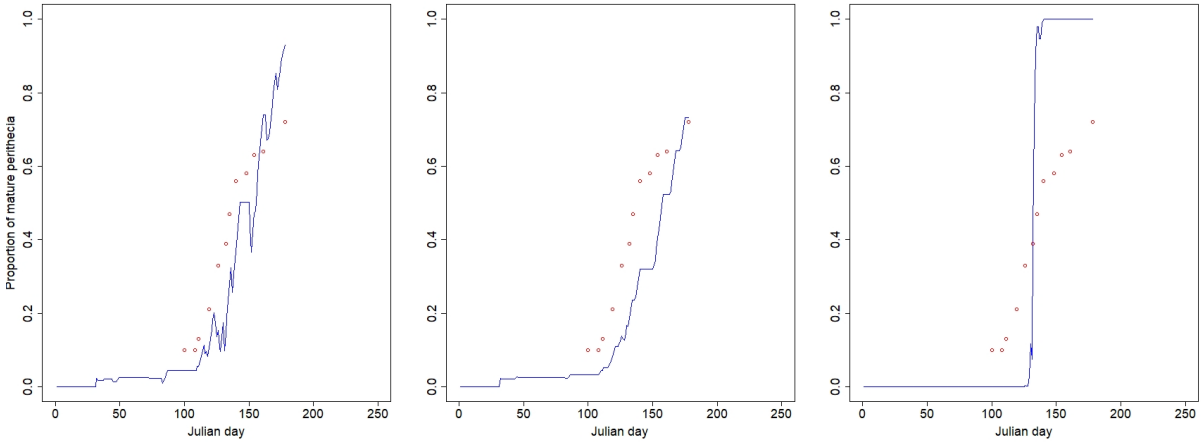
766

767 **Figure A.8. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
768 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
769 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
770 **2 (Table 7). H: Balma (1994).**



771

772 **Figure A.9. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
773 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
774 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
775 **2 (Table 7). I: Balma (1995).**



776

777 **Figure A.10. Comparison between observed and predicted pseudothecial maturation dynamics. Blue lines are**
778 **predictions, red points are observations. Left figures: before optimization (Table 7). Central figures: after**
779 **parameter optimization using procedure 1 (Table 7). Right figures: after parameter optimization using procedure**
780 **2 (Table 7). Auzeville (1997).**

781 Before estimation, the prediction of pseudothecial maturation dynamics of all the situations began earlier
782 than observations. SimMat-*Diaporthe* was unable to predict correctly the pseudothecial maturation dynamics
783 for Revel (left Figure A.1. and A.2). These predictions were the most over-estimated of all the situations.

784 After estimation of parameters with procedure 1, the prediction of pseudothecial maturation rates of all the
785 situations began earlier than with initial parameter values. The predictions were globally much less
786 overestimated, especially for Revel. The observation lower than 0.60 were under-estimated for Quint-
787 Fonsegrives in 1990 (central Figure A.4.). They were over-estimated for observation higher than 0.70 for
788 Quint-Fonsegrives in 1991 and Auzeville in 1997 (central Figure A.5 and A.10). The observations lower than
789 0.4 were over-estimated for Balma in 1995 (central Figure A.9).

790 After estimation of parameters with procedure 2, the predictions of maturation dynamics were not as good as
791 those obtained after estimation of parameters with procedure 1, especially for Revel (right Figure A.1. and
792 A.2). It is logical since procedure 1 aimed at improving the quantitative prediction of pseudothecial maturation
793 dynamics, whereas procedure 2 aimed at improving the quality of prediction of precocity classes.

2. Application à SimMat et *Leptosphaeria lindquistii*

2.1 Introduction

Leptosphaeria lindquistii est responsable du phénomène de pieds secs du tournesol. Son cycle a été présenté dans la partie introductive de cette thèse. Il provoque, en général trois semaines à un mois avant la maturité physiologique normale, le noircissement du collet. Ce noircissement peut s'accompagner d'un dessèchement précoce des pieds : les plantes se dessèchent prématurément, et les capitules restent de très petite taille, secs, noirs et recroquevillés. Les attaques débutent par une petite tache au collet qui se développe puis l'encercle. Le champignon pénètre peu à peu à l'intérieur des tissus de la tige, jusqu'aux vaisseaux conducteurs. On assiste alors à un arrêt progressif du fonctionnement physiologique de la plante (transpiration, photosynthèse), qui entraîne la dégradation du système racinaire. L'apparition du dessèchement précoce se manifeste environ 45 jours après le début de l'attaque au collet et stoppe le remplissage des graines. Il existe des traitements fongicides efficaces contre les taches sur tiges, mais ils ne sont pas efficaces pour maîtriser les attaques au collet. Les effets d'une modification de la date de semis ne permettent pas de mettre en œuvre des stratégies d'esquive efficaces (Mestries *et al.*, 2011). Le premier enjeu de modéliser la production d'inoculum primaire de *L. lindquistii* est donc de mieux raisonner les traitements fongicides contre les taches sur tiges, dans les situations pour lesquelles les méthodes prophylactiques n'ont pas été suffisamment efficaces. Le second enjeu est la production de connaissances sur le cycle biologique de l'agent pathogène.

2.2 Matériel et méthodes

2.2.1 Présentation du modèle SimMat

Le modèle épidémiologique SimMat-Leptosphaeria maculans est un modèle dynamique de prédiction de la maturation des pseudothèces et de l'émission d'ascosporess responsable de la nécrose au collet du colza (phoma du colza) en fonction du climat (Aubertot *et al.*, 2006). Ce modèle repose sur l'hypothèse selon laquelle les fructifications du champignon considéré ont besoin de cumuler un certain nombre de jours favorables à la maturation pour atteindre le stade de maturité. Une fois ce stade atteint, les pseudothèces vont émettre leurs ascosporess à la faveur d'un événement pluvieux. Le formalisme mécaniste générique de SimMat ne comporte que sept paramètres (Tableau III.1) qui permettent de définir ce qu'est un jour favorable à la maturation des pseudothèces (la température de ce jour doit être comprise entre une température minimale et une température maximale, et les précipitations cumulées sur un nombre déterminé de jours avant le jour considéré doivent être supérieures à un seuil minimal). Il fonctionne uniquement avec deux variables d'entrée météorologiques journalières : température et pluviométrie).

Paramètre	**Signification**	**Valeurs des paramètres et unités**
θ_{min}^{SM}	seuil de température minimale journalière	6,0°C
θ_{max}^{SM}	seuil de température maximale journalière	29,0°C
R_{min}^{SM}	seuil de précipitation minimale cumulée	11,6 mm
N_d	Nombre de jours favorables qui doivent précéder le jour considéré	8 jours
N_{FD}	Nombre moyen de jours favorables nécessaires pour une maturation	22,9 jours
σ_{FD}	écart-type de jours favorables nécessaire pour une maturation	15,0 jours
k	Proportion d'émission par mm de précipitation	0.046 mm^{-1}

Tableau III.1 – Paramètres, définitions et unités des paramètres de SimMat.

Dans le modèle, une variable binaire « jour favorable » représente les conditions de maturation. Elle prend la valeur 1 si le jour est favorable, 0 sinon. Une autre variable est construite pour cumuler les jours favorables. Puis, la proportion de pseudothèces matures n'ayant pas émis d'ascosporess P_{MP} (équation 2.1) jusqu'au jour considéré est calculée. Cette proportion est décrite par une loi normale (Aubertot *et al.*, 2006) d'espérance N_{FD} et d'écart-type σ_{FD}. Le taux de pseudothèces vides P_{PV} au jour d (équation 2.2) est calculé, puis le taux de maturation sur les pseudothèces non

vides P_{MPNV} est déduit (équation 2.3). Cette variable est la principale variable de sortie du modèle.

$$P_{MP}(d) \simeq \frac{1}{\sqrt{2\pi}\sigma_{FD}} \int_0^{CDFM(d)} e^{\frac{-(x-N_{FD})^2}{2\sigma_{FD}^2}} dx \tag{2.1}$$

$$P_{PV}(d) = P_{PV}(d-1) + (P_{MP}(d) - P_{PV}(d-1))(1 - e^{-k*r(d)}) \tag{2.2}$$

$$P_{MPNV}(d) = \frac{P_{MP}(d) - P_{PV}(d)}{1 - P_{PV}(d)} \tag{2.3}$$

Les pseudothèces de *L. lindquistii* commencent à mûrir dès le mois de février et sont capables d'émettre des ascosporess tout le long du cycle du tournesol (Délos *et al.*, 1998 ; Descorps, 2010). La date de démarrage des simulations est donc la même que pour *D. helianthi* : le 1er février. L'optimisation de la qualité de prédiction n'a été réalisée que sur les dynamiques de taux de maturation des pseudothèces.

2.2.2 Valeurs initiales des paramètres

Le phomopsis et le phoma présentent des similarités biologiques (survie du mycélium et fructification sur les cannes infectées, période de maturation des fructifications et d'éjections). Par conséquent, certaines valeurs de paramètres obtenues lors de l'adaptation de SimMat au phomopsis ont servi de base pour l'adaptation de SimMat au phoma. Les valeurs initiales des paramètres de température maximale ($\theta_{max}^{SM-L.\ lindquistii}$), de pluviométrie ($R_{min}^{SM-L.\ lindquistii}$), nombre de jour $N_d^{SM-L.\ lindquistii}$, l'esperance ($N_{FD}^{SM-L.\ lindquistii}$) et la variance ($\sigma_{FD}^{SM-L.\ lindquistii}$) sont issues des valeurs de paramètres obtenues après une première optimisation de SimMat-*Diaporthe*. Les valeurs initiales sont : $\theta_{max}^{SM-L.\ lindquistii}$ = 29,0°C ; $R_{min}^{SM-L.\ lindquistii}$ = 11,6 mm ; $N_d^{SM-L.\ lindquistii}$ = 8 jours ; $N_{FD}^{SM-L.\ lindquistii}$ = 22,9 jours ; $\sigma_{FD}^{SM-L.\ lindquistii}$ = 15,0 jours. Les autres valeurs proviennent de la littérature ou du modèle SimMat - *L. lindquistii* : $\theta_{min}^{SM-L.\ lindquistii}$ = 6°C (Muhlberger, 1996) et k = 0,046 mm^{-1} (Aubertot *et al.*, 2006).

2.2.3 Jeux de données pour la calibration de SimMat-*L. lindquistii*

Le tableau III.3 est un récapitulatif des données ayant servi à la calibration de SimMat-*L. lindquistii*. Ce jeu de données provient d'essais de Terres Inovia, d'essais issus du Projet ANR

MicMac de l'INRAE et de relevés relevés par le **S**ervice **R**égional de la **P**rotection des **V**égétaux (SRPV) de Midi-Pyrénées. Une partie de ce jeu de données a été conçue pour étudier l'impact de la pluviométrie et de la température sur la maturation des pseudothèces et l'émission d'ascospores. La maturation de pseudothèces a été notée dans 40 champs dans l'Ariège (09), l'Aude (11), le Cher (18), la Haute-Garonne (31), le Gers (32), la Gironde (33), l'Ille-et-Vilaine (35) et le Lot-et-Garonne (37), entre 1996 et 2016. Ce jeu de données est composé de 189 observations. Les taux entre 0,0 et 0,1 et entre 0,9 et 1,0 sont les plus observés (respectivement 25 et 56 observations sur 189, figure III.1).

Région	Dép.	Provenance géographique des résidus infectés	Année	Nombre d'observations	Période d'observation des pseudothèces	Taux de maturation de pseudothèces observés (valeur minimum et maximum)
Bourgogne-Franche-Comté	34	Dijon	1998	12	12/02/2011 - 30/07/2011	0,0 - 0,65
Bretagne	35	Nancy	1998	11	12/02/1998 - 01/07/1998	0,0 - 0,80
Centre-Val de Loire	18	Saint-Florent	1998	12	12/02/1998 - 30/07/1998	0,27 - 1,0 ; 0,0 - 0,95
			2011	4	17/06/2011 - 21/07/2011	0,16 - 1,0 ; 0,52 - 1,0
Nouvelle-Aquitaine	33	Saint-Pathus	1998	12	12/02/2011 - 30/07/2011	0,0 - 0,8
	47	Agen	2011	5	16/05/2011 - 12/07/2011	0,76 – 1,0
Occitanie	09	Troyes-d'Ariege	2011	1	15/06/2011	1,0 ; 1,0
	11	Alairac	2011	1	15/06/2011	1,0
		Cintegabelle	2011	1	15/06/2011	0,84
		Fanjeaux	2011	1	15/06/2011	0,72 ; 1,0
	31	Albiac	2011	1	16/06/2011	0,84
		Auzeville	1996	18	30/04/1996 - 30/06/1996	0,5 - 1,0
			1998	12	26/03/1998 - 30/06/1998	0,01 - 0,71
			1999	6	07/04/1999 - 12/05/1999	0,04 - 0,78
			2000	6	14/04/2000 - 04/05/2000	0,064 - 0,508

		2011	3	23/05/2011 - 07/06/2011	0,92 - 1,0
		2011	1	20/06/2011	1,0 ; 1,0 ; 1,0
	Balma	1996	10	30/04/1996 - 18/06/1996	0,14 - 0,844
	En-Crambade	1998	12	12/02/2011 - 30/07/2011	0,0 - 0,93
		2016	7	20/01/2016 - 01/04/2016	0,16 - 0,6
	Lauragais	2011	1	15/06/2011	0,92
		2011	1	16/06/2011	1,0 ; 1,0 ; 0,84 ; 1,0 ; 0,64 ; 0,96
	Mauremont	2011	7	06/04/2011 - 29/06/2011	0,76 - 1,0
	Montastruc-la-Conseillère	2011	1	16/06/2011	0,92
	Venerque	2011	7	30/03/2011 - 22/06/2011	0,28 - 0,80
		2011	1	15/06/2011	0,92
32	Ligardes	2011	5	16/05/2011 - 12/07/2011	0,88 - 1,0
	Pergain-Taillac	2011	1	16/06/2011	1.0

Tableau III.3 – Présentation du jeu de données (189 observations) utilisé pour calibrer le modèle SimMat-*L. lindquistii*. Provenance géographique des résidus infectés (09 : Ariège, 11 : Aude, 18 : Cher, 31 : Haute-Garonne, 32 : Gers, 33 : Gironde, 35 : Ille-et-Vilaine, 47 : Lot-et-Garonne). Les taux de maturation séparés d'un « ; » correspondent à des parcelles différentes du même site et de la même année

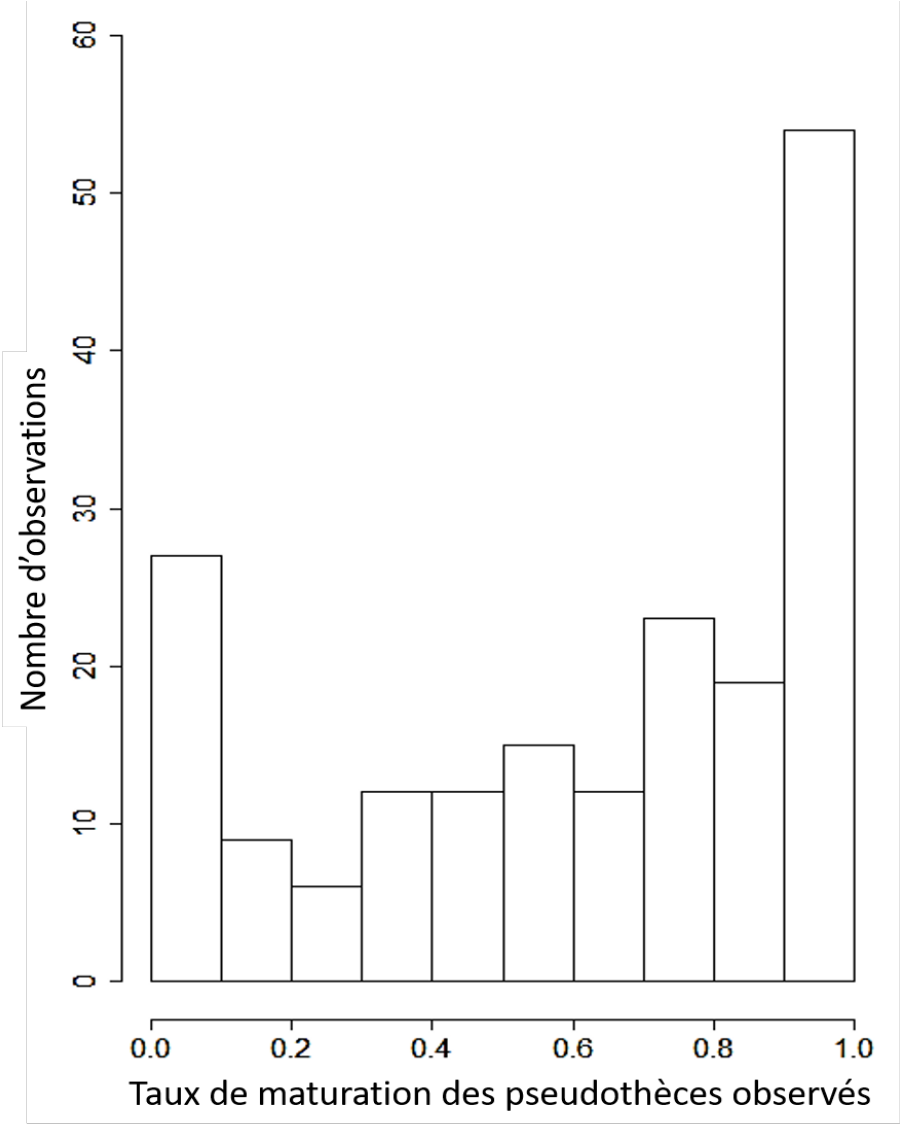


Figure III.1 – Distribution des taux de maturation des pseudothèces observés dans le jeu de données (189 observations issues de 9 départements et pour 5 années).

Les variables climatiques (température moyenne et pluviométrie journalières) proviennent de quatre bases de données : SAFRAN (Météo), et de réseaux de stations météorologiques de l'INRAE, de Terres Inovia et d'Arvalis.

2.2.4 Estimation des paramètres

L'algorithme utilisé pour identifier la combinaison de paramètres (parmi les $2^7 = 128$ combinaisons) à optimiser pour améliorer la qualité de prédiction du modèle est celui présenté dans la Chapitre 5 de la partie II de ce manuscrit.

Les critères utilisés sont les mêmes que ceux décrits pour le phomopsis dans le chapitre précédent. Ils sont exposés dans la Tableau III.4.

Symbole	Critère	Type de qualité	Etape
$MSE^{P_i}_{SM-L.\,l.}$	*MSE*	Prédiction	Initiale
$MSE^{P_{cv1}}_{SM-L.\,l.}$	*MSE*	Prédiction	C.V.
$MSE^{A1}_{SM-L.\,l.}$	*MSE*	Ajustement	Finale
$Biais^{P_i}_{SM-L.\,l.}$	*Biais*	Prédiction	Initiale
$Biais^{P_{cv1}}_{SM-L.\,l.}$	*Biais*	Prédiction	C.V.
$Biais^{A1}_{SM-L.\,l.}$	*Biais*	Ajustement	Finale
$Efficience^{P_i}_{SM-L.\,l.}$	*Efficience*	Prédiction	Initiale
$Efficience^{P_{cv1}}_{SM-L.\,l.}$	*Efficience*	Prédiction	C.V.
$Efficience^{A1}_{SM-L.\,l.}$	*Efficience*	Ajustement	Finale

Tableau III.4 – Critères utilisés dans l'étude, type de qualité et étape de chaque symbole.
SM : SimMat, *L. l.* : *Leptosphaeria lindquistii* ; C.V. : Cross-Validation.

2.3 Résultats

2.3.1 Évaluation de la qualité de prédiction initiale de SimMat-*L. lindquistii* sur les dynamiques de taux de maturation de pseudothèces

La Figure III.2A représente la comparaison entre les observations et les résidus des taux de maturation de pseudothèces prédits en utilisant les valeurs initiales des paramètres. Les résidus ont tendance à décroître avec l'augmentation des taux de maturation observés. Ceci est cohérent : pour les taux observés proches de 0, les prédictions ne peuvent être que supérieures ou égales à 0 tandis que pour les taux observés proches de 1, les prédictions ne peuvent être qu'inférieures ou égales à 1. En revanche, les résidus devraient être centrés autour de 0, or ils sont nettement négatifs, ce qui signifie que les observations sont sous-estimées. Ces observations (Figure III.2) sont confirmées par le critère statistique $Biais^{P_i}_{SM-L.\,l.} = 0,15$ (Tableau III.6).

Paramètres	**Intervalle de valeurs testées**	**Valeur initiale**	**Origine de la valeur**	**Valeur finale**	**Unité**
$\theta_{max}^{\textbf{SM-}\textit{L. lindquistii}}$	**]20 ; 40[**	**29,0**	**Optimisation préliminaire de SimMat-Diaporthe**	**25,0**	**°C**
$\theta_{min}^{SM-\textit{L. lindquistii}}$	]0 ; 20[	6,0	Muhlberger, 1996	6,0	°C
$R_{min}^{SM-\textit{L. lindquistii}}$	]0 ; 20[	11,6	Optimisation préliminaire de SimMat-Diaporthe	11,6	mm
N_d	**[1 ; 30]**	**8**	**Optimisation préliminaire de SimMat-Diaporthe**	**12**	**jour**
N_{FD}	]1; 100[	22,9	Optimisation préliminaire de SimMat-Diaporthe	**12**	jour
σ_{FD}	]1; 100[	15,0	Optimisation préliminaire de SimMat-Diaporthe	15,0	jour
k	**]0; 1[**	**0.046**	**Aubertot *et al.*, 2006**	**0.064**	**mm^{-1}**

Tableau III.5 – Résultats obtenus avec la procédure d'optimisation des paramètres de SimMat-*L. lindquistii*. Intervalle de valeurs testées, valeurs de paramètres avant estimation, origine des valeurs, valeurs de paramètres après estimation, et leurs unités. En gras : nouvelles valeurs obtenues après optimisation.

MSE	**Valeurs**	**Biais**	**Valeurs**	**Efficience**	**Valeurs**	**Étape**
$MSE_{SM-L.\ l.}^{P_i}$	0,099	$Biais_{SM-L.\ l.}^{P_i}$	0,146	$Efficiency_{SM-L.\ l.}^{P_i}$	0,145	Qualité de prédiction initiale
$MSE_{SM-L.\ l.}^{P_{cv1}}$	0,064	$Biais_{SM-L.\ l.}^{P_{cv1}}$	0,032	$Efficiency_{SM-L.\ l.}^{P_{cv1}}$	-1,850	Qualité de prédiction en validation-croisée
$MSE_{SM-L.\ l.}^{A1}$	0,078	$Biais_{SM-L.\ l.}^{A1}$	0,014	$Efficiency_{SM-L.\ l.}^{A1}$	0,327	Qualité d'ajustement

Tableau III.6 – Valeurs de *MSE*, *Biais* et $Efficience$ calculées pour la qualité de prédiction initiale de SimMat-*L. lindquistii*, pour la qualité de prédiction obtenue après la procédure d'optimisation et pour la qualité d'ajustement.

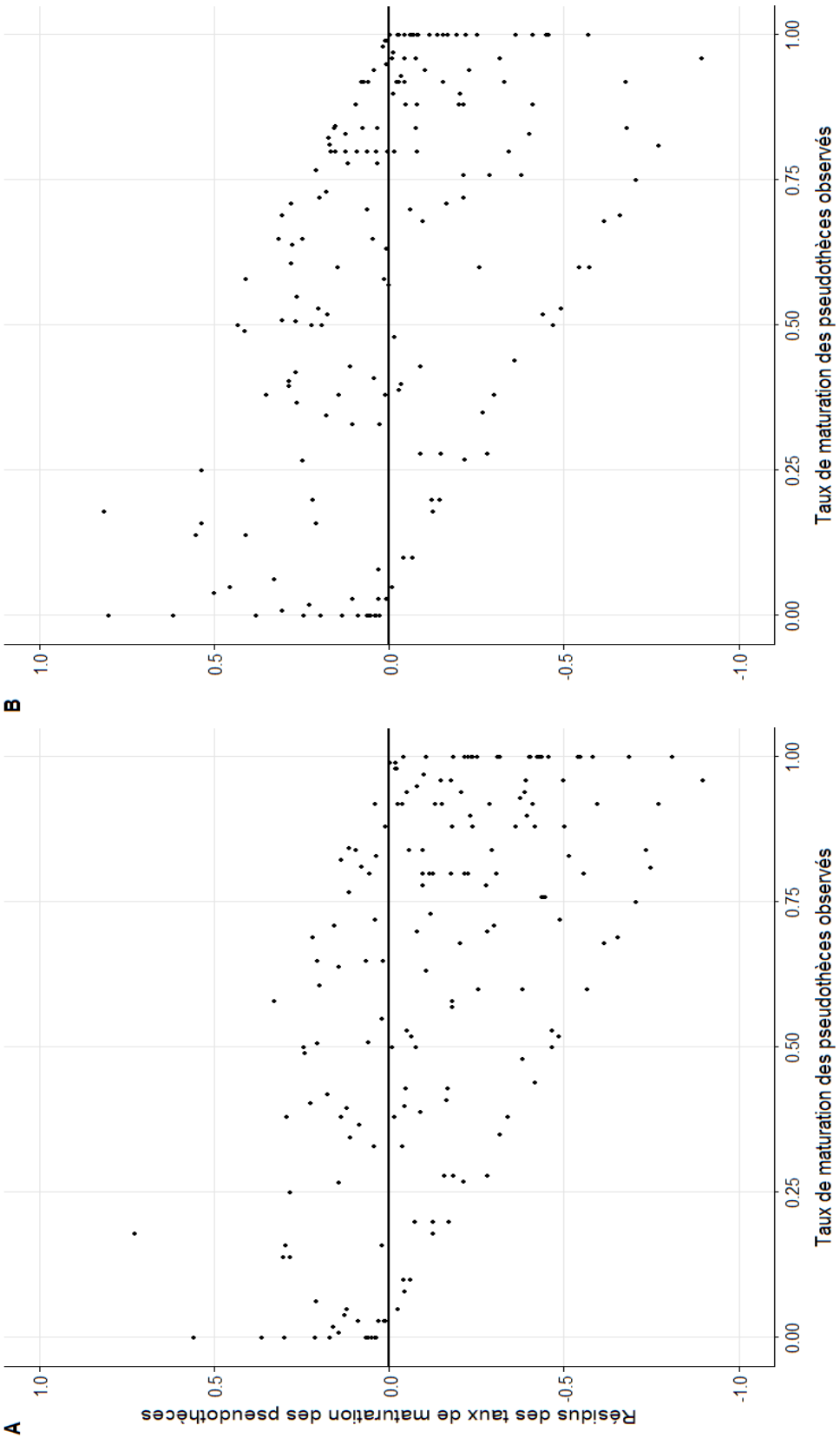


Figure III.2 – Résidus des prédictions de dynamique de SimMat-*L. lindquistii* sur le jeu de données entier (189 observations) avec : les valeurs initiales des paramètres (A) ; les valeurs finales des paramètres après optimisation.

2.3.2 Identification de la combinaison des paramètres de SimMat-*L. lindquistii* à ajuster pour minimiser les erreurs de prédiction des dynamiques de maturation des pseudothèces

L'algorithme a identifié trois paramètres à modifier parmi les sept possibles. $\theta_{max}^{SM\text{-}L.\ lindquistii}$; N_d; k. Les valeurs optimisées ne sont pas restées proches des valeurs initiales, excepté pour k (0,064 contre 0,046 mm^{-1}, Tableau III.5).

L'optimisation des paramètres au cours de la validation croisée a un peu amélioré la mauvaise qualité de prédiction initiale de SimMat-*L. lindquistii* (Figure III.2A et B). La MSE a été réduite d'un tiers ($MSE_{SM-L.\ l.}^{P_{CV1}} = 0,062$ versus $MSE_{SM-L.\ l.}^{P_i} = 0,099$), le biais a été divisé par 4,6 ($Biais_{SM-L.\ l.}^{P_{cv1}} = 0,032$ versus $Biais_{SM-L.\ l.}^{P_i} = 0,146$, Tableau III.6), les observations sont moins sous-estimées (Figure III.2B). En revanche l'efficience est devenue plus mauvaise. Initialement positive, elle est devenue négative ($Efficiency_{SM-L.\ l.}^{P_{CV1}} = -1,85$ versus $Efficiency_{SM-L.\ l.}^{P_i} = 0,111$, Tableau III.6). Ce changement de signe peut être expliqué par une difficulté du modèle à prédire certaines situations durant la validation croisée et les raisons sous-jacentes devraient être explorées par une analyse des résidus de la validation-croisée. La minimisation porte sur le critère MSE dans l'algorithme et non sur l'efficience.

2.3.3 Optimisation des paramètres de SimMat-*L. lindquistii* en utilisant le jeu de données entier

La Figure III.2B représente les résidus des taux de maturation des pseudothèces prédit par SimMat-*L. lindquistii* après ré-estimation des 3 paramètres en utilisant le jeu de données complet. Cette figure semble indiquer qu'il y a eu moins de sous-estimations des observations après optimisation. Le biais confirme cette tendance ($Bias_{SM-L.\ l.}^{A1} = 0,014$ contre $Bias_{SM-L.\ l.}^{P_i} = 0,146$, Tableau III.6). La $MSE_{SM-L.\ l.}^{A1}$ était faible ($MSE_{SM-L.\ l.}^{A1} = 0,078$, Tableau III.6) et l'efficience indique que le modèle explique un tiers de la variabilité des observations ($Efficency_{SM-L.\ l.}^{A1} = 0,327$).

2.4 Discussion

2.4.1 Jeu de données utilisé pour l'estimation des paramètres

Le jeu de données utilisé pour le phoma avait été en partie conçu pour développer un modèle de prévision de la production d'inoculum de *L. lindquistii*. Ce jeu de données est légèrement plus grand que celui utilisé pour le phomopsis (189 observations) et présente l'avantage de bien balayer l'ensemble des valeurs possible des taux de maturation de pseudothèces. En revanche, il est déséquilibré puisque 13% des observations (25 sur 189) sont comprises entre 0,0 et 0,1 ;

30% des observations (56 sur 189) sont comprises entre 0,9 et 1,0 (toutes dates confondues). Les 142 autres observations (57%) sont réparties entre 0,1 et 0,9. Au sein de ce jeu de données, 20 situations ne comportent qu'une seule observation du taux de maturation de pseudothèces, car ces observations n'avaient pas été mises en place pour le développement d'un modèle. La limite de la taille du jeu de données est due au coût d'observation de la maturation des pseudothèces. Ce jeu est représentatif de la situation du phoma en France, car il est constitué de situations réparties dans plusieurs départements français touchés par la maladie, sur plusieurs années. De plus, l'observation la plus récente ayant été réalisée par l'auteure de ce manuscrit au printemps 2016, une éventuelle évolution du cycle de vie de *L. lindquistii* a pu être prise en compte au sein du jeu de données.

2.4.2 Résultats obtenus lors de l'évaluation et de la procédure d'optimisation des paramètres

Le seuil de pluviométrie n'a pas été modifié par l'algorithme, ce qui signifierait que cette valeur (11,6 mm) serait pertinente pour la reproduction sexuée de *L. lindquistii*. Le besoin de pluviométrie pour les émissions d'ascosporess confirme donc l'étude de Délos *et al.* en 1998. Cependant, ce besoin de pluviométrie ne fait pas consensus dans la littérature technique. Une étude préliminaire menée dans plusieurs essais (Descorps, 2010) a montré qu'il n'y aurait pas de relation entre précipitations journalières et émissions d'ascosporess, ce qui conforte les résultats d'une autre étude préliminaire à ce sujet (Guerin, 1997). L'adaptation du modèle SimMat au phoma du tournesol avait alors été proposée en raison de sa capacité à prendre en compte l'alternance de phénomènes pluvieux et de sécheresse sur plusieurs jours.

Le seuil de température minimal initial provient de la littérature ($\theta_{min}^{SM-L.\ m}$ = 6°C, Muhlberger, 1996) et n'a pas été modifié par l'algorithme. Ce seuil issu de la littérature est assez proche de la température minimale observée (8°C) pour *D. helianthi* par Fabregue (1992) et Moinard et Eychenne (1998), mais est un peu plus éloigné de celui observé (10,0°C) pour ce même champignon par Terekhov *et al.* (2004) ; Mathew *et al.* (2018).

Le seuil de température maximale initial provient de la valeur optimisée préliminaire de SimMat-*Diaporthe* ($\theta_{max}^{SM-D.\ h}$ = 29°C). Il a diminué pour le phoma (25°C contre 29°C, Tableau III.5). La valeur de température maximale de notre jeu de données était de 27,5°C. Cette diminution de la valeur du seuil de température maximale n'est pas cohérente avec ce que l'on sait de la biologie de *L. lindquistii* car les émissions d'ascosporess peuvent être observées pendant l'été, saison au cours de laquelle les températures dépassent le seuil obtenu.

La valeur finale de N_d (12 jours) pour *L. lindquistii* est différente de la valeur initiale correspon-

dant à l'ancienne valeur optimisée pour le phomopsis (8 jours, SimMat-*Diaporthe*). La valeur du paramètre k est restée proche de la valeur initiale provenant de SimMat pour le phoma du colza (k = 0,064 mm^{-1} contre k = 0.046 mm^{-1}).

Sur les 5 paramètres issus de SimMat-*Diaporthe*, seul σ_{FD} a gardé la même valeur. La reproduction sexuée de *L. lindquistii* est donc impactée différemment par le climat que celle de *Diaporthe helianthi*. Les valeurs des critères obtenues semblent indiquer que la maturation et les projections d'ascosporess du phoma du tournesol ne fonctionnent pas de la même manière que pour le phomopsis du tournesol et le phoma du colza. Les résultats insuffisants de cette adaptation pourraient être dû aux variables d'entrée (pluviométrie et température moyenne journalières) inadaptées, ou au formalisme mathématique non adapté pour le phoma du tournesol. En effet, une comparaison préliminaire avait été menée entre les performances d'ASPHODEL (sans adaptation) et SimMat sur le phoma du tournesol (avec un jeu de données moins fourni que celui présenté) et avait montré qu'ASPHODEL obtenait de meilleurs résultats (Desanlis, 2013). Mais pour des raisons conceptuelles et pratiques, nous n'avons pas tenté d'adapter ASPHODEL au phoma.

2.4.3 Conclusion

De manière générale, le modèle SimMat-*L. lindquistii* a donné de moins bonnes performances que le modèle SimMat-*Diaporthe* puisqu'il n'explique qu'un tiers de la variabilité observée alors que SimMat-*Diaporthe* en explique deux-tiers (ajustement à l'ensemble des données)..

Il a donc été décidé de ne pas mobiliser le modèle SimMat-*L. lindquistii* pour le modèle IPSIM correspondant (*c.f.* partie suivante). Cependant, si ce modèle n'est pas exploitable, le protocole présenté dans la seconde partie de ce manuscrit a démontré encore une fois sa capacité à améliorer la qualité prédictive d'un modèle quantitatif de faible complexité Cet algorithme peut donc être utilisé pour adapter SimMat à des champignons ayant une biologie proche de celle de *D. helianthi* sur le tournesol ou d'autres cultures.

IV Modèles IPSIM tournesol

1 IPSIM - Tournesol - Phomopsis 171

1.1 Design and evaluation of IPSIM-Sunflower-Phomopsis, a model that predicts the severity of Phomopsis stem canker on Sunflower

1.2 Amélioration de la qualité de prédiction d'IPSIM - Tournesol - Phomopsis

2 IPSIM - Tournesol - Dessèchement précoce 225

2.1 Identification, organisation des attributs, définition de leurs valeurs et de leurs poids

2.2 Description du convertisseur

2.3 Évaluation future du modèle IPSIM - Tournesol - Dessèchement précoce

3 IPSIM - Tournesol - Orobanche 247

3.1 Construction sur le modèle IPSIM - Tournesol - Orobanche

3.2 Commentaires sur le modèle, synthèse bibliographique

3.3 Variables à développer

3.4 Convertisseur à développer

1. IPSIM - Tournesol - Phomopsis

Ce modèle a été développé selon la méthode présentée en chapitre 2 de la deuxième partie de ce manuscrit. Cinq ateliers mobilisant Mme. Bret-Mestries (phytopathologiste des maladies du tournesol, Terres Inovia), M Debaeke (agronome spécialiste du tournesol, INRAE), Mme. Seassau (enseignante-chercheuse agronome spécialiste des maladies du tournesol, EI Purpan), Mme. Robin (enseignante-chercheuse agronome, EI Purpan), M Dechamp-Guillaume (phytopathologiste spécialisé sur les maladies du tournesol, ENSAT) et M Aubertot (agronome modélisateur, INRAE), ont été nécessaires pour le finaliser. Il fera l'objet d'une soumission à *PLOS One* en 2020.

1.1 Design and evaluation of IPSIM-Sunflower-Phomopsis, a model that predicts the severity of Phomopsis stem canker on Sunflower

Ce chapitre présente le développement et l'évaluation de la qualité de prédiction d'un nouveau modèle, baptisé IPSIM-Tournesol-Phomopsis (IPSIM-Sunflower-Phomopsis) prédisant la sévérité du phomopsis du tournesol. Un groupe d'experts a proposé un modèle de prédiction qualitatif en mobilisant l'approche IPSIM présentée dans la deuxième partie de la thèse. Une base de données comportant 1073 parcelles, réparties dans 16 départements de sept régions françaises et sur 12 années (entre 2002 et 2018) a été utilisée pour évaluer sa qualité de prédiction. Le modèle pourra être utilisé pour aider à la conception d'itinéraires techniques du tournesol limitant le risque Phomopsis.

Design and evaluation of IPSIM-Sunflower-Phomopsis, a model that predicts the severity of Phomopsis stem canker on Sunflower

Marie-Anne Vedy-Zecchini[a], Emmanuelle Bret-Mestries[b], Marie-Hélène Robin[a], Philippe Debaeke[a], Celia Seassau[a], Gregory Dechamp-Guillaume[a], Jay Ram Lamichhane[a], Jean-Philippe Deguine[c], Jean-Noël Aubertot[a*]

[a]*INRAE-INPT-ENSAT-EI-Purpan, University of Toulouse, UMR 1248 AGIR, Centre de recherche Occitanie-Toulouse, Auzeville, France*

[b] *Terres Inovia, Centre de recherche Occitanie-Toulouse, Auzeville, France*

[c]*CIRAD, UMR PVBMT, F-97410 Saint-Pierre, La Réunion, France*

* I am the corresponding author
Email address: jean-noel.aubertot@inrae.fr

To be submitted to PLOS One, 2020

Abstract

Phomopsis stem canker, caused by *Diaporthe helianthi*, is one of the main diseases of sunflower. Preventive and curative methods are available to manage this fungal disease. However, there is a need for a tool that helps design integrated management of this disease. It is hypothesized that the severity of the disease is determined only by the combinations effects of various cropping practices, in addition to the effects of weather and field environment. A new model, called IPSIM-Sunflower-Phomopsis, was designed that uses the generic IPSIM (Injury Profile SIMulator) platform. This model is a deterministic hierarchical qualitative model using ordinal variables as input variables. The model was developed according to information published in scientific and technical literature, besides feedback of specialized phytopathologists and agronomists. In order to ease the model's use, a converter was designed to convert quantitative (e.g. sowing density) or nominal variables (e.g. the name of the cultivar) into ordinal variables. The model and its converter were then evaluated using an independent dataset with 1073 situations observed across seven 7 French regions from 2002 to 2018. The model IPSIM-Sunflower-Phomopsis and its converter correctly predicted 79\% of the observed phomopsis stem canker severities, and explained 52\% of the variability of the dataset. Despite its fair quality of prediction, this model can be used to help design sustainable sunflower crop management strategies across various production situations, through a better disease management.

Keywords: *Helianthus annuus*, Epidemiology, Qualitative modelling, DEXi, cultural management

1. Introduction

Sunflower (*Helianthus annuus L.*) as oil crop ranks fourth globally after palm, soybean and oilseed rape (Pilorgé, 2016). In France, this is the second oil crop in terms of the acreage following rapeseed. Sunflower is considered as an environmental-friendly crop for several reasons including a good tolerance to water and nutrient stresses (Debaeke *et al.*, 2017). This crop represents a good preceding crop as it improves soil structure and nutrient leaching risks (Lecomte and Nolot, 2011). Despite its several environmental and agronomic advantages, the sunflower acreage has reduced by 500,000 ha in the last decade. Likewise, the sunflower yield in France ranges between 20 and 25 $q.ha^{-1}$ (Vear, 2016) although it has potential to reach 45 $q.ha^{-1}$ (Skoric *et al.*, 2007). Such a poor crop yield is mainly due to unfavorable soil characteristics as well as biotic stresses (Jouffret *et al.*, 2011; Mestries *et al.*, 2011; Debaeke *et al.*, 2017).

Fungal diseases represent the most important biotic stresses limiting sunflower production (Mestries *et al.*, 2011, Debaeke *et al.*, 2017). Overall downy mildew (caused by *Plasmopara halstedii*), sclerotinia (caused by *Sclerotinia sclerotiorum*), phomopsis stem canker (caused by *Diaporthe helianthi*) and phoma black stem (caused by *Leptosphaeria lindquistii*) represents the most important fungal diseases in decreasing order of importance. Nevertheless, the economic importance of these fungal diseases may differ from one region to another, due mainly to different pedo-climatic conditions and management practices. Here, our focus is phomopsis stem canker because the disease is present across all sunflower growing areas in France and to a higher extent in Southwestern region where the disease was first reported (Perny, 1992). This disease causes up to 3 $q.ha^{-1}$ when only 10% of plants in a field have girdling spots, the characteristic symptom of the disease (Masirevic and Gulya, 1992; Estragnat, 1992). Previous studies provided important insights into biology of the causal agent and epidemiology of the disease. It has been reported that *D. helianthi* can survive up to a year on infected stubble present on the soil surface where the pathogen performs its sexual and asexual reproductions (Fayret and Assemat, 1987; Slyusa *et al.*, 1988). *D. helianthi* is considered as a monocyclic pathogen since its sexual reproduction is the only infectious phase of the cycle (Mihaljcevic *et al.*, 1985). The mycelium performs its sexual reproduction by producing pseudothecia during autumn (Muntanola-Cvetkovic

97 *et al.*, 1988). The maturation of these pseudothecia halts during the winter but it re-begins at the end of
98 winter. Pseudothecia produce asci that contain ascospores (primary inoculum), which are discharged during
99 rainfall events from May to August (Fayret and Assemat, 1987; Petrov *et al.*, 1981). Ascospores are either
100 splash or air-dispersed (Perny, 1992; Peres and Régnault, 1988) and may contaminate sunflower leaves up to
101 10 days after their release (Bertrand and Tourvieille, 1987; Fabregue and Eychenne, 1998). Under suitable
102 environmental conditions, ascospores germinate on leaf surfaces and the fungal mycelium penetrates into
103 leaf organs where it has a latent phase. The duration of this phase depends on the phenological stage of
104 sunflower at the time of infection (Penaud *et al.*, 1995). The first disease symptom on leaves appears at flower
105 bud stage (Penaud *et al.*, 1995), which is the most receptive crop phase to phomopsis stem canker (Delos *et*
106 *al.*, 1995). Subsequently, the fungal mycelium starts to move towards the stem that can take over three weeks
107 (over five weeks in case of water stress, Penaud *et al.*, 1995; Desanlis *et al.*, 2013) thereby causing girdling
108 symptoms within two weeks. This finally affects nutrients flows and water circulations killing the entire plant.
109 Following harvest, fungal mycelium continue to survive on crop residues and re-start its cycle again. Several
110 disease management levers are deployed aiming at reducing the impact of phomopsis stem canker on
111 sunflower productivity. Chemical control is a frequent practice although sunflower is less reliant on pesticides
112 than many other field crops (Mihaela, 2019). Four active ingredients (azoxystrobin, difenoconazole, boscalid
113 and dimoxystrobin) are authorized in UE to control this disease with a satisfactory level of disease control
114 (prevention of yield losses up to 10 $q.ha^{-1}$) (Duroueix and Penaud, 2013). Nevertheless, increasing awareness
115 on human health and the environment, due to negative effects of conventional pesticides, call for alternative
116 disease management levers to these pesticides (Lamichhane *et al.*, 2016}. Consequently, several public policy
117 initiatives are ongoing in the EU, and, in France, there is a national action plan called Ecophyto that aims at
118 reducing reliance on conventional pesticides in French agriculture (Lamichhane *et al.*, 2019). The Ecophyto
119 plan thus prioritizes research and development initiatives in order to develop alternatives to conventional
120 pesticides. Because disease management relying on a single solution is unlikely to work in the long term,

121 independent of chemical or biological origin, a combination of different levers is needed that can be adapted
122 to local and site-specific contexts to achieve the ultimate goal of sustainable crop protection.

123 Agroecological Crop Protection (ACP) is an approach that aims at providing ecologically sustainable
124 management of crop pests (pests *sensu lato* that includes animal pests, weeds and pathogens). ACP is the
125 application of agroecological principals to crop protection (Gliessman, 2014). ACP aims at fostering the
126 resilience of agroecosystems through application of ecological practices (Deguine *et al.*, 2017). Its
127 implementation in the field is based on a strictly planned phytosanitary strategy, giving priority to preventive
128 measures, based on non-chemical tactics. Curative methods, relying on pesticides, however, can be
129 considered, but only as a last option, and under strict monitoring, in order to foster the effectiveness of non-
130 chemical methods.

131 ACP gives priority attention to the functioning of the agroecosystem and promotes the interactions of
132 vegetal, animal and microbial communities, both above- and below-ground (Aubertot *et al.*, in press). In
133 addition, this approach aims to promote soil health, which in turn enhance crop health. Soil health and
134 biodiversity are thus the basis of ACP. Early detection of a risk due to phomopsis stem canker can allow the
135 implementation of agroecological techniques to limit crop damage.

136 This is especially true when the system considered include a spatial scale (agroecosystem or landscape
137 scales). This requires information sharing and training programs for sunflower growers. At the same time, the
138 implementation of ACP principles is challenging due to number of factors involved and their interactions
139 within an agro-ecosystem. Consequently, the prediction regarding effects of one or different cropping
140 practices modifications on the overall system can be complex. In such a case, farm advisers need a model to
141 take into account this complexity that can effectively predict these combined effects on phomopsis stem
142 canker and help ease their decision-making process. This is feasible only with a Decision Support System (DSS)
143 model.

144 A few models have been developed and used for sunflower that include both generic crop models such
145 as STICS (Brisson *et al.*, 1998) as well as specific crop model including SUNFLO (Casadebaig *et al.*, 2011) and
146 ASPHODEL (Délos *et al.*, 1995). While the previous two are crop-models that do not take into account pests,
147 the latter model predicts pseudothecial maturation, ascospores dissemination and contamination of
148 sunflower as a function of hourly and daily weather variables. In addition, ASPHODEL is included into DSS to
149 help design pesticides applications. Recently, we developed a model called SimMat-*Diaporthe* (Vedy-Zecchini
150 *et al.*, submitted) that predicts pseudothecial maturation and ascospore dissemination as a function of mean
151 daily temperature and daily rainfall. Nevertheless, none of these models predicts the injury caused by
152 phomopsis nor they consider interactions between the production situation (Robin *et al.*, 2013), cropping
153 practices (Sébillotte, 1990a) or the pathosystem.

154 They do not consider interactions between the situation of production (defined as the abiotical and
155 biotical components of a field; De Wit and Penning de Vries, 1982; Breman and De Wit, 1983; Rabbinge,
156 1993), the cropping practices (cropping practices are the temporal and spatial sequence of activities applied
157 by a farmer, the social-economic context is not considered in this study; Sebillotte, 1990a), the sunflower,
158 the phomopsis stem canker and the tremendous interactions. Injury Profile SIMulator (IPSIM) is a modelling
159 platform that permits to develop qualitative aggregative models (Aubertot and Robin, 2013). The temporal
160 scale of the variables of these models are the growing season of a given crop while the spatial scale is a field.
161 IPSIM models aim at predicting injury level caused by one pest in a field as a function of a set of variables
162 related to the production situation and cropping system. IPSIM models integrate a toll called “converters”,
163 which are informatic protocols that translate quantitative data into qualitative ones for IPSIM. This modelling
164 framework has been successfully used to develop and test IPSIM models such as eyespot (Robin *et al.*, 2013)
165 and brown (Robin and Aubertot, 2017) of wheat. Based on this, we chose this modelling framework to create
166 IPSIM - Sunflower - Phomopsis stem canker. This study presents this model and its converters including the
167 evaluation of its prediction quality.

168

2. Materials and methods

2.1 Description of IPSIM-Sunflower-Phomopsis stem canker, identification and organisation of attributes and definition of their Scales

2.1.1 Description of IPSIM-Sunflower-Phomopsis stem canker

IPSIM-Sunflower-Phomopsis stem canker is based on the DEX (Decision EXpert) method and implemented with the software DEXi (http://kt.ijs.si/MarkoBohanec/dexi.html, Bohanec *et al.*, 2013). The DEX method is a qualitative hierarchical aggregative method. It has an output depending of several variables called aggregated attributes (internal nodes). These attributes are sub-divided in other variables until input variables of the model. Each attribute has a name, a description and a scale. All the variables of the tree are qualitative. Aggregative tables composed of aggregative rules also compute aggregative attributes. The entire model is defined using available information in scientific and technical literature and as well as expert's knowledge (Table 1). The principle of IPSIM and the generic structure are described by Aubertot and Robin in 2013 and Robin *et al.*, in 2013.

The synthesis of knowledge is presented Table 1.

Phomopsis stem canker				Effects	Dir.	Int.	Literature
Injuries				10 to 30% of injured sunflower in a field result in a loss 10 to 30 g.m^{-2}			Estragnat, 1992 ; Pinochet, 1995; Jouffret, 2005
Phomopsis stem canker incidence at landscape level	Level risk			The level of phomopsis stem canker's presence in the field depends on its presence at the farm level. The ascospores are dispersed by rain and wind. The presence at the farm level is well represented by maps designed by Terres Inovia	-	++	Petrov *et al.*, 1981; Fayret and Assemat, 1987; Maric *et al.*, 1988; Perny, 1992;
	Tillage after the last sunflower			Tillage and burying of infected stubble increase the decaying and decrease primary inoculum quantity	+	++	Masirevic and Gulya, 1992; Jouffret, 2005
Weather	Infection			Regular or frequent rainfall and irrigation increase infection success. Two successive days with ejection during vegetative stage are considered as sufficient to give a contamination. An ascospore germinate if there is water (>=20 mm) on leaves. They can germinate until 10 days.	-	+++	Petrov *et al.*, 1981; Fayret and Assemat, 1987; Jinga *et al.*, 1987; Masirevic and Gulya, 1992; Delos *et al.*, 1995; Penaud, 1995 ; Debaeke *et al.*, 2003b; Debaeke and Estragnat., 2003; Debaeke and Moinard, 2010; Desanlis *et al.* 2013
	Symptoms development			Mycelia can develop on leaves and petiole during 25 days if infection occurred before anthesis and temperature is comprised between 15 and 30°C and if rainfall is available (35 mm).It can take 40 days in case of abiotic stresses. Four expositions at 32°C kill *D. helianthi*	-	+++	
Cropping practices	Cropping control	Mitigation through crop status		LAI at flowering is related to girdling symptom proportion. LAI depends on density, N- fertilization Nitrogen fertilization together stimulates the leaf area index, increasing the probability of ascospores contact. Plant density affects the stem diameter and the rapidity of canopy closure. Thinner the stem diameter, the fastest girdling symptoms are. The mitigation control impact the number of infested plants by creating a microclimate favourable for sunflower infection.	-	++	Iliescu and Baicu, 1984; Stanojevic, 1985; Blanchet *et al.*, 1987; Jinga *et al.*, 1992; Debaeke and Raffaillac, 1996; Connor and Hall, 1997; Debaeke *et al.*, 2000; Debaeke *et al.*, 2003; Debaeke and Estragnat, 2003; Taverne, 2005; Debaeke and Estragnat, 2009; Debaeke *et al.*, 2009; Debaeke and Moinard, 2010; Berard, 2019 (p.14); Desanlis *et al.* 2013; (Debaeke *et al.*, 2001)
		Escape	Sowing date	A late sowing (end of April instead of beginning of April) can decrease 30% the number of injured sunflower. A cultivar that have an early flowering is more injured than a cultivar that late flowering	+	+	Fazekas, 1989; Jinga *et al.*, 1992; Perny *et al.*, 1995; Délos and Moinard, 1997 ; Gulya *et al.*, 1997; Debaeke *et al.*, 2003; Debaeke and Estragnat, 2003; Taverne, 2005; Debaeke and Estragnat, 2009; Desanlis *et al.* 2013; (Debaeke *et al.*, 2001)
			Primary inoculum production	Mycelium survives on infected stubble stayed at the soil surface and produces primary inoculum at spring. 0.5 mm of rain (or more), a temperature above 6°C (optimal temperature: 18°C) and a relative humidity above 85% are favourable for inoculum production. The Bulletin de Santé du Végétal (BSV) provides information on situations where there is a risk of contamination. SimMat-*Diaporthe* is a model predicting ascospores maturation and releases	-	++	Muntanola-Cvetkovic *et al.*, 1981; Muntanola-Cvetkovic *et al.*, 1985; Mihaljcevic *et al.*, 1985; Fayret et Assemat, 1987; Vedy-Zecchini *et al.*, 2019, submitted;
	Cultivar choice			There are five levels of susceptibility: very susceptible, susceptible, less susceptible, very less susceptible and resistant. The choice must be adapted in function of level risk at farm level. A susceptible cultivar will be more injured in situation of uncontrolled density and N-fertilization	+	+++	Bertrand and Tourvieille de Labrouhe, 1987; Vear *et al.*, 1997; Viguié *et al.*, 2000b; Pinochet, 2005; Debaeke and Estragnat, 2009; Debaeke et Moinard, 2010; www.myvar.fr (Penaud et Jouffret, 1997)
	Chemical control			Two active substances (azoxystrobin, difenoconazole) are officially registered against phomopsis stem canker in France. Their respective efficacies are considered as sufficient as they can increase yield by 100 g.m^{-2} (10 q/ha) on average in a field with 50% of sunflower with girdling spots. The reasoning behind the fungicide application is related to the BSV of phomopsis. The efficacy of the products makes it possible to reduce from 50% of harmful symptoms (untreated control) to 20% of harmful symptoms and with a yield gain of around.	+	++	Debaeke and Estragnat, 2009; Duroueix et Penaud, 2013; https://www.terresinovia.fr/-/phomopsis-sur-tournesol-ne-traiter-que-si-necessaire?p_r_p_categoryId=130441&p_r_p_tag=69906&p_r_p_tags=282115

Table 1 Summary knowledge on phomopsis stem canker including key factors, their effect on the disease, the potential outcome (Dir), the impact on disease intensity (Int) and References. Dir: (-) the increase of value of the variable increase the phomopsis risk, (+) the increase of value of the variable decrease the phomopsis risk. Int: (+) the variable has a low effect, (++) the variable has an intermediate effect, (+++) has a strong effect

192 The hierarchical structure of IPSIM-Sunflower-Phomopsis stem canker is composed of five aggregative (in
193 bolded terms) and eight basic (in non-bolded terms) attributes (Figure 1). The output of the model is
194 represented by the attribute "Phomopsis stem canker incidence" (girdling spots incidence at the maturity;
195 Lancashire *et al.*, 1991). It is determined by three sub-attributes: "Cropping practices", "weather" and
196 "Phomopsis stem canker pressure in the field".

Attribute	Scale
Phomopsis stem canker incidence	High; Intermediate; *Low*
├**Phomopsis stem canker pressure in the field**	High; Intermediate; *Low*
├Risk level at the farm level	High; Intermediate; *Low*
└Mechanical grinding and burying in fields around	No burying and griding; *Burying and griding*
├**Weather during Spring and/or Summer**	Unfavourable; Moderately favourable; *Favourable*
├Weather for infection during Spring and/or Summer	Unfavourable; *Favourable*
└Weather for girdling symptoms during Summer	Unfavourable; Moderately favourable; *Favourable*
└**Cropping practices**	Unfavourable; Moderately unfavourable; *Favourable*
├**Cropping control**	Unfavourable; Moderately favourable; *Favourable*
├Escape strategy	Inefficient; *Efficient*
└Mitigation through crop status	Unfavourable; Moderately favourable; *Favourable*
├Genetic control	Susceptible and few susceptible; *Very few susceptible and resistant*
└Chemical control	No treatment; *Treatment*

197

Figure 1. Structure and scale of attributes of IPSIM-Sunflower-Phomopsis stem canker attributes.

198

199 The first main sub-tree "Phomopsis stem canker pressure in the field" considers that the risk of phomopsis
200 stem canker's presence in a given field depends on its presence at the farm level (the farm rotation and farms
201 around) and tillage. Fungal mycelium can survive on infested stubble left at soil surface during *ca.* one year
202 and ascospores are known to be rain and wind- dispersed (Fayret and Assemat, 1987; Petrov *et al.*, 1981) at
203 a high distance (ten kilometers; Perny, 1992; Maric *et al.*, 1988). We arbitrarily considered that the risk at
204 farm level is well represented by risk level in the department of the field. The degradation of infested stubble
205 in a given field can be facilitated by mechanical grinding and their burying into the soil before sowing. This
206 helps decrease the mycelium survival and the consequent production of primary inoculum (Masirevic and
207 Gulya, 1992; Jouffret, 2005}. Therefore, crop sequences with two successive sunflower cropping are not
208 allowed.

209 The second main sub-tree "Weather effect" represents the effects of weather (daily mean temperature and
210 rainfall) on two phases of the sexual reproduction of *D. helianthi*: the success of leaf infections and

progression from leaves toward stem. This generally occurs during spring and summer (May and June). Frequent or abundant rainfall (together with irrigation) is the most important factor for successful infection (Debaeke *et al.*, 2003). An irrigation (or rainfall) of 20 mm seems to be sufficient for an infection to occur, in absence of warm periods (Debaeke *et al.*, 2003a; Debaeke and Moinard, 2010). An irrigation applied at flowering stage can increase the infected sunflower proportion. Girdling symptoms appear and causes severe losses only when higher than 70 mm of water is available for the plant between the leaf infection and senescence, and under 35 mm the symptoms do not appear (Debaeke and Moinard, 2010). Irrigation can increase girdling symptoms proportion by stimulating evapotranspiration (Debaeke *et al.*, 2003; Desanlis *et al.*, 2013). Temperatures higher than 32°C can kill the fungus (Jinga *et al.*, 1987). The stem lesion formation halts with the beginning of senescence (Debaeke and Estragnat, 2003).

The third main sub-tree ”Effect of cropping practices” illustrates the way a farmer can manage the agro-ecosystem at several stages: cultivar choice, escape strategy, mitigation through crop status, and chemical control. Several experiments showed that these variables interact with each other and finally affect the proportion of infected sunflower (Debaeke *et al.*, 2003).

i) Cropping strategy consists of the mitigation of crop infection through crop status and escape strategies. The mitigation can be impacted by plant density and/or nitrogen fertilization (Desanlis *et al.*, 2013). These variables increase the speed of canopy closure (Blanchet *et al.*, 1987) and create a humid microclimate favorable for ascospore germination and the consequent infection (Stanojevic, 1985; Masirevic and Gulya, 1992, Debaeke and Raffaillac, 1996, Connor and Hall, 1997, Debaeke *et al.*, 2000). An increase in planting density from 5 to 7.5 plants.m^{-2} increase the rate of infected plants by 22% and that of girdling symptoms by 82%. A thinner stem is more sensitive to a rapid destruction due to girdling symptoms (Debaeke and Moinard, 2010). Therefore, plant density should not exceed 6.5 plants.m^{-2} (Debaeke *et al.*, 2003) to slow down the canopy closure. The proportion of plants with girdling symptoms can be limited with a balanced nitrogen fertilization (Iliescu and Baicu, 1984; Jinga *et al.*, 1992; Debaeke *et al.*, 2003; Debaeke and Moinard, 2010; Desanlis *et al.*, 2013) that depends on the initial nitrogen level. For example, a soil with an initial nitrogen

content of 60 units is considered as moderately rich in nitrogen and thus needs less than 40 units. The combination of plant density and nitrogen fertilization is synthesized by the Leave Area Index (*LAI*) value at the anthesis (Debaeke *et al.,* 2003a, Debaeke and Estragnat, 2009}. *LAI* is used as an indicator of the effect of mitigation through crop status on phomopsis stem canker since a relation can be established between *LAI* and the proportion of injured sunflower (Debaeke *et al.*, 2003). A dry season can decrease the canopy closure, reducing the proportion of infected leaves and stem (Debaeke *et al.*, 2003). An escape strategy can be adopted to shift the receptive period (flower bud) of sunflower (Acimovic and Straser, 1981) to ascospores of *D. helianthi*. A late sowing (for example beginning of May instead of beginning of April) can delay the sunflower bud development stage (the most receptive to phomopsis stem canker; Delos, 1996) even in case of a high *LAI* (Jinga *et al.*, 1992; Perny *et al.*, 1995; Delos and Moinard, 1997; Gulya *et al.*, 1997; Desanlis *et al.*, 2013; Debaeke *et al.*, 2003). Furthermore, a cultivar characterized by early flowering is more exposed to spore release than a late flowering cultivar (Desanlis *et al.*, 2013; Fazekas, 1989).

ii) Genetic resistance against phomopsis stem canker has been well developed in France since the beginning of 1990 (Jouffret, 1997). Several studies showed that the resistance of sunflower to phomopsis stem canker has a polygenetic nature (Deglene *et al.*, 2015; Viguie *et al.*, 2000), meaning that it is based on the accumulation of minor genes. The resistance can act at several level: prevention of ascospore germination (Bertrand and Tourvieille, 1987), retardation of parasite progression (Vear *et al.*, 1997) and blockage at the petiole-stem passage (Bertrand and Tourvieille, 1987). New crop genotypes with a good resistance level are regularly released (Jouffret, 2005; http://www.myvar.fr) and farmers choose their cultivars based on the level of phomopsis risk (http://www.myvar.fr) in their region. Overall, sunflower cultivars are classified into five levels of susceptibility to the disease: resistant, very less susceptible, less susceptible, susceptible, and very susceptible (http://www.myvar.fr).

iii) Lastly, a curative method (*i.e.* use of fungicides) can be adopted at the flower bud stage at the latest and when applied in association with irrigation can increase the yield gain (Debaeke and Estragnat, 2003). It is

260 justified when mitigation through crop status is not enough to contain infection risks due to girdling
261 symptoms (Debaeke *et al.*, 2003).
262

263 **2.1.2 Scales of IPSIM-Sunflower-Phomopsis stem canker**

264 Each of the basic and aggregated attributes has a scale. These scales are nominal or ordinal; they are
265 described by word or sentence. IPSIM-Sunflower-Phomopsis stem canker uses a three-grade value scale (*i.e.*
266 "Unfavorable", "Moderately favorable", "Favorable") for aggregative tables and a two to three-grade value
267 scale for basic variables. The value "Unfavorable" means that the attribute is detrimental to the crop and
268 therefore favorable for the disease. The scales used in the model are presented Figure 1 at right-hand side.
269 The values in red, black and green represent values favorable, neutral and favorable for phomopsis stem
270 canker, respectively. The five aggregative tables are presented Table 2.

	Phomopsis stem canker pressure in the field	Weather during Spring and/or Summer	Cropping practices	Phomopsis stem canker incidence
	27%	38%	35%	
1	<=Intermediate	Unfavourable	<=Moderately unfavourable	High
2	<=Intermediate	<=Moderately favourable	Unfavourable	High
3	High	<=Moderately favourable	Favourable	Intermediate
4	<=Intermediate	Unfavourable	Favourable	Intermediate
5	High	Moderately favourable	>=Moderately unfavourable	Intermediate
6	<=Intermediate	Moderately favourable	Moderately unfavourable	Intermediate
7	<=Intermediate	Favourable	Unfavourable	Intermediate
8	Low	Unfavourable	<=Moderately unfavourable	Intermediate
9	Low	<=Moderately favourable	Unfavourable	Intermediate
10	*	Favourable	>=Moderately unfavourable	Low
11	>=Intermediate	>=Moderately favourable	Favourable	Low
12	Low	*	Favourable	Low
13	Low	>=Moderately favourable	>=Moderately unfavourable	Low
14	Low	Favourable	*	Low

271

	Risk level at the farm level	Mechanical grinding and burying in fields around	Phomopsis stem canker pressure in the field
	33%	67%	
1	High	No burying and griding	High
2	Intermediate	No burying and griding	Intermediate
3	*	Burying and griding	Low
4	Low	*	Low

272

	Weather for infection during Spring and/or Summer	Weather for girdling symptoms during Summer	Weather during Spring and/or Summer
	67%	33%	
1	Unfavourable	Unfavourable	Unfavourable
2	Unfavourable	Moderately favourable	Moderately favourable
3	*	Favourable	Favourable
4	Favourable	*	Favourable

273

	Cropping control	Genetic control	Chemical control	Cropping practices
	27%	36%	36%	
1	<=Moderately favourable	Susceptible and few susceptible	No treatment	Unfavourable
2	Unfavourable	Susceptible and few susceptible	*Treatment*	Moderately unfavourable
3	Unfavourable	*Very few susceptible and resistant*	No treatment	Moderately unfavourable
4	*	*Very few susceptible and resistant*	*Treatment*	*Favourable*
5	>=Moderately favourable	*	*Treatment*	*Favourable*
6	>=Moderately favourable	*Very few susceptible and resistant*	*	*Favourable*
7	*Favourable*	*	*	*Favourable*

	Escape strategy	Mitigation through crop status	Cropping control
	47%	53%	
1	*	Unfavourable	Unfavourable
2	Inefficient	>=Moderately favourable	Moderately favourable
3	*Efficient*	>=Moderately favourable	*Favourable*

Table 2 Aggregative table of incidence of IPSIM - Sunflower - Phomopsis stem canker. Legend: in red: favorable to phomopsis stem canker, in dark: intermediate, in green: unfavorable to phomopsis stem canker. * indicate all the possible values of the concerned variable

The attributes “Phomopsis stem canker incidence”, “Weather during Spring and/or Summer”, “Cropping practices” and “Cropping control” are composed of 27, 6, 12, and 6 aggregative rules, respectively. This means that there are in total 2*3*2*2*2*3*3*2= 864 combinations possible of input values.

*2.1.3 Weights of IPSIM-Sunflower-*Phomopsis *stem canker*

All variables have two weights. They define their relative importance in the model. They are described Table 3.

- Local weight of a variable represents the weight of that variable for the aggregative variable by considering the number of values of the variable.
- Global weight of a variable represents the weight of that variable for the output variable by considering the number of values of the variable.

For example, weather variable, Cropping Practices and the Phomopsis stem canker pressure in the field have a global weight of 38%, 35% and 27%, respectively. This means that Weather during spring and/or summer has the higher influence on output variable.

Variable	Local level 1	Local level 2	Local level 3	Global level 1	Global level 2	Global level 3
Phomopsis stem canker pressure in the field	**27**			**27**		
Risk level at the farm level			43			12
Mechanical grinding and burying in fields around			57			15
Weather during Spring and/or Summer	**38**			**38**		
Weather for infection during spring and/or summer			57			22
Weather for girdling symptoms during Summer			43			16
Cropping practices	**35**			**35**		
Cropping controls		**36**			**12**	
Escape strategy			37			5
Mitigation through crop status			63			8
Genetic control			32			11
Chemical control			32			11

Table 3 Local and global weights of attributes of IPSIM-Sunflower-Phomopsis stem canker

292

293 **2.2 Description of converters**

294 The input values provided by a user can be quantitative or qualitative and need to be translated into the
295 qualitative values of the model. A link called converters was created between the quantitative or qualitative
296 variables of a user and the input variable of the model. Converters need for a set of input variables (Table 4)
297 that include cropping practices, site of the landscape level, and weather dataset of the cropping season (daily
298 temperature, rainfall and evapotranspiration). These converters are composed of algorithm that use
299 quantitative or qualitative models. These algorithms are based on literature and expertise Table 4.

300

301

IPSIM's Input variable	Value of IPSIM's input variable	Model used in converter	Input variable(s) needed in the converter	Unit
Importance of phomopsis stem canker risk at landscape	High, Intermediate, Low	-	Department of production	-
Weather for infection	Unfavorable, favorable	SimMat-*Diaporthe*	Escape strategy's output	-
			Sowing date of sunflower	Day
			Cultivar name	-
			Irrigation (date and quantity)	Day and mm
			Daily mean temperature	°C
			Daily rainfall	mm
Weather for girdling symptoms	Unfavorable, Moderately favorable, Favorable	SimMat-*Diaporthe*	Weather for Infection's output	-
			Sowing date of sunflower	Day
			Cultivar name	-
			Irrigation (date and quantity)	Day
			Daily mean temperature	°C
			Daily rainfall	mm

Escape strategy	Inefficient, Efficient	SimMat-*Diaporthe*	Sowing date of sunflower	Day	
			Cultivar name	cultivars are referenced at http://www.myvar.fr	
			Irrigation (date and quantity	day and mm	
			Daily mean temperature	°C	
			Daily rainfall	mm	
Crop management	Unfavorable,	CART-*LAI* model	Cultivar name	-	
	Moderately favorable, favorable		Sowing date	Day	
			Crop density	plants.m^{-2}	
			Soil depth	mm	
			Soil N content at sowing	u	
			N-fertilization before anthesis	u	
			daily mean temperature	°C	
			daily rainfall	mm	
			irrigation (date and quantity)	mm	
			ETPP	mm	

304

Genetic control	susceptible and few susceptible, Very less susceptible and resistant	Cultivar name	-
Chemical control	No treatment, Treatment	Treatment number (0 or 1)	-

Table 4 IPSIM 's input variable, their values, information needed for the converter, and their unit

The life cycle of sunflower is simulated using the sum of temperature. Each stage is associated to a specific sum of temperature (for a base temperature of 6°C, Table 5), and a Julian day by the converter. Five maturity groups are proposed in the converter for the anthesis and the maturity stage: very early, early, mid early, mid late and late. This converter needs the sowing date and a daily main temperature for the chosen cultivar.

Phenological stage	Sum of temperatures (°C)				
	VE	E	ME	ML	L
Germination	90	90	90	90	90
Vegetative growth	121	121	121	121	121
Flower bud	482	482	482	482	482
Anthesis	750	790	840	900	950
Maturity	1500	1570	1640	1700	1770

Table 5 Sum of temperature required to reach the phenological stages. There are different precocities for anthesis and maturity. VE: Very Early; E: Early; ME: Mid Early; ML: Mid Late; L: Late

2.2.1 Converter of the escape strategy, the infection and the progression of symptoms on stem

This converter simulates the ascospore releases, infection and progression of symptoms on stem. A farmer can adapt his escape strategy by choosing a sowing date and a cultivar characterized by different flowering and maturity time. We considered that the escape strategy (Equation 1) is determined by the presence of significant ascospore releases for two successive days during flower bud stage of sunflower. ASPHODEL efficiently predicts ascospore releases (Desanlis *et al.*, 2013), this model could be used to our objective, but it needs hourly data. Consequently, we chose SimMat-*Diaporthe* (Vedy-Zecchini *et al.*, 2019 submitted), which needs daily mean temperatures and rainfall data. SimMat-*Diaporthe* simulates the daily rates of mature pseudothecia and that releasing their ascospores. Therefore, ASPHODEL was used to calibrate the ascospore release output of SimMat-*Diaporthe*. A protocol was developed to determine the threshold of

340 the rate of ascospore release predicted by SimMat-*Diaporthe* from which a release could be considered
341 significant for infection to occur. An ascospore release is considered as significant if the ascospore release
342 rate $R_{MFP}(d)$ is higher than a given threshold $R_{MFP}^{\min}$(d).

343
$$\begin{cases} ES = "Inefficient" \; if \; \sum_{d=S}^{d=E} \sum_{j=0}^{j=d-1} \delta_{rs}(j) \geq 2 \\ else \\ ES = "Efficient" \end{cases} \quad (1)$$

344 Where S is the start day of flower bud stage and E is the end day of vegetative stage; *j* is a daily summation
345 index; δ_{rs} (j) is the Kronecker symbol where:

346
$$\begin{cases} r = s \; if \; RMFP(d) \geq R_{MFP}^{\min} \\ and \\ r \neq s \; otherwise \end{cases} \quad (2)$$

347 The parameter $R_{MFP}^{\min}$ was estimated by using an algorithm comparing the SimMat-*Diaporthe*'s and
348 ASPHODEL's outputs using a dataset and the ascospore release variables of the two models. The dataset is
349 composed of 3301 sites/years from 1984 to 2018 across sunflower growing areas in France. For each
350 sites/years, the period considered extended from the first projection predicted by ASPHODEL to 31 July.
351 There were in all 182407 days. Both hourly and daily weather variables were used for ASPHODEL while daily
352 weather variables were used for SimMat-*Diaporthe* (mean hourly temperature and hygrometry, daily
353 minimal, maximal and mean temperature and rainfall). The mathematical formalism of ASPHODEL is
354 presented by Vedy *et al.*, 2019 (submitted). Ascospore release begin when the pseudothecia maturation
355 percentage reach 50\%. The ascospore releases predictions of ASPHODEL is a daily binary variable
356 (presence/absence of ascospore releases). Each absence and presence of ascospore release was considered.
357 Prediction of SimMat-*Diaporthe* ($R_{MFP}^{\min}$) is quantitative and therefore an algorithm was designed to translate
358 the quantitative variable into a binary one. When the projection rate reached at least a given threshold, the
359 projection was present. One thousand thresholds of $R_{MFP}^{\min}$ from 0.0 to 1.0 with a step of 0.001 were tested.
360 The ascospore release predictions of ASPHODEL and SimMat-*Diaporthe* were compared using the Receiver
361 Operator Characteristic (ROC) curve technique Fawcett, 2006. This curve is used to determine threshold

362 implicated into binary problematic. Each threshold represents a point of this curve and is determined by its
363 ordinate, the true positive rate (proportion of ascospore release presence correctly predicted, *TPR* Equation
364 3) and its abscissa, the false positive rate (proportion of ascospore release absent predicted as present, *FPR*
365 Equation 4).

366 $$TPR = \frac{TP}{TP+FN} \quad (3)$$

367 Where *TP* is the number of True Positive and *FN*, the number of False Negative.

368 $$FPR = 1 - \frac{FP}{FP+TN} \quad (4)$$

369 Where *FP* is the number of False Positive and *TN*, the number of True Negative.

370 The threshold retained is the one with the highest true positive rate and the lowest false positive rate. On
371 the ROC curve, this is the point that is closest to the coordinate point (0.1). The threshold retained was
372 included into the converter designed to compute significant ascospore release.

373

374 After detection of two successive significant ascospore releases, another algorithm compute the success
375 of an infection and symptom progression.

376 An infection (denoted *I*, Equation 5), can occur (I = "Unfavorable") until ten days after two successive
377 significant ascospore releases (Bertrand and Tourvieille, 1987; Fabregue and Eychenne, 1998). An infection is
378 successful if leaves are wetted during at least 48 hours (Yang *et al.*, 1987; Mihaljcevic *et al.*, 1985; Peres and
379 Régnault, 1986). We supposed three successive days were sufficient to reach the quantity of rainfall (R_2^{min})
380 necessary for the success of an infection.

381 $$\begin{cases} I = "Unfavorable" \; if \; R(d_i) \geq R_2^{min} \\ else \\ I = "Favorable" \end{cases} \quad (5)$$

382 Where $d_{i0} \leq d_i \leq d_{i10}$, d_{i0} is the julian day of the second significant ascospore releases and d_{i10} is the 10th julian
383 day after the second significant ascospore releases; and R_2^{min} = 20 mm (Debaeke and Estragnat, 2003;

384 Debaeke and Moinard, 2010). The value of R_2^{min} was not the subject of these two previous studies, it is a
385 parameter that can vary and we proposed an algorithm to adjust it. In total, 2230 values (from 0.0 to
386 223.0mm) were tested. We selected the value of R_2^{min} for which the predicted injuries were as close as
387 possible to the observed injuries.

388 From an infection, the fungal mycelium can enter in a latent phase until the beginning of flower bud stage
389 (Penaud *et al.*, 1995). It has been demonstrated that in greenhouses the mycelium can infect the stem 40
390 days after the contamination (Penaud *et al.*, 1995). The progression of the mycelium is possible as long as the
391 sunflower's leaves are not senescent or there are not unusual warm temperatures during four successive
392 days after infection (Table 1). The fungal mycelium can cause girdling symptoms 10 days after reaching the
393 stem. The girdling symptoms denoted GS is given Equation 4.

394
$$\begin{cases} GS = \text{"Unfavorable"}\ if\ \sum_{k=d_i}^{k=d_{i40}} R(k) \geq R_3^{min}\ and\ \sum_{k=d_i}^{k=d_{i37}} \sum_{m=0}^{m=k+3} \delta_{rs}(m) < 3 \\ GS = \text{"M. f."}\ if\ R_4^{min} \leq \sum_{k=d_i}^{k=d_{i40}} R(k) < R_3^{min}\ and\ \sum_{k=d_i}^{k=d_{i37}} \sum_{m=0}^{m=k+3} \delta_{rs}(m) < 3 \\ else \\ GS = "Favorable" \end{cases} \quad (6)$$

395 Where d_i the day of I., d_{i37} the 37th day after I., d_{i40} the 40th day after I.; R_3^{min}= 75 mm, R_4^{min}= 20 mm
396 (Debaeke and Moinard, 2010).
397 Where *m* is a daily summation index; *k* is the considered day; $\delta_{rs}(m)$ is the Kronecker symbol with:

398
$$\begin{cases} r = s\ if\ \theta(m) \geq \theta^{max} \\ and \\ r \neq s\ otherwise \end{cases}$$

399 Where $\theta^{max} = 33$ °C (Fabregue, 1991).

400

401 **2.2.2 Converter of the variable ``Risk level in the department''**

402 A map of phomopsis stem canker frequency in France is regularly updated using surveys on cultural
403 practices (www.terresinovia.fr). Phomopsis stem canker is considered as "Frequent", "Moderately
404 frequent", "Infrequent" and "Undetermined" in 17, 32, 21 and 32 departments, respectively. The user has

405 to fill in the name of his department in the converter that gives the correspondence with the input variable
406 of the model. If the phomopsis stem canker is frequent, the input value is “H”, if phomopsis stem canker is
407 moderately frequent, the input value is “I” and “L” otherwise.
408

409 **2.2.3 Converter of the variable ``Crop management''**

410 The period the most receptive to phomopsis stem canker is the flower bud stage (Délos *et al.*, 1995;
411 Penaud *et al.*, 1995; Moinard and Eychenne, 1998) when all leaves are developed and vegetation cover closes.
412 Debaeke and Estragnat (2003) showed that *LAI* at the anthesis was a good indicator characterising the
413 sunflower canopy development. This indicator is not easily accessible but it is a synthesis of crop density, soil
414 depth, nitrogen and water stresses experienced by the crop during its vegetative stage (Debaeke *et al.*, 2000).
415 We arbitrarily considered that *LAI* can be classified into three categories: ”Unfavorable”, “Intermediate” and
416 “Favorable” depending of *LAI* thresholds. These thresholds were identified using an algorithm and a dataset.
417 All the values from 1.1 to 4.0 with a step of 0.1 were tested for the first threshold and all the values from x +
418 0.1 to 5.0 (with x the value of the first threshold) were tested for each value of the first threshold. For each
419 combination, a Classification And Regression Trees (CART, Breiman *et al.*, 1984) model (package {rpartScore}
420 using R) were considered to predict *LAI* categories as a function of abiotic input variables. CART designs
421 decision tree by learning using a dataset. CART selected variables and values that best discriminated *LAI*’s
422 categories of the dataset. The adjustment quality of the CART model designed is then evaluated on the same
423 dataset. The model with the highest quality of adjustment was identified and selected. The input and output
424 variables are presented in (Table 6). The dataset used (5167 sites/years) was generated by SUNFLO
425 Casadebaig *et al.*, 2011, a model that predicts the final yield as a function of abiotic stresses. The distribution
426 of *LAI* related to this dataset is described in Figure 2.
427
428

Symbol	Definition	Formula	Phenological period	Values	Units	Type
CD	Plant density	-	sowing	4 and 6	plants.m^2	Input
Ncredit	Soil N content at sowing	-	sowing	30, 60 and 90	u	Input
Nferti	N-fertilization	-	vegetative	0, 40 and 80	u	Input
RETPP	Water-stress status	$\sum_{i=d_G}^{i=d_v} R(i) - ETP(i)$	vegetative	[-291.7; -3.2]	mm	Input
SD	Soil depth	-	-	500 and 1500	mm	Input
LAI-F1	*LAI* at anthesis	-	anthesis	1 (Favorable to sunflower), 2 (Moderately favorable) 3 (Unfavorable to sunflower)	-	Output

Table 6 Symbol, definition, phenological period, values and units used in CART model predicting LAI's categories. With d_G the julian day of germination, dv the julian day of the end of vegetative stage, R(i) the daily sum of rainfall plus irrigation and ETP(i) the daily evapo-transpiration

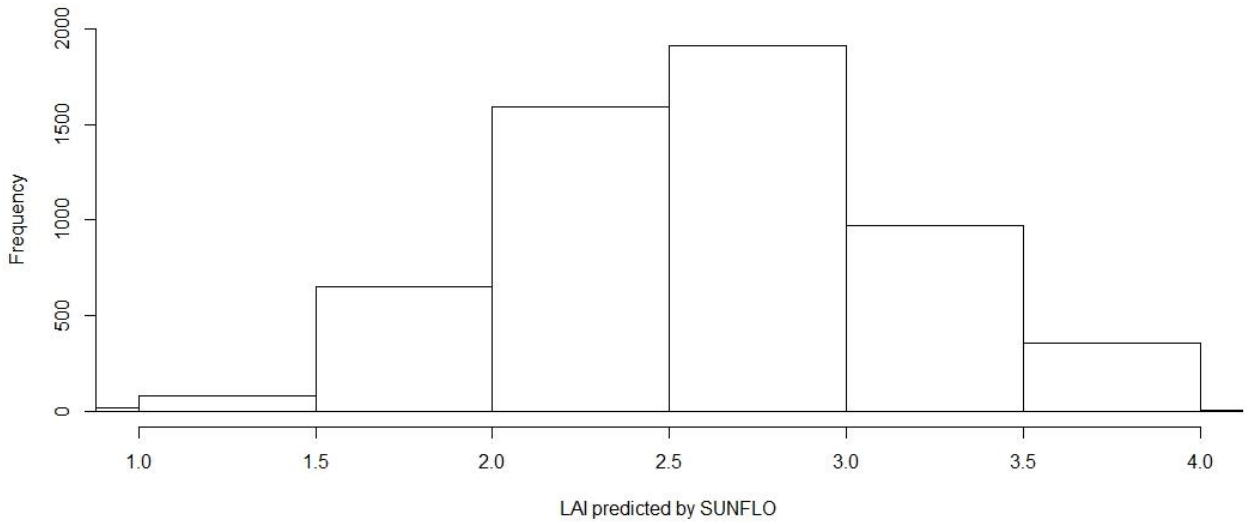


Figure 2 Distribution of *LAI* observed on the SUNFLO dataset (5616 situation).

2.2.4 Converter of the variable ``Genetic control''

The converter proposes 714 cultivars of sunflower (from Terres Inovia, data not shown). They are characterized by their susceptibility to several diseases. There are 41 very susceptible, 119 susceptible, 224 less susceptible, 219 very less susceptible and 13 resistant. The cultivars very susceptible, susceptible and less susceptible are considered as “susceptible and few susceptible” in the converter, they are “very less susceptible and resistant” otherwise. The input variables of the model and that of the converter they depend are described in (Table 4).

2.3 Evaluation of predictive quality of IPSIM-Sunflower-Phomopsis stem canker, SimMat-*Diaporthe* and the CART-*LAI* model

2.3.1 Description of the dataset used to evaluate the predictive quality of IPSIM-Sunflower-Phomopsis stem canker and the CART-*LAI* model

A dataset of 81 sites/years was used to evaluate the model. This dataset was created by the Centre Technique Interprofessionnel des Oléagineux en Métropole in France and the Institut National de la

447 Recherche Agronomique of Toulouse to analyse the impact of environment and cropping practices on
448 incidence and severity of different pests (Table 7). The input variables of converters described previously
449 were observed on one site/year in Allier (03), in Charente (16), in Côte-d'Or (21), in Deux-Sèvres (79), in Indre
450 (36), in Marne (51), in Moselle (57), in Puy-de-Dôme (63), in Saône et Loire (71), in Seine-et-Marne (77), in
451 Vendée (85) and in Yonne (89), 2 in Jura (39), in Meuse (55) and in Rhône (69); 6 in Charente-Maritime (17),
452 6 in Cher (18), 3 in Gers (32), in Meurthe-et-Moselle (54), in Nièvre (58) and 3 in Vienne (86), 5 in Lot-et-
453 Garonne (47) and 34 in Haute-Garonne (31), between 1994 and 2018. In all sites, spots on stem incidence
454 was observed for 1029 on the 1074 situations. These spots included girdling symptoms and little spot on stem
455 that evolved into girdling symptoms after observations. There were 478 situations observed as "Low", 186
456 observed as "Intermediate" and 365 as "High".

457 The *LAI* at anthesis was observed for 771 out of the 1074 previous situations Table 7. This part of the dataset
458 was used to evaluate the CART-*LAI* model. The distribution is described in Figure 3. The *LAI* values of this
459 dataset were classified into three classes “Unfavorable”, “Intermediate” and “Favorable”. This classification
460 was defined using the thresholds selected by the algorithm that designed the CART-*LAI* model.

461 There were 300 situations without any measure of *LAI* at anthesis. This has generated an uncertainty in the
462 input variable “Crop Management” as well as in the outputs of the model. We considered three types of
463 evaluations: i) an evaluation with all the data, where the 300 missing *LAI* values were predicted by the CART-
464 *LAI* model; ii) an evaluation with only the 729 situations with a *LAI* value and iii) an evaluation with only the
465 300 situations without a *LAI* value (Table 8).

466

467

468

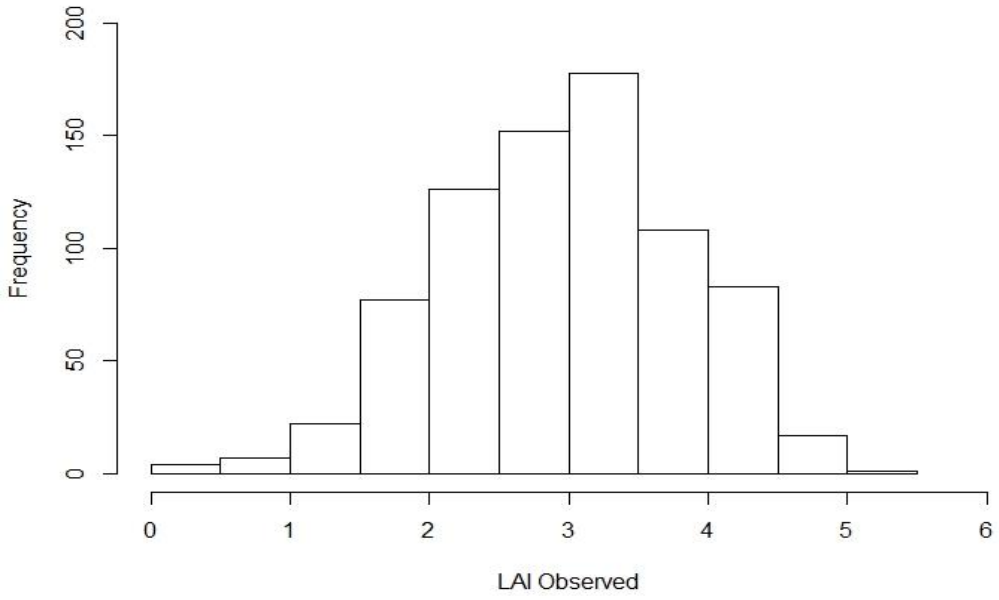
469

Region	Dep.	Location	Year	Number of site-years	Cropping practices	Values of spot on stem observed	Number of *LAI* values	Ref.
Auvergne-Rhône-Alpes	69	Lyon St Exupéry	2009-2013	2	sowing date, crop density, soil depth, N-fertilisation, use of fungicide and weather	Low: 4	-	Terres Inovia
Bourgogne-Franche-Comté	21	Chatillon	2013	1	crop density	Low: 2	2	Terres Inovia
	39	Chemin			crop density, use of fungicide	Low: 2	2	Terres Inovia
	71	Champforgeuil			use of fungicide	Intermediate: 1, High: 1	-	Terres Inovia
Centre-Val de Loire	18	Saint-Florent	2003-2011	6	sowing date, crop density, soil depth, crop status at sowing, N-fertilisation, cultivar, use of fungicide and weather	Low: 18	13	Terres Inovia
Grand Est	54	Doncourt-les-Conflans	2015	1	fungicide	Low: 2	-	Terres Inovia
	57	Dieuze	2014	1	Fungicide	Low: 2	-	Terres Inovia

Nouvelle-Aquitaine	16	Chalais	2002	1	sowing date, crop density, N-fertilisation, cultivar and weather	Low: 4	4	Terres Inovia
	17	La Rochelle, Le Magneraud, Nuaillé-sur-Boutonne	2002-2007	10	sowing date, crop density, soil depth, crop status at sowing, N-fertilisation, cultivar, use of fungicide and weather	Low: 52	21	Terres Inovia
	47	Agen-Estillac, Nérac	2002-2008	4	sowing date, crop density, soil depth, crop status at sowing, N-fertilisation, cultivar, tillage, use of fungicide and weather	Low: 18, Intermediate: 4, High: 1	19	Terres Inovia
	79	Melle, Thure	2002	2	sowing date, crop density, N-fertilisation and cultivar and use of fungicide	Low: 6	6	Terres Inovia
	86	Loudun	2002-2003	2	sowing date, crop density, N-fertilisation, cultivar, use of fungicide and weather	Low: 2, Intermediate: 1, High: 4	5	Terres Inovia

Occitanie	31	Auzeville, En-Crambade	1994-2018	20	sowing date, crop density, soil depth, crop status at sowing, N-fertilisation, cultivar, use of fungicide and weather	Low: 331, Intermediate: 163, High: 166	632	Terres Inovia and INRAE
	32	Auch, Condom	2005-2013	4	sowing date, crop density, soil depth, crop status at sowing, N-fertilisation, cultivar, use of fungicide and weather	Low: 45	10	Terres Inovia
Pays de la Loire	85	Saint-Gemme	2002	1	sowing date, crop density, crop status at sowing, N-fertilisation	Low: 3, Intermediate: 1	3	Terres Inovia

Table 7 A representative summary of the dataset used for model evaluation. The observed variables were the following: sowing date, crop density, soil depth, crop status at sowing, previous crop, fertilisation, cultivar, crop sequence, tillage, fungicide and weather. The variables with different values are indicated in third column. Location of fields (03: Allier, 16: Charente, 17: Charente-Maritime, 18: Cher, 21: Côte-d'Or, 79: Deux-Sèvres, 32: Gers, 31: Haute-Garonne, 36: Indre, 39: Jura, 47: Lot-et-Garonne, 51: Marne, 54: Meurthe-et-Moselle, 55: Meuse, 57: Moselle, 58: Nièvre, 63: Puy-de-Dôme, 69: Rhône, 71: Saône et Loire, 77: Seine-et-Marne, 85: Vendée, 86: Vienne and 89: Yonne); Year; Number of site-years; Cropping practices; Values of spot on stem observed; Number of LAI values; Ref. All: no variations. L: Low, I: Intermediate, H: High. Dep.: Department. Ref.: Reference.



473

Figure 3 Distribution of *LAI* observed on the 772 situations from the dataset used to evaluate IPSIM-Sunflower-*Diaporthe*.

474

Scenario	Nature of scenario	Number of situation	Use of CART-*LAI* model
1	Observed and predicted *LAI*	1029	Yes, on situation without *LAI* value
2	Observed LAI	729	No
3	Predicted LAI	300	Yes

Table 8 Description, use of CART-*LAI* model for each scenario

475

476 The SAFRAN database (Météo France) was used to extract mean daily and hourly temperature and rainfall;
477 along with hourly relative humidity for each site presented in the sections 2.2.1 and 2.2.3. This national
478 database uses a 8 x 8 km^2 square grid pattern. One square was used per site and in the case of a site inside
479 more than one square, the square considered was the one with weather variables closest to the sample of
480 observed weather variables.

481

2.3.2 Criteria used for qualitative predictions

The outputs of model and converter are qualitative ordinal. The agreements between their predictions and the observations were evaluated using confusion matrices (table representing the distribution of qualitative prediction in function of observation, they represent the agreement between predictions and observations), the precision and the squared weighted κ of Cohen. Its interpretation was realized using metric of Landis and Koch (Landis and Koch, 1977).

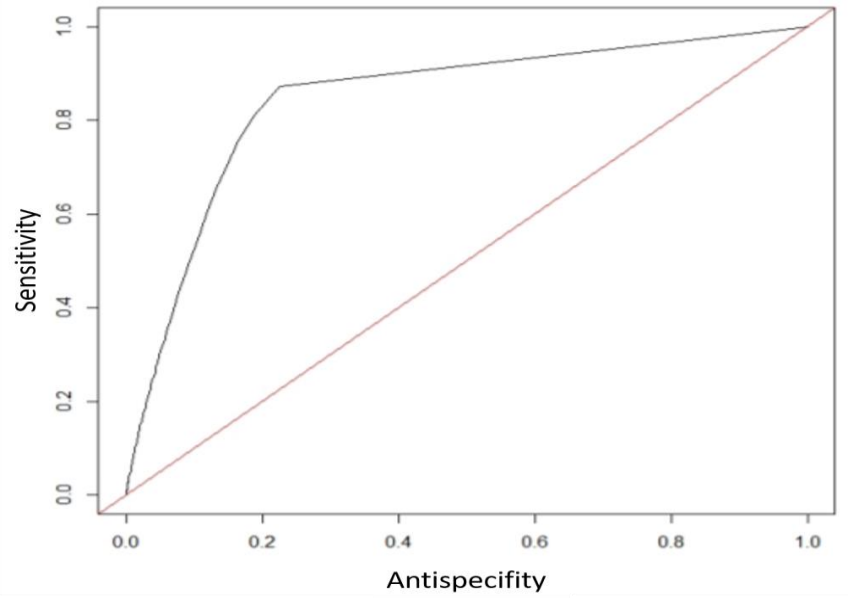
2.3.3 Computer implementation

IPSIM-Sunflower-Phomopsis stem canker was designed using the DEXi software tool (Bohanec *et al.*, 2013). The converters were implemented in VBA (Microsoft ® Office Excel). The evaluation was realized using R (https://www.R-project.org). The irr and vcd packages were used to compute the squared weighted kappa of Cohen's and the stats package was used to perform the Wilcoxon test.

3. Results

3.1. Thresholding of ascospore releases rates of SimMat-*Diaporthe* and necessary rainfall for infection

The trade-off between sensitivity and antispecificity was obtained when the true positive rate was equal to 0.872 and the false negative rate was equal to 0.226. The threshold of ascospore release rate was equal to 0.001 (Figure 4).



500

Figure 4 ROC curve with sensitivity in function of antispecificity (3301 sites/years).

501

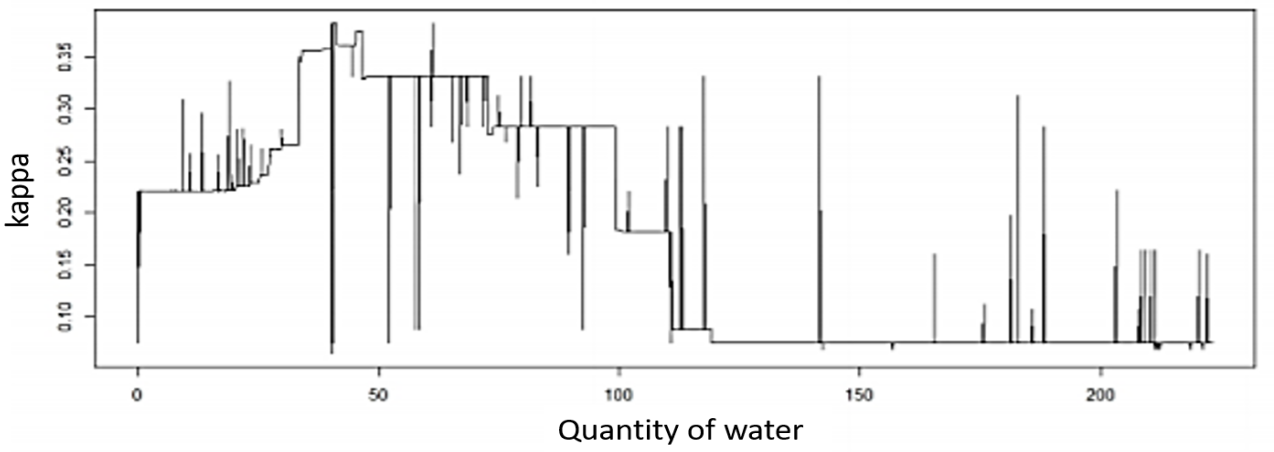
502 The squared weighted Cohen's Kappa computed was κ= 0.464 (Table 9). This indicates a low consistency
503 between SimMat-*Diaporthe*'s and ASPHODEL's predictions. The associated precision was P= 0.79. This
504 moderate kappa and the good precision result from the high proportion (77%) of absence and presence (88%)
505 of ascospore release correctly predicted by SimMat-*Diaporthe*.

		ASPHODEL predictions		
		Absence	Presence	Total
SimMat-*Diaporthe* predictions	Absence	116777	3845	120622
	Presence	34622	27163	61785
	Total	151399	31008	182407

Table 9 Confusion matrix with predictions of presence/absence of ascospore release of ASPHODEL and SimMat-Diaporthe.

506 The evolution of kappa obtained between observations and predictions as a function of the necessary rainfall
507 for an infection to occur, is shown in Figure 6. The best value of κ (κ= 0,38) was obtain for minimal value of
508 R_2^{min}= 40.4 mm.

509



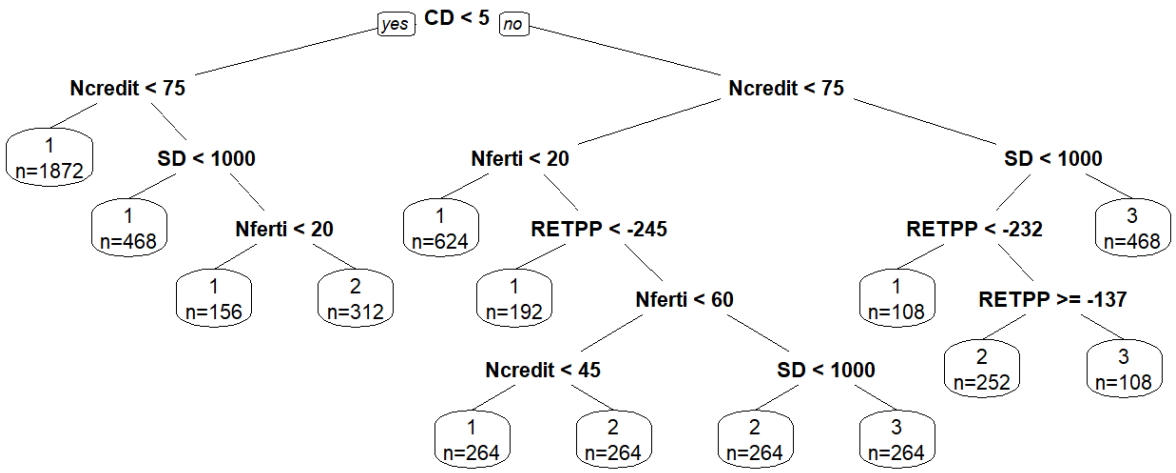
510

Figure 5 kappa between observations and predictions as a function of the necessary rainfall for an infection to occur

511 **3.2. Structure, evaluation of adjustment and prediction qualities of CART-*LAI* model**

512 The thresholds of *LAI* considered by the algorithm that fed the CART model were 2.8 and 3.1. The crop
513 management was considered as favorable if the *LAI* at anthesis was lower than 2.8. It is considered as
514 favorable, intermediate and unfavorable if the LAI at anthesis was <2.8, >2.8 <3.1, and >3.1, respectively. The
515 CART-*LAI* model is presented Figure 6.

516



517

Figure 6 An overview of the CART-*LAI* model (5167 sites/years). Legend: *CD*= Crop Density, *Ncredit*= Soil N content at sowing, *SD*= Soil Depth, *Nferti*= N fertilisation, *RETPP*= irrigation and rainfall minus potential evapotranspiration (Table 6). 1= favorable, 2= Intermediate and 3= Unfavorable, 1000 is in mm

518

519 The confusion matrix of the adjustment quality of CART-*LAI* model on SUNFLO dataset and that of the
520 predictive quality on observed *LAI* are presented Table 11 A and B. The adjustment quality on SUNFLO dataset
521 was high (κ= 0.74) while its predictive quality on observed LAI was low (κ= 0.129 and P= 0.47). These results
522 mainly from the 115 unfavorable *LAI* predicted as favorable and the 150 favorable *LAI* predicted as
523 unfavorable.

524

525

A		Predicted precocity			
		Fav.	Inter.	Unfav.	Total
Observed precocity	Fav.	3150	352	46	3548
	Inter.	470	378	140	988
	Unfav.	64	362	654	1080
	Total	3684	1095	840	5619

526

B		Predicted precocity			
		Fav.	Inter.	Unfav.	Total
Observed precocity	Fav.	149	29	150	328
	Inter.	40	7	52	99
	Unfav.	115	25	208	348
	Total	304	61	410	775

Table 10 Confusion matrix with A: classes of *LAI* simulated by SUNFLO and predicted by CART-*LAI* model (5167 situations), B: classes of *LAI* observed and predicted by CART-*LAI* model (775 situations).

Fav.= Favorable, Inter.= Intermediate and Unfav.= Unfavorable

527

528 Despite of its performance on observed *LAI*, the CART-*LAI* model had a very good performance on SUNFLO
529 dataset and was introduced into converter.
530
531 **3.3. Evaluation of the quality of adjustment for girdling incidence classes and total spot on stem**
532 The confusion matrices from all three evaluations are presented Table 11. Their associated kappa and the
533 precisions for each scenario are described in Table 12.
534 The associated precision of the first scenario (with the entire dataset) was *P*= 0.52 (Table 11 A and 12). The
535 squared weighted Cohen's Kappa computed on the 1029 situations was κ= 0.366, which is the highest value
536 of κ among all scenarios. This fair κ resulted from 138 situations observed as high and predicted as low, the
537 118 situations observed as high and predicted as moderate, the 109 situations observed as low and predicted
538 as moderate and the 106 situations observed as low and predicted as moderate.
539 The precision of the second scenario was *P*= 0.46 (Table 11B and 12). The squared weighted Cohen's Kappa
540 computed on the 729 situations was κ= 0.342, which is a fair agreement between observations of total spot
541 on stem and predictions.
542 The precision of the third scenario was *P*= 0.67 (Table 11C and 12) while the squared weighted Cohen's Kappa
543 computed on the 300 situations was κ= 0.0175.
544 These indicate fair agreements between converter's predictions and observations. These consistencies
545 results from the numerous sites/years that were very bad predicted (unfavorable instead of favorable for
546 example), and represents the worst scenario.
547

A

		Predicted precocity			
		Low	**Moderate**	**High**	**Total**
Observed precocity	**Low**	362	106	10	478
	Moderate	109	63	14	186
	High	138	118	109	365
	Total	609	287	133	1029

548

B

		Predicted precocity			
		Low	**Moderate**	**High**	**Total**
Observed precocity	**Low**	168	77	8	253
	Moderate	83	57	14	154
	High	102	111	109	322
	Total	353	245	131	729

549

C

		Predicted precocity			
		Low	**Moderate**	**High**	**Total**
Observed precocity	**Low**	194	29	2	225
	Moderate	26	6	0	32
	High	36	7	0	43
	Total	256	42	2	300

Table 11 Confusion matrix between observations and predictions with A: the first scenario; B: the second scenario; C: the last scenario Table 8.

550

Scenario	**κ**	***P***
1	0.366	0.52
2	0.342	0.46
3	0.0175	0.67

Table 12 kappa and precision computed for each scenario (Table 8).

551

552 **4. Discussion**

553 **4.1 Thresholding and evaluation of predictive quality of converter for infection and development of**
554 **symptoms**

555 The results of SimMat-*Diaporthe* were close to those of ASPHODEL for an ascospore release rate of 0.001.
556 Each ascospore release was considered as significant. The pseudothecia maturation is included into the
557 ascospore release rate so we considered that this parameter did not need a threshold. To the best of our
558 knowledge, there were no reports of equivalence of this type of rates in other models simulating the
559 reproduction of ascomycota.

560 **4.2 Consistency and performances of CART-*LAI* model**

561 Although the thresholds of *LAI* (*LAI*= 2.8 and 3.1) considered by the algorithm to discriminate the classes
562 of *LAI* seemed a little high, they were consistent with thresholds obtained in scientific (*LAI*= 2.4 and 3.3;
563 Debaeke and Estragnat, 2003) and technical (*LAI*= 2.5; Taverne, 2005) literature. The CART-*LAI* model took
564 into account all the proposed variables highlighting their relevance. Similarly, the value of the thresholds of
565 input variables were consistent with the literature: the critical value of density employed by the algorithm
566 was 5 plants.m^{-2} and it is advised to avoid crop density higher than 6.5 plants.m^{-2} (Debaeke *et al.*, 2000,
567 Debaeke and Estragnat, 2003, Debaeke and Moinard, 2010). The values of threshold considered by the
568 algorithm for soil depth and fertilisation correspond to the orders of known magnitude of values that are
569 usually advised to farmers (Lartigot, 2015). Two thresholds were taken into account for the soil N content at
570 sowing (45 and 75 u), these values correspond to low and high values that could be measured under field
571 conditions (Lartigot, 2015). Our dataset was not designed for this study and, consequently, soil N content at
572 sowing was difficult to provide (creating the uncertainty in the *LAI* variable). Therefore, the future users of
573 IPSIM-Sunflower-Phomopsis stem canker will be advised to obtain this variable with a tool designed by farm
574 advisers. This tool need information such as the previous crop, the yield, and the soil depth. CART-*LAI* model
575 was very good to classify *LAI* at anthesis with SUNFLO dataset (κ= 0.74). Nevertheless, this model was not
576 that precise to predict the observed values of *LAI* at anthesis (κ= 0.129 and *P*= 0.47) given that one third of
577 observed LAI were correctly predicted and it predicted LAI either lower or higher than what was observed.
578 The good result on SUNFLO dataset could be explained by the large dataset (5619 sites-years) although the
579 dataset to evaluate the predictive quality was smaller and unbalanced. Indeed, 53%, 8% and 39% of
580 observations with *LAI* values were favorable, intermediate, and unfavorable, respectively. This was due to a
581 difference in the collection and formatting of the dataset.

582

583

584

585 **4.3 Evaluations of converters and IPSIM-Sunflower-Phomopsis stem canker**

586 As expected, predictions quality was the best for the scenario with the entire dataset (κ= 0.366 and *P*=
587 0.52) than for other scenarios with smaller dataset (κ= 0.342 and 0.0175 and *P*= 0.46 and 0.67). Overall, 35.6%
588 of the observations were under-predicted while 12.6% of the observations were over-predicted, thereby
589 highlighting that this model underestimates the injuries. The lowest quality of evaluation was obtained for
590 the scenario using only a part of observations and the CART-*LAI* model. This was, however, not surprising
591 since the CART-*LAI* model had a low agreement (κ= 0.129). As a consequence, despite a fair result on *LAI*
592 provided by the CART-*LAI* model, this tool still needs further improvement.

593

594 **4.4 Methodological considerations and model improvements**

595 **4.4.1 Improvement of the dataset**

596 The dataset used to evaluate the predictive quality of IPSIM - Sunflower - Phomopsis stem canker and
597 converters was small (81 sites/years) and unbalanced: 46%, 18% and 35% of observations were low,
598 moderate and high, respectively. A uniform dataset would have been useful to perform a consistent model
599 evaluation. Furthermore, most of the dataset (77%, 791 on 1029) originated from Auzeville (31) where
600 phomopsis stem canker was well studied. This is also because the disease was first reported from this region
601 where it is responsible for important economic losses. However, future studies should focus on a detailed
602 field survey across other French regions to improve the uniformity and representativeness of the dataset on
603 a national scale.

604

605 **4.4.2 Improvement of the model**

606 The IPSIM model was chosen as this emerges a simple and innovative approach (Aubertot and Robin,
607 2013). Moreover, this model has been successfully applied to other pathosystems including IPSIM-Wheat-

608 Eyespot (Robin *et al.,* 2013), on IPSIM-Wheat-Brown-Rust (κ= 0.68, Robin and Aubertot, 2015) and soon on
609 IPSIM-Chayotte-Flies (κ >0.90, Deguine, com. pers.).

610 Although not currently considered in its present form, the model structure could evolve to integrate
611 highlighting the flexibility of this modeling framework.Several information reported in scientific literature
612 have not been included included into IPSIM - Sunflower - Phomopsis due to poor data availability. For
613 example, it seems that K value affects severity of phomopsis stem canker. An increase of 80 u of K decreases
614 the *D. helianthi* infection possibility by 69.6% (Iliescu and Baicu, 1984). The variable soil has been proposed
615 in the generic IPSIM structure (Aubertot and Robin, 2013) but no information is available to date to feed the
616 model with this information. For instance, the effect of soil depth and type on survival ability of the pathogen
617 in infested sunflower stubble and its capacity to produce pseudothecia is still poorly known.

618 Acimovic and Straser (1981) attempted to study in this regard by burying stubble at different depths in
619 different soil types but the results were inconsistent. Likewise, no study yet focused on biological interactions
620 of *D. helianthi* with other organisms (either beneficial or harmful) present in the field. Future studies are
621 needed to fill this gap that may result useful to improve the prediction quality of the model. As for the model
622 structure, aggregative rules can also be modified, by identifying different values of aggregative rules into the
623 IPSIM aggregative tables that finally allow to optimise the predictive quality of IPSIM. An algorithm will be
624 proposed to this objective, based on an adaptation of the method used to optimize the prediction quality of
625 quantitative models (Wallach *et al.*, 2006). The aggregative rule will be considered as parameters to modify
626 into this future original algorithm.
627 Finally, the converter can be also modified by taking into account regular updates and potential evolutions
628 of *D. helianthi* in France. Terres Inovia regularly updates the list of sunflower cultivars with their sensitivity
629 level to the disease, anthesis and maturity precocities as well as maps representing the frequency of
630 phomopsis stem canker girdling symptoms in each French region. The converter can then be fitted to take
631 into account the weather effect on pathogen infection based on observations. To this aim, the value of R^{min}
632 (40.4 mm) found here was consistent with the reported experimental value (R^{min}= 20 mm, Debaeke and

633 Estragnat, 2003; Debaeke and Moinard, 2010). However, the converter simulating the weather effects was
634 not fitted for the other parameters (weather for the expression of symptoms of stem) due to poor data
635 availability in literature. Future studies thus should aim to generate these data through experiments that will
636 finally allow to strengthen the converter.

637

638 **4.5 Conclusion**

639 Although the predictive quality of IPSIM-Sunflower-Phomopsis stem canker needs further improvements,
640 this model in its current form represents a summary knowledge of literature and expertise regarding the
641 effect of the production situation and cropping practices on phomopsis stem canker. This is the first IPSIM
642 model designed for the sunflower and phomopsis stem canker, using a complete converter, including other
643 qualitative and quantitative models. The key advantage of this model is that it includes a wide range of
644 information available to date to evaluate prediction quality of the injury of phomopsis stem canker. This
645 model with its converter is simple and its interface is user-friendly. Currently, the model is implemented
646 under a common langage (VBA) although it can be implemented in computerized decision support system
647 for Phomopsis stem canker Management used by adviser services. The reliability of this model must be
648 proven before a routine use. The adviser will have to use this model before sowing to assess whether a given
649 cropping practice can pose higher risks in a given production situation. In addition, this model can be used to
650 improve cropping practices by choosing those more sustainable for the environment. IPSIM models designed
651 for individual pests of a given crop can theoretically be combined to predict the injury profile caused by their
652 interactions (for example, phoma stem canker that can halt the development of phomopsis disease), as a
653 function of cropping practices and production situation (Aubertot and Robin, 2013). In the long term, IPSIM-
654 Sunflower-Phomopsis stem canker will be included into a global model IPSIM-Sunflower merging IPSIM
655 models developed on individual fungal pathogens.

656

Acknowledgements

We thank Benoît Persyn from the Agroclim unit (INRAE, Avignon) and Météo France for providing the weather dataset used in this study. This work was supported by Terres Inovia, the Occitanie region and INRAE through the doctoral research fellowship of the first author.

References

Acimovic M., Straser N. 2003. Phomopsis sp. A newAgronomie 23 (7): 581–592.

Aubertot J.N., Robin M.H. 2013. Injury Profile SIMulator, a qualitative aggregative modelling framework to predict crop injury profile as a function of cropping practices, and the abiotic and biotic environment. I. Conceptual bases. PLoS ONE 8(9), DOI 10.1371/journal.pone.0073202

Aubertot J.N., Deguine J.P., Lamichhane J.R., Robin M.H., Sarthou J.P., Steinberg C. 2019. Agroécosystème semences / lit de semences : les bases d'une protection agroécologique des cultures en phase d'implantation. In: Boiffin, J., Laurent F. and Richard, G. (eds) Implantation des cultures, quae edn, Versailles, France

Bertrand F., Tourvieille D. 1987. Phomopsis tournesol : Tests de selection. Information Techniques CETIOM 98(I) :12-18

Blanchet R., Gelfi N., Puech J. 1987. Alimentation azotée, surface foliaire et formation du rendement du tournesol (*Helianthus annuus L.*). Agrochimica 31:233–244

Bohanec M., Znidarzic M., Rajkovic V., Bratko I., Zupan B. 2013 DEX methodology: Three decades of qualitative multi-attribute modeling. Informatica (Slovenia) 37(1):49–54

Breiman L., Friedman J., Stone C., Olshen R. 1984 Classification and Regression Trees. The Wadsworth and Brooks-Cole statistics-probability series, Taylor & Francis

679 Breman H., De Wit C. 1983. Rangeland productivity and exploitation in the Sahel. Science 221:1341–1347

680 Brisson N., Mary B., Ripoche D., Jeuffroy M.H., Ruget F., Nicoullaud B., Gate P., Devienne-Barret F.,
681 Antonioletti R., Durr C., H élène M., Françoise J., Nicoullaud R.B., Richard G., Beaudoin N., Recous S., Tayot
682 X., Plenet D., Cellier P., Machet J.M., Meynard J.M., Delécolle R. 1998. STICS: a generic model for the
683 simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to
684 wheat and corn. 18(5):311–346

685 Casadebaig P., Guilioni L., Lecoeur J., Christophe A., Champolivier L., Debaeke P. 2011. SUNFLO, a
686 model to simulate genotype-specific performance of the sunflower crop in contrasting environments.
687 Agricultural and Forest Meteorology 151(2):163–178, DOI 10.1016/j.agrformet.2010.09.012

688 Connor D., Hall A. 1997. Sunflower physiology. In: Schneiter, A.A. (Ed.), Sunflower Technology and
689 Production. Agronomy Monograph 35, pp 113–182

690 De Wit C., Penning de Vries W. 1982. L'analyse des systèmes de production primaires. In: Penning de Vries
691 W., Djiteye M., 918 ARR (eds) La productivité des pâturages sahéliens, Pudoc, Wageningen, pp 275–283

692 Debaeke P., Estragnat A. 2003. A simple model to interpret the effects of sunflower crop management on the
693 occurrence and severity of a major fungal disease: Phomopsis stem canker. Field Crops Research
694 83(2):139–155, DOI 10.1016/S0378-4290(03)00071-6

695 Debaeke P., Estragnat A. 2009 Crop canopy indicators for the early prediction of Phomopsis stem canker
696 (*Diaporthe helianthi*) in sunflower. Crop Protection 28(9):792–801, DOI 10.1016/j.cropro.2009.04.011

697 Debaeke P., Moinard J. 2010. Effect of crop management on epidemics of phomopsis stem canker (*Diaporthe*
698 *helianthi*) for susceptible and tolerant sunflower cultivars. Field Crops Research 115(1):50–60, DOI
699 10.1016/J.FCR.2009.10.002

700 Debaeke P., Raffaillac D. 1996. Light interception as an indicator of leaf area index and risk of diseases in
701 sunflower. Helia 19:1–15

Debaeke P., Délos M., Moinard J., Bérault S., Lambert R. 2000. Prise en compte du couvert de tournesol dans la simulation des épidémies de *Diaporthe helianthi* par le modèle Asphodel. In: Proceedings of the Sixth International Conference on Plant Diseases, 68 December 2000, Tours, France, pp 251–258

Debaeke P., Estragnat A., Reau R. 2003. Influence of crop management on sunflower stem canker (*Diaporthe helianthi)*. Agronomie 23 (7): 581–592.

Debaeke P., Bedoussac L., Bonnet C., Bret-Mestries E., Seassau C., Gavaland A., Raffaillac D., Tribouillois H., Véricel G., Justes E. 2017. Sunflower crop: environmental-friendly and agroecological. Ocl 23(4), DOI 10.1051/ocl/2017020

Deglène L., Tourvieille D., Lesigne P., Sarafi A. 2015. Contrôle génétique de la résistance du tournesol à *Phomopsis helianthi* (June 2000)

Deguine J.P., Gloanec C., Laurent P., Ratnadass A., Aubertot J. 2017 Agroecological Crop Protection, springer edn. Dordrecht, The Netherlands, DOI https://doi. org/10.1007/978-94-024-1185-0

Délos M., Moinard J. 1996. ASPHODEL: Moèle de simulation des épidémies de Phomopsis du tournesol (*Diaporthe helianthi*). Med Fac Landbouww Univ Gent 61(2b):555–565

Délos M., Moinard J. 1997. ASPHODEL : modèle de simulation des épidemies de phomopsis du tournesol (*Diaporthe helianthi*). In: ANPP (ed) Vème Conférence Internationale sur les maladies des plantes, Tours, France, pp 413–420

Délos M., Moinard J., Jacquin D. 1995. Etude et surveillance du phomopsis : des pièges au modèle, Phytoma - La Défense des Végétaux 473: 25–27.

Desanlis M, Aubertot J.N., Mestries E., Debaeke P. 2013. Analysis of the influence of a sunflower canopy on *Phomopsis helianthi* epidemics as a function of cropping practices. Field Crops Research 149: 63–75.

Duroueix F., Penaud A. 2013. Les fongicides en renfort contre phoma et phomopsis, Perspectives Agricoles 400: 18–21.

725 Estragnat A. 1992. Phomopsis du tournesol. Oléoscope 8: 8–10.

726 Fabregue C., Eychenne N. 1998. GUIDE ”Grandes-Cultures”. Dossier support des Avertissements Agricoles
727 routés et télécopiés, URL https://archive.org/stream/RN_GC_Phomopsis_du_tournesol 1990/Phomopsis
728 du tournesol 1990_djvu.txt

729 Fawcett T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27(8):861–874, DOI
730 10.1016/j.patrec.2005.10.010

731 Fayret J., Assemat P. 1987. Evolution du *Diaporthe helianthi* (*Phomopsis helianthi*) Munt-Cvet et al. et
732 différenciation des organes reproducteurs sur les plants de tournesol après la période de végétation,
733 *Informations Techniques CETIOM* 98: 2–11.

734 Fazekas M. 1989. Study of relationships between the time and intensity of sunflower flowering and
735 infection rate of Phomopsis (*Diaporthe*) *helianthi* Munt.-Cvet et al. Novenyvedelem 25:158–162

736 Gliessman S. 2014. Agroecology: The ecology of sustainable food systems, 3rd edn. CRC Press, Boca Raton,
737 United States

738 Gulya T., Rashid K., Masirevic S. 1997. Sunflower Technology and Production. In: Schneiter AA (ed) Sunflower
739 Technology and Production, Madison, Wisconsin, USA, chap Sunflower, pp 319–322

740 Iliescu H, Baicu T (1984) Mesures de lutte intégrées contre les agents pathogènes du tournesol. Helia
741 7:39–42

742 Jinga V., Iliescu H., Rafaila C., Ciurea A., Ionita A., Csep N. 1987. Investigations on forecasting the attack of
743 brown spot and stem canker (Diaporthe-Phomopsis) in sunflower. Analele Institutul de Cercetari pentru
744 Protectia Plantelor 20:57–68

745 Jinga V., Iliescu H., Ionita A., Csép N., Iordache, E. 1992. Technological factors in the integrated control of
746 sunflower diseases, in: ISA (Ed.), Proceedings of 13th International Sunflower Conference, Pisa, Italy. 761–
747 766.

748 Jouffret P. 1997. Tournesol : phomopsis et phoma. CETIOM - Oléoscope 38:21

749 Jouffret P. 2005 Stratégie de lutte durable contre phomopsis et phoma. Oléoscope 80:30

750 Jouffret P., Labalette F., Thibierge J. 2011. Atouts et besoins en innovations du tournesol pour une
751 agriculture durable. Innovations agronomiques 14:1–17

752 Lamichhane J., Dachbrodt-Saaydeh S., Kudsk P., Messean A. 2016. Toward a reduced reliance on conventional
753 pesticides in European agriculture. Plant Disease 100(1):24, DOI https://doi.org/10.1094/PDIS-05-15-
754 0574-FE

755 Lamichhane J., Messean A., Ricci P. 2019. Research and innovation priorities as defined by the Ecophyto plan
756 to address current crop protection transformation challenges in France. Advances in Agronomy 154:81–
757 152

758 Lancashire P.D., Bleiholder H., Van Den Boom T., Langelüddeke P., Stauss R., Elfriede Weber A., Witzen-berger
759 A. 1991. A uniform decimal code for growth stages of crops and weeds. Ann Appl Biol 119:561–601

760 Landis J. R., Koch G. G. 1977. The measurement of observer agreement for categorical data. Biometrics 33:
761 159-174

762 Lartigot I. 2015. Tournesol. Guide de culture p 29

763 Lecomte V., Nolot J.M. 2011. Place du tournesol dans le système de culture TT - Sunflower in cropping system.
764 Innovations Agronomiques 14:59–76

765 Maric A., Camprag D., Masirevic S. 1988. Sunflower diseases and pests and their control

766 Masirevic S. M., Gulya T. 1992. Sclerotinia and Phomopsis two devastating sunflower pathogens. Field Crops
767 Research 30 (3-4): 271–300.

768 Mestries E., Desanlis M., Seassau C., Moinard J., Debaeke P., Dechamp-Guillaume G. 2011 Impact de la
769 conduite de culture sur les maladies du tournesol TT - Impact of management practices on sunflower
770 diseases. Innovations Agronomiques 14:91–108

771 Mihaela C. 2019. Pratiques culturales en grandes cultures 2017 IFT et nombre de traitements. Agreste Chiffres
772 et Données 2019-3:22

773 Mihaljcevic M., Muntanola-Cvetkovic M., Vukojevic J., Petrov M. 1985. Source of infection of sunflower plants
774 by *Diaporthe helianthi* in Yugoslavia. Journal of Phytopathology 113(4):334–342, DOI 10.1111/j.1439-
775 0434.1985.tb04834.x

776 Moinard J.., Eychenne N. 1998. Le Phomopsis. Guide Grandes Cultures. Dossier support des Avertissements
777 Agricoles routés et telecopiés DGAL-Service Régional de la Protection des Végétaux Midi-Pyrénées,
778 dossier support des Avertissements Agricoles routés et télécopiés, Tech rep

779 Muntanola-Cvetkovic M., Vukojevic M., Mihaljcevic M. 1988. Pathohistology of sunflower stems attacked by
780 *Diaporthe helianthi*. Canadian Journal of Botany 67: 1119–1125.

781 Penaud A., Pérès A., Régnault Y. 1995. Rappels sur la maladie. Paris

782 Peres A., Régnault Y. 1986. *Phomopsis helianthi*: production d'inoculum et mise au point d'une méthode de
783 contamination artificielle. Information Techniques CETIOM 95:24

784 Peres A., Régnault Y. 1988. *Diaporthe helianthi* Munt.Cvet. *et al.*, Éléments de biologie et d'épidémiologie
785 appliqués aux essais de lutte. In: Proceedings of the 12th International Sunflower Conference, pp 90–93

786 Perny A., Lagarde F., Jouffret P., Deverchère J. 1995. Campagnes 1993 et 1994 : vers une extension de la
787 maladie. Oléosope Numéro Spécial 14:7–15

788 Perny R.A. 1992. L'évolution du phomopsis en France. Oléoscope 8:7

789 Petrov M., Muntanola-Cvetkovic M., Mihaljcevic M. 1981. New findings on the disease of sunflower plants
790 caused by *Phomopsis helianthi* Munt.-Cvet. *et al.* (*Diaporthe helianthi* Munt.-Cvet *et al.*). Archives des
791 Sciences Biologiques (Beograd) 33:13–19

792 Pilorgé E. 2016. Sunflower: some examples of current research - Introduction. Oilseeds & fats Crops and Lipids
793 23(2):1, DOI 10.1051/ocl/2016012

794 Pinochet X. 2005. Progrès génétique : le phomopsis en avance sur le phoma. Oléoscopoe 80:18–20

795 Rabbinge R. 1993. The ecological background of food production. In: Chadwick D, Marsh J (eds) Crop
796 Protection and Sustainable Agriculture, john wiley edn, Chistester, UK, chap The ecolog, pp 2–29

797 Robin M.H., Aubertot J.N. 2015. IPSIM-WHEAT brown rust, (Injury Profile SIMulator), a hierarchical,
798 aggregative and qualitative model to predict brown rust severity as a function of cropping practices, soil,
799 climate and field environment. In: 14th International Cereal Rusts and Powdery Mildews Conference,
800 Helsingor, Denmark, 5-8 July 2015, Helsingør, Denmark

801 Robin M.H., Aubertot J.N. 2017. IPSIM-WHEAT brown rust, (Injury Profile SIMulator), a hierarchical,
802 aggregative and qualitative model to predict brown rust severity as a function of cropping practices, soil,
803 climate and field environment. In: 14th International cereal rusts and powdery mildews conference,
804 Helsingør, Denmark

805 Robin M.H., Colbach N., L. P., Montfort F., Cholez C., Debaeke P., Aubertot J.N. 2013. Injury Profile SIMulator,
806 a hierarchical aggregative modelling framework to predict an injury profile as a function of cropping
807 practices, and abiotic and biotic environment. II. Proof of concept: design and evaluation of IPSIM-Wheat-
808 Eyespot, a model that. PLoS ONE 8(10), DOI e75829

809 Sébillotte M. 1990a. Système de culture: un concept opératoire pour les agronomes. In: Combe L., Picard D.
810 (eds) Les systèmes de culture, INRA, Paris, pp 165–196

811 Skoric D., Jocic S., Hladni N., Vannozzi G. 2007. An analysis of heterotic potential for agronomically important
812 traits in sunflower (*Helianthus annuus L.*). Helia 30(46):55–74

813 Slyusa E., Antonova L., Bochkarev N. 1988. On the spread of Phomopsis infection on sunflower, Zashchita I
814 Karantin Rastenii 1: 33–35.

815 Stanojevic D. 1985. Relationship between stand density and microclimate and their effect on some characters
816 of sunflower. In: Proceding 11th International Sunflower Conference, ISA, Mar del Plata (Argentina), pp
817 01–306

818 Taverne M. 2005. Effet du milieu sur le phomopsis et le phoma. Oléoscope 80:16–18

819 Vear F. 2016. Changes in sunflower breeding over the last fifty years. OCL - Oléagineux, Corps Gras, Lipides
820 23(D203):8

821 Vear F., Garreyn M., Tourvieille De Labrouhe D. 1997. Inheritance of resistance to phomopsis (*Diaporthe*
822 *helianthi*) in sunflower. Plant Breeding 116(3):277–281, DOI 10.1111/j.1439-0523.1997.tb00996.x

823 Viguié A., Tourvieille De Labrouhe D., Vear F. 2000b. Inheritance of several sources of resistance to Phomopsis
824 stem canker (*Diaporthe helianthi* Munt.-Cvet.) in sunflower (*Helianthus annuus L.*). Euphytica 116:167–
825 179

826 Wallach D., Makowski D., Jones J. 2006. Working with dynamic crop models: evaluation, analysis,
827 parameterization, and applications. Elsevier, 2nd edition.

828 Yang S., Berry R., Lutrell E., Vongkaysone T. 1984. A new sunflower disease in Texas caused by *Diaporthe*
829 *helianthi*. Plant Disease 68:254–2

1.2 Amélioration de la qualité de prédiction d'IPSIM - Tournesol - Phomopsis

L'algorithme d'amélioration de la qualité de prédiction des modèles IPSIM présenté dans la deuxième partie de ce manuscrit a été appliqué au modèle IPSIM - Tournesol - Phomopsis. C'est le même jeu de données que celui qui a servi à son évaluation qui a été utilisé. Cependant, l'algorithme a bien respecté le principe d'indépendance entre données utilisées pour construire le modèle et données utilisées pour son évaluation. Le critère Kappa de Cohen pondéré quadratiquement κ obtenu pour la qualité de prédiction initiale est $\kappa = 0{,}366$. Ceci signifie que l'accord entre incidences observées et simulées est modéré (« fair », Landis et Koch, 1977). La meilleure qualité de prédiction est obtenue pour 3, 4 et 5 règles d'agrégations modifiées (Figure IV.1).

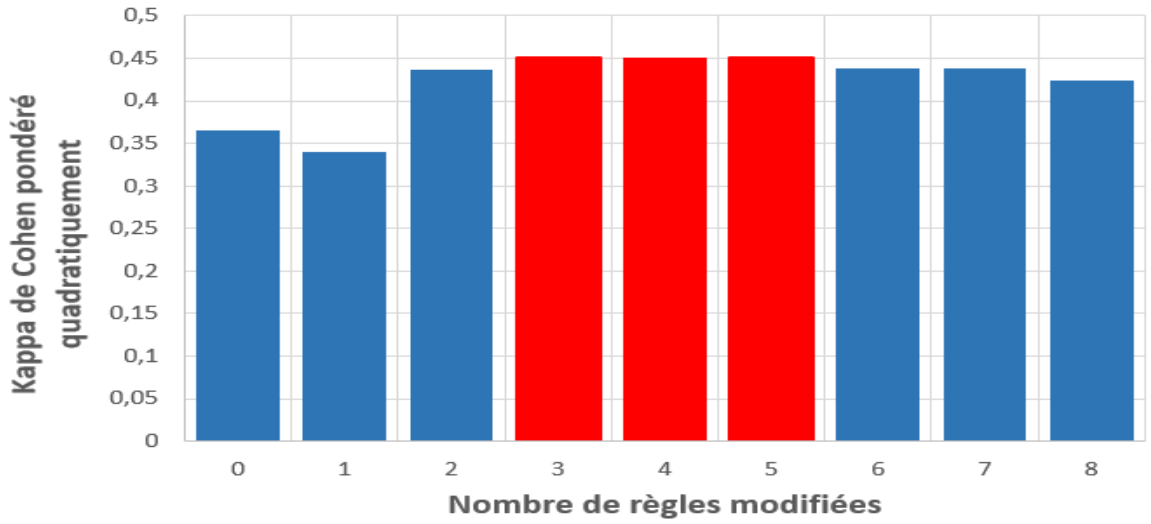


Figure IV.1 – Évolution du kappa de Cohen pondéré quadratiquement en fonction du nombre de règles d'agrégation ajustées simultanément pour le modèle IPSIM - Tournesol - Phomopsis. En rouge, le kappa maximum obtenu en validation-croisée.

Pour rester dans une démarche de parcimonie des modifications du modèle initial, nous considérons les modifications de 3 règles d'agrégation (Tableau IV.1). Les tables d'agrégation concernées sont présentées Tableau IV.2 à IV.4.

Attribut agrégé	Attribut sous-jacents	Règle d'agrégation	Valeur possible	Valeur initiale	Valeur finale
« Pression du phomopsis au niveau de la parcelle »	« Niveau de risque à l'échelle de la parcelle » « Travail du sol et enfouissement »	3	« Elevée » « Intermédiaire » « Faible »	« Intermédiaire »	« Faible »
« Pratiques culturales »	« Contrôle cultural » « Contrôle génétique » « Contrôle chimique »	4	« Défavorable » « Moyennement Défavorable » « Favorable »	« Favorable »	« Défavorable »
« Incidence du phomopsis »	« Pression du phomopsis au niveau de la parcelle » « Climat » « Pratiques culturales »	8	« Elevée » « Intermédiaire » « Faible »	« Faible »	« Modérée »

Tableau IV.1 – Synthèse des sorties de l'algorithme pour optimiser le modèle IPSIM - Tournesol - Phomopsis en ajustant quatre règles d'agrégation. Les valeurs initiales et optimisées sont les résultats des règles d'agrégation.

	Pression_du_phomopsis_au_niveau_de_la_parcelle	Climat	Pratiques_culturales	Incidence_du_Phomopsis
1	Eleve	Defavorable	Defavorable	**Elevee**
2	Eleve	Defavorable	Moyennement_favorable	**Elevee**
3	Eleve	Defavorable	Favorable	**Moderee**
4	Eleve	Moyennement_Favorable	Defavorable	**Elevee**
5	Eleve	Moyennement_Favorable	Moyennement_favorable	**Moderee**
6	Eleve	Moyennement_Favorable	Favorable	**Moderee**
7	Eleve	Favorable	Defavorable	**Moderee**
8	Eleve	Favorable	Moyennement_favorable	**Faible**
9	Eleve	Favorable	Favorable	**Faible**
10	Intermediaire	Defavorable	Defavorable	**Elevee**
11	Intermediaire	Defavorable	Moyennement_favorable	**Elevee**
12	Intermediaire	Defavorable	Favorable	**Moderee**
13	Intermediaire	Moyennement_Favorable	Defavorable	**Elevee**
14	Intermediaire	Moyennement_Favorable	Moyennement_favorable	**Moderee**
15	Intermediaire	Moyennement_Favorable	Favorable	**Faible**
16	Intermediaire	Favorable	Defavorable	**Moderee**
17	Intermediaire	Favorable	Moyennement_favorable	**Faible**
18	Intermediaire	Favorable	Favorable	**Faible**
19	Faible	Defavorable	Defavorable	**Moderee**
20	Faible	Defavorable	Moyennement_favorable	**Moderee**
21	Faible	Defavorable	Favorable	**Faible**
22	Faible	Moyennement_Favorable	Defavorable	**Moderee**
23	Faible	Moyennement_Favorable	Moyennement_favorable	**Faible**
24	Faible	Moyennement_Favorable	Favorable	**Faible**
25	Faible	Favorable	Defavorable	**Faible**
26	Faible	Favorable	Moyennement_favorable	**Faible**
27	Faible	Favorable	Favorable	**Faible**

Tableau IV.2 – Table d'agrégation de l'incidence du modèle IPSIM - Tournesol - Phomopsis avant optimisation de la qualité de prédiction telle que présentée sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables.

	Niveau_de_risque_dans_le_departement	Enfouissement_des_residus_apres_anteprecedent_si_tournesol	Pression_du_phomopsis_au_niveau_de_la_parcelle
1	Eleve	Non_Retournement	**Eleve**
2	Eleve	Retournement	**Faible**
3	Intermediaire	Non_Retournement	**Intermediaire**
4	Intermediaire	Retournement	**Faible**
5	Faible	Non_Retournement	**Faible**
6	Faible	Retournement	**Faible**

Tableau IV.3 – Table d'agrégation de la pression du phomopsis du modèle IPSIM - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables.

	Controle_cultural	Lutte_chimique	Controle_genetique	Pratiques_culturales
1	Defavorable	Non_Traitement	Sensible_et_peu_sensible	**Defavorable**
2	Defavorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Moyennement_favorable**
3	Defavorable	Traitement	Sensible_et_peu_sensible	**Moyennement_favorable**
4	Defavorable	Traitement	Tres_peu_sensible_et_resistante	**Favorable**
5	Moyennement_favorable	Non_Traitement	Sensible_et_peu_sensible	**Defavorable**
6	Moyennement_favorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Favorable**
7	Moyennement_favorable	Traitement	Sensible_et_peu_sensible	**Favorable**
8	Moyennement_favorable	Traitement	Tres_peu_sensible_et_resistante	**Favorable**
9	Favorable	Non_Traitement	Sensible_et_peu_sensible	**Favorable**
10	Favorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Favorable**
11	Favorable	Traitement	Sensible_et_peu_sensible	**Favorable**
12	Favorable	Traitement	Tres_peu_sensible_et_resistante	**Favorable**

Tableau IV.4 – Table d'agrégation des Pratiques culturales du modèle IPSIM - Sunflower - Tournesol après optimisation de la qualité de prédiction telles que présentées sous DEXi. En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables.

La modification de ces trois règles a entraîné une augmentation du kappa qui est passé de $\kappa = 0{,}366$ (un accord modéré ; « slight » Landis et Koch, 1977) à $\kappa = 0{,}451$ (un accord assez bon ; « fair » selon Landis et Koch, 1977). Une fois de plus, l'algorithme a été capable d'améliorer la qualité prédictive d'un modèle IPSIM.

La première modification (première ligne de la Tableau IV.1) a été réalisée dans la table de la pression du phomopsis au niveau de la parcelle. Elle porte sur la troisième règle. Elle correspond à un niveau de risque de phomopsis « Intermédiaire » au sein de la parcelle avec un non enfouissement de résidus potentiellement contaminés. La valeur de cette agrégation (« Pression du phomopsis ») est passée de « Intermédiaire » à « Faible ». Le jeu de données est composé de 60 situations pour lesquelles l'attribut « Niveau de risque dans le département » a pour valeur « Intermédiaire » et dont l'attribut « Enfouissement des résidus après antéprécédent si tournesol » a pour valeur« Non-retournement ». Sur ces 60 situations, 45 ont une incidence « Faible », soit les trois-quarts. Ces incidences « Faible » sont aussi expliquées par les attributs climatiques et les pratiques culturales défavorables à la maladie. Cette troisième règle d'agrégation était la seule de cette table dont l'agrégation adoptait la valeur « Intermédiaire ». Le remplacement de cette valeur par une autre valeur implique la réduction de l'échelle de l'attribut « Pression du phomopsis au niveau de la parcelle ». Cet attribut peut adopter deux valeurs « Élevée » ou « Faible ».

La deuxième modification réalisée par l'algorithme concerne la table des « Pratiques culturales ». La règle d'agrégation correspond à la combinaison d'attributs suivants : attribut « Contrôle cultural »

avec une valeur « Défavorable » : attribut « lutte chimique » avec une valeur « Traitement », et attribut « Contrôle génétique » avec une valeur « Très peu sensible et résistante ». Le changement de valeur de l'attribut agrégé « Pratiques culturales » (qui passe de « Favorable » à « Défavorable ») semble indiquer que le modèle devrait accorder davantage de poids à l'attribut « Contrôle cultural ».

Enfin, la dernière modification est réalisée pour la règle d'agrégation 8 de la table « Incidence ». Avant optimisation, une pression du phomopsis dans la parcelle « Élevée », un climat « Favorable » et des pratiques culturales « Moyennement favorables » conduisait à une incidence « Faible ». Après optimisation, la valeur de l'attribut agrégé est passé à « Modérée », signifiant que davantage de poids est accordé à la pression et aux pratiques culturales. Cette modification semble cohérente.

Les tables avec les nouvelles règles et les nouveaux poids sont présentées en Annexe 7.

Les nouveaux poids sont présentées en Tableau IV.5

Variable	**Niveau local 1**	**Niveau local 2**	**Niveau local 3**	**Niveau global 1**	**Niveau global 2**	**Niveau global 3**
Phomopsis stem canker pressure in the field	**31**			**31**		
Risk level at the farm level			53			16
Mechanical grinding and burying in fields around			47			14
Weather during Spring and/or Summer	**35**			**35**		
Weather for infection during Spring and/or Summer			57			20
Weather for girdling symptoms during Summer			43			15
Cropping practices	**35**			**35**		
Cropping control		**63**			**22**	
Escape strategy			37			8
Mitigation through crop status			63			14
Genetic control			19			6
Chemical control			19			6

Tableau IV.5 – Local and global weights of attributes of IPSIM-Sunflower-Phomopsis stem canker

Les poids des principaux attributs agrégés (« Pression du phomopsis au niveau de la parcelle », « Climat au printemps et/ou en été » et « Pratiques culturales ») sont plus équilibrés. Leurs poids respectifs sont passés de 27%, 38% et 35% à 31%, 35% et 35%. Le climat était considéré comme l'attribut ayant le plus d'impact avant optimisation, il en a autant que les pratiques culturales après optimisation. La pression du phomopsis au niveau de la parcelle en a un peu plus. La lutte chimique et le contrôle génétique ont un poids moins important après optimisation (6% contre 11%). Le contrôle par esquive et l'atténuation ont des poids plus élevés (respectivement 8 et 14% contre 5 et 8%) donnant ainsi un plus grand poids au contrôle cultural.

2. IPSIM - Tournesol - Dessèchement précoce

2.1 Identification, organisation des attributs, définition de leurs valeurs et de leurs poids

2.1.1 Modèle IPSIM - Tournesol - Dessèchement précoce sous DEXi

Ce modèle a été développé selon la méthode présentée au chapitre 2 de la deuxième partie de ce manuscrit. Trois ateliers mobilisant Mme. Bret-Mestries, M Debaeke, Mme. Seassau, Mme. Robin et M Aubertot, ont été nécessaires pour le finaliser.

La structure hiérarchique d'IPSIM - Tournesol - Dessèchement précoce et les valeurs de chaque attribut sont présentées Tableau IV.6 à gauche. Il a dix attributs agrégés (en gras) et neuf attributs basiques (police normale). Les différents poids attribués à chaque attribut sont présentés Tableau IV.7. Ce modèle dispose de huit tables d'agrégation présentées Tableau IV.8 à IV.17. Il y a 62 règles d'agrégation dans ce modèle. Il y a au total 3*2*2*2*2*2*3*2*2 = 1152 combinaisons possibles de valeurs d'entrée.

Attribute	Scale
Incidence du dessechement precoce ([M1.2-M1.3])	Tres elevee; Elevee; *Nul a faible*
├Incidence phoma au collet	Elevee; Moyenne; *Faible*
├Pression du phoma à l'echelle de l'exploitation	Defavorable ; Moyennement favorable; *Favorable*
├Niveau de risque dans le département	Frequente; Moyennement frequente; *Peu frequente*
└Enfouissement des résidus après anteprecedent (si tournesol) à l'année n	Non enfouissement; *Enfouissement*
├Sol et climat	Defavorable ; *Favorable*
└Climat pour l'infection au collet	Defavorable ; *Favorable*
└Pratiques culturales	Defavorable ; Moyennement favorable; *Très favorable*
├Lutte chimique	Non-traité; *Traité*
└Esquive (date de semis)	Tardif; *Non tardif*
└Conditions pour l'expression du dessechement précoce	Defavorables; *Moyennement favorables*; *Favorables*
├Sol et climat	Defavorable; *Favorable*
└Conditions hydriques pour l'apparition du dessechement precoce	Defavorable ; *Favorable*
└Pratiques culturales	Defavorable ; Moyennement favorable; *Tres favorable*
├Controle cultural	Defavorable ; Moyennement favorable; *Tres favorable*
└Attenuation par etat du couvert	Peu efficace; Moyennement efficace; *Efficace*
├Densite de peuplement	d> 7; 5<=d<7; *d<5 plantes*
└Statut azote du tournesol	Defavorable; *Favorable*
└Controle genetique	NA; NA

Tableau IV.6 – Structure et échelle (Scale) des attributs (Attribute) d'IPSIM - Tournesol - Dessèchement précoce. À gauche, la structure de l'arbre. À droite, les valeurs des attributs. En rouge, les valeurs sont défavorables au tournesol, en noir, les valeurs sont intermédiaires et en vert les valeurs sont favorables au tournesol

	Poids locaux normalisés					Poids globaux normalisés				
	Niv. 1	**Niv. 2**	**Niv. 3**	**Niv. 4**	**Niv. 5**	**Niv. 1**	**Niv. 2**	**Niv. 3**	**Niv. 4**	**Niv. 5**
Incidence phoma au collet	50					50				
Pression du phoma à l'échelle de l'exploitation		25					12			
Niveau de risque dans le département			63					8		
Enfouissement des résidus après antéprécédent (si tournesol) à l'année n			37					5		
Sol et climat		50					25			
Climat pour l'infection au collet			100					25		
Pratiques culturales		25					13			
Lutte chimique			50					6		
Esquive (date de semis)			50					6		
Conditions pour l'expression du Dessèchement précoce	50					50				
Sol et climat		64					32			
Conditions hydriques pour l'apparition du déssechement précoce			100					32		
Pratiques culturales		36					18			
Contrôle cultural			100					18		
Atténuation par état du couvert				100					18	
Densité de peuplement					22					4
Statut azoté du tournesol					78					14
Contrôle génétique			0					0		

Tableau IV.7 – Attributs d'IPSIM - Tournesol - Dessèchement précoce et leurs poids locaux et globaux normalisés. Niv. = Niveau

	Incidence phoma au collet	Conditions pour l'expression du dessechement précoce	Incidence du dessechement precoce ([M1.2-M1.3])
1	Elevee	Defavorables	Tres elevee
2	Elevee	Moyennement favorables	Elevee
3	Elevee	Favorables	Nul a faible
4	Moyenne	Defavorables	Elevee
5	Moyenne	Moyennement favorables	Nul a faible
6	Moyenne	Favorables	Nul a faible
7	Faible	Defavorables	Nul a faible
8	Faible	Moyennement favorables	Nul a faible
9	Faible	Favorables	Nul a faible

Tableau IV.8 – Table d'agrégation de la sortie du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi avant optimisation. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Pression du phoma à l'echelle de l'exploitation	Sol et climat	Pratiques culturales	**Incidence phoma au collet**
1	Defavorable	Defavorable	Defavorable	**Elevee**
2	Defavorable	Defavorable	Moyennement favorable	**Elevee**
3	Defavorable	Defavorable	Très favorable	**Moyenne**
4	Defavorable	Favorable	Defavorable	**Faible**
5	Defavorable	Favorable	Moyennement favorable	**Faible**
6	Defavorable	Favorable	Très favorable	**Faible**
7	Moyennement favorable	Defavorable	Defavorable	**Elevee**
8	Moyennement favorable	Defavorable	Moyennement favorable	**Moyenne**
9	Moyennement favorable	Defavorable	Très favorable	**Faible**
10	Moyennement favorable	Favorable	Defavorable	**Faible**
11	Moyennement favorable	Favorable	Moyennement favorable	**Faible**
12	Moyennement favorable	Favorable	Très favorable	**Faible**
13	Favorable	Defavorable	Defavorable	**Moyenne**
14	Favorable	Defavorable	Moyennement favorable	**Faible**
15	Favorable	Defavorable	Très favorable	**Faible**
16	Favorable	Favorable	Defavorable	**Faible**
17	Favorable	Favorable	Moyennement favorable	**Faible**
18	Favorable	Favorable	Très favorable	**Faible**

Tableau IV.9 – Table d'agrégation de l'incidence du phoma du collet du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Niveau de risque dans le département	Enfouissement des résidus après anteprecedent (si tournesol) à l'année n	Pression du phoma à l'echelle de l'exploitation
1	Frequente	Non enfouissement	Defavorable
2	Frequente	Enfouissement	Defavorable
3	Moyennement frequente	Non enfouissement	Defavorable
4	Moyennement frequente	Enfouissement	Moyennement favorable
5	Peu frequente	Non enfouissement	Moyennement favorable
6	Peu frequente	Enfouissement	Favorable

Tableau IV.10 – Table d'agrégation de la pression du phoma à l'échelle de l'exploitation du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Conditions hydrique pour l'apparition du dessechement precoce	Sol et climat
1	Defavorable	Defavorable
2	Favorable	Favorable

Tableau IV.11 – Table d'agrégation du sol et du climat du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Lutte chimique	Esquive (date de semis)	Pratiques culturales
1	Non-traité	Tardif	**Moyennement favorable**
2	Non-traité	Non tardif	**Defavorable**
3	Traité	Tardif	**Très favorable**
4	Traité	Non tardif	**Moyennement favorable**

Tableau IV.12 – Table d'agrégation des pratiques culturales du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Sol et climat	Pratiques culturales	Conditions pour l'expression du dessechement précoce
1	Defavorable	Defavorable	**Defavorables**
2	Defavorable	Moyennement favorable	**Defavorables**
3	Defavorable	Tres favorable	**Moyennement favorables**
4	Favorable	Defavorable	**Moyennement favorables**
5	Favorable	Moyennement favorable	**Favorables**
6	Favorable	Tres favorable	**Favorables**

Tableau IV.13 – Table d'agrégation des conditions pour le dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Climat pour l'infection au collet	Sol et climat
1	Defavorable	**Defavorable**
2	Favorable	**Favorable**

Tableau IV.14 – Table d'agrégation du sol et du climat pour les conditions permettant l'expression du dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Contrôle cultural	Contrôle genetique	Pratiques culturales
1	Defavorable	NA	**Defavorable**
2	Defavorable	NA	**Defavorable**
3	Moyennement favorable	NA	**Moyennement favorable**
4	Moyennement favorable	NA	**Moyennement favorable**
5	Tres favorable	NA	**Tres favorable**
6	Tres favorable	NA	**Tres favorable**

Tableau IV.15 – Table d'agrégation des pratiques culturales pour les conditions pour l'expression du dessèchement précoce du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Attenuation par etat du couvert	Contrôle cultural
1	Peu efficace	**Defavorable**
2	Moyennement efficace	**Moyennement favorable**
3	Efficace	**Tres favorable**

Tableau IV.16 – Table d'agrégation du contrôle cultural du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

	Densite de peuplement	Statut azote du tournesol	Attenuation par etat du couvert
1	d> 7	Defavorable	**Peu efficace**
2	d> 7	Favorable	**Moyennement efficace**
3	5<=d<7	Defavorable	**Peu efficace**
4	5<=d<7	Favorable	**Moyennement efficace**
5	d<5 plantes	Defavorable	**Peu efficace**
6	d<5 plantes	Favorable	**Efficace**

Tableau IV.17 – Table d'agrégation de l'atténuation par l'état du couvert du modèle préliminaire IPSIM - Tournesol - Dessèchement précoce telle que présentée sous DEXi. En rouge, les valeurs sont favorables au phoma au collet ; en noir, les valeurs sont intermédiaires ; en vert, elles sont défavorables au phoma au collet.

2.1.2 Commentaires du modèle, synthèse bibliographique

Le tableau IV.18 synthétise les informations disponibles dans la bibliographie.

Incidence du dessèchement précoce ((M1.2-M1.3))

La variable de sortie d'IPSIM - Tournesol - Dessèchement précoce est la valeur de l'attribut « Incidence du dessèchement précoce ([M1.2-M1.3]) » (Figure IV.6). Le stade M1.2-M1.3 (mi-remplissage des graines, sur l'échelle BBCH Lancashire *et al.*, 1991) permet de différencier le dessèchement précoce de la sénescence naturelle. La sortie du modèle peut prendre trois valeurs associées à un seuil de pourcentage de tournesols desséchés précocement dans le champ. Une étude aux États-Unis a présenté une régression basée sur 15 observations s'étalant sur deux années entre la perte de rendement et l'incidence du dessèchement précoce. Elle indique que 1% de plantes touchées entraînerait 0,5% de perte de rendement (Carson, 1991). Une étude française réalisée entre 1995 et 1997 dans les régions Centre et Poitou-Charentes sur le lien entre perte de rendement et pourcentage de tournesols touchés indique que 10% de plantes atteintes conduit à une perte de 0,5 $q.ha^{-1}$, valeur qui serait cohérente avec des observations réalisées précédemment dans d'autres régions (Penaud *et al.*, 1998). Une enquête a fait état de pertes atteignant 10 à 15% du rendement (Pérès et Pilorge, 1997). Il est admis que les pertes de rendement atteignent généralement 10 à 30% (ce qui est déjà jugé comme considérable), voire 50% en cas de forte attaque (Pérès et Poisson-Bammé, 1999). Le choix a été fait de considérer de manière plus large que l'incidence pouvait être considérée comme « Faible » si le pourcentage de plantes touchées par le dessèchement précoce atteint 10% (soit 5% de pertes), puis « Moyenne » si 10 à 30% de plantes sont touchées (soit 5 à 15% de pertes) et enfin « Élevée » lorsque plus de 30% des plantes sont touchées.

L'incidence du dessèchement précoce est l'agrégation de deux attributs principaux, l'« Incidence phoma au collet » et les « Conditions pour l'expression du dessèchement précoce ». Le choix a été fait de représenter ces deux étapes du cycle séparément. Le phoma au collet se manifeste en bas de tige et induit le dessèchement précoce (appelé aussi pied sec) en conditions favorables ce qui génère des pertes plus importantes. Les valeurs de l'attribut « Incidence phoma au collet » sont « Élevée », « Moyenne », « Faible » et les valeurs de l'attribut « Conditions pour l'expression du dessèchement précoce » sont « Défavorables », « Moyennement favorables », « Favorables ».

Incidence du phoma au collet

L'attribut « Incidence phoma au collet » dépend des trois attributs « Pratiques culturales », « Sol et climat » et « Pression de la maladie au niveau de l'exploitation » avant la floraison, au

cours du stade bouton floral (Figure IV.6, Tableaux IV.18). L'attribut « Conditions pour l'expression du dessèchement précoce » ne dépend que des deux attributs « Pratiques culturales » et « Sol et climat » puisque le dessèchement précoce est provoqué par le phoma au collet à partir de la floraison (Seassau, 2010).

Pression du phoma à l'échelle de l'exploitation

L'attribut « Pression du phoma à l'échelle de l'exploitation » (Tableaux IV.10 et IV.18) dépend du niveau de risque dans le département. Nous supposons que ce dernier représente correctement le risque à l'échelle de l'exploitation. Le risque dans le département est une information qui sera facilement accessible aux futurs utilisateurs du modèle puisqu'une carte de France a été établie par Terres Inovia avec les fréquences du phoma au collet dans chaque département. Toutes les zones de production du tournesol étant touchées par le phoma, cet attribut dispose d'un faible poids (8%, Tableau IV.7).

Le mycélium du phoma peut survivre sur les résidus infectés (Pérès et Poisson-Bammé, 2000 ; Pinochet, 2003) pendant un an (Pérès et Poisson-Bammé, 1999). Les ascosporess peuvent être dispersées par la pluie et le vent (Maric *et al.*, 1988) à l'échelle de l'exploitation et des exploitations avoisinantes (Jouffret, 2005). La présence du bioagresseur dépend aussi des travaux du sol réalisés dans les champs alentours ayant contenu du tournesol. La dégradation des résidus de tournesol est favorisée par le broyage des résidus et leur enfouissement. Cette pratique permet de diminuer la survie du mycélium et la production d'inoculum primaire (Masirevic et Gulya, 1992 ; Poisson-Bammé et Pérès, 2000 ; Jouffret, 2005 ; Remidi, 2013). Cette pratique est appliquée, mais il est fréquent que des résidus entiers puissent encore être présents dans un champ. Pour cette raison, l'attribut intitulé « Enfouissement des résidus après antéprécédent (si tournesol) à l'année n » ne dispose que d'un faible poids (5%, Tableau IV.7).

Sol et Climat 1 - Incidence du phoma au collet

La partie « Sol et climat » n'est constituée que du climat pour l'infection au collet. Il n'y a pas, à notre connaissance, d'étude sur l'effet du type du sol sur le phoma au collet. Plusieurs études préliminaires ont été réalisées sur l'effet de la pluviométrie sur les émissions d'ascosporess au cours de la phase végétative du tournesol et présentent des contradictions. La pluviométrie et la température ne seraient pas des variables essentielles pour l'émission d'ascosporess (Guerin, 1997, Descorps, 2010), tandis que d'autres études suggèrent que la pluviométrie serait la principale variable à l'origine de fortes éjections (Délos *et al.*, 1998 ; Bordat *et al.*, 2011). Malgré des études rapportant des résultats apparemment contradictoires, cette variable a été conservée car elle peut

être à l'origine d'éjections de faibles quantités d'ascosporess qui pourraient provoquer de graves attaques (Bordat *et al.*, 2011). De plus, la pluviométrie permet la germination d'ascosporess sur le tournesol (Roustaee *et al.*, 2000 ; Weeraratne et Priyantha, 2003). La pluviométrie est considérée comme importante et a le poids le plus important (60%, Tableau IV.7).

Pratiques culturales 1 - Incidence du phoma au collet

L'attribut « Pratiques culturales » dépend des deux attributs « Lutte chimique » et « évitement (date de semis) ». Cette partie ne prend pas en compte l'effet du couvert (dépendant de la densité de peuplement et de la fertilisation) sur le phoma au collet, en effet, la relation entre l'indice foliaire à la floraison et le pourcentage de tournesol atteint par le phoma au collet n'a pu être établie (Debaeke *et al.*, 2003).

La lutte chimique a été intégrée dans le modèle mais ne dispose pas d'un poids très élevé (6%, Tableau IV.7). En effet, les traitements contre le phoma se font en même temps que les traitements contre le phomopsis, à la **L**imite **P**assage **T**racteur (LPT) qui correspond à la période de bouton floral, qui est la phase la plus réceptive au phomopsis. Or, pour être efficace, le traitement doit être positionné juste avant la contamination et le phoma au collet peut se manifester tout au long du cycle jusqu'à la floraison (après la LPT). Elle n'est donc pas toujours efficace, et est économiquement injustifiée (Pinochet, 2003).

La stratégie d'évitement se fait par un semis tardif (avant le 1^{er}mai, Taverne, 2005). Toutefois, si cette technique est appliquée pour des maladies comme le phomopsis (Debaeke *et al.*, 2001) dont les éjections d'ascosporess sont susceptibles de donner lieu à une infection au cours du bouton floral, cette technique serait moins efficace pour le phoma dont l'agent pathogène responsable peut émettre et contaminer le tournesol durant tout son cycle. En revanche, les semis précoces permettent une diminution du stress hydrique en fin de cycle favorable au dessèchement précoce. Cet attribut a un faible poids (6%, Tableau IV.7) dans le modèle.

Conditions pour l'expression du dessèchement précoce

Le dessèchement précoce s'exprime sur la proportion de plantes atteintes par le phoma au collet. Son expression dépend des attributs « Sol et climat » et « Pratiques culturales ».

Sol et Climat 2 - Conditions hydrique pour l'apparition du dessèchement précoce

Le dessèchement précoce se développant après la floraison, l'attribut « Sol et climat » représente l'effet des conditions hydriques après la floraison sur le développement du pied sec. Les conditions hydriques dépendent de la disponibilité en eau (issue des précipitations et de l'irrigation) dans le sol (Debaeke, 2002). Une bonne disponibilité post-floraison en eau est d'autant plus importante que

le flux transpiratoire du tournesol peut être limité par la présence du champignon (Haefele *et al.*, 2008 ; Seassau, 2010). La topographie semblerait avoir un impact sur le dessèchement précoce (Bordat *et al.*, 2011), mais celui-ci serait d'avantage lié à la disponibilité en azote et en eau qu'à la pente. Le statut azoté n'est pas traité dans cette partie.

Pratiques culturales 2 - Conditions pour l'expression du dessèchement précoce

Cet attribut regroupe le « Contrôle génétique » et le « Contrôle cultural ».

1. Le contrôle génétique a fait l'objet de plusieurs études à l'ENSAT et à l'INRAE (Bordat *et al.*, 2011). Il en est ressorti que la résistance du tournesol au phoma au collet varie en fonction de la variété et de l'organe touché et qu'elle est de type quantitative. Des résultats d'expérimentations menées en champ et en serre n'ayant pas concordés (Bordat *et al.*, 2011), l'identification de variétés résistantes en champ n'a pas été possible. Si certaines variétés sont connues pour leur résistance partielle (Jouffret, 2005), il n'existe actuellement aucune classification variétale officielle. Cet attribut a été inséré pour la progression future des connaissances. Un poids de 0% lui a donc été attribué (Tableau IV.7). L'attribut a néanmoins été conservé afin de rester homogène avec d'autres bioagresseurs pour lequel le contrôle génétique a un certain poids.
2. Le contrôle cultural ne dépend que de l'atténuation par l'état du couvert (fertilisation azotée et densité de peuplement). La disponibilité en azote est un des principaux facteurs influençant le dessèchement précoce (Seassau, 2010). Il agit sur le développement foliaire, créant un micro-climat favorable au phoma (Stanojevic, 1985). Le couvert est à la base de sa demande hydrique (Jouffret, 2005). Lorsque ce dernier est trop développé (dû à un apport azoté non limitant), le stress hydrique favorise le dessèchement précoce et le tournesol puise les réserves en eau du sol (Debaeke, 2002). La densité de peuplement agit sur le diamètre de la tige (Massey, 1971 ; Sedghi *et al.*, 2008). Une forte densité (par exemple 6,2 plantes.m^{-2}) entraîne un petit diamètre. Une forte densité couplée à un fort apport en azote entraîne la prolifération du champignon (Debaeke, 2002), favorisant le dessèchement précoce par embolie des vaisseaux conducteurs (Huber et Gillespie, 1992 ; Seassau, 2010 ; Seassau *et al.*, 2012). Le contrôle cultural a un poids de 18% (Tableau IV.7).

Dessèchement précoce					Sens de l'effet	Sources
Incidence du dessèchement précoce ([M1.2-M1.3])					Le dessèchement peut entraîner des pertes de 13q.ha^{-1}. Le stade [M1.2-M1.3] permet de différencier un dessèchement dû à la maladie ou la sénescence. 10% de plantes atteintes conduit à une perte de 0,5 q.ha^{-1}.	Gulya *et al.*, 1984 ; Donald *et al.*, 1987 ; Penaud *et al.*, 1998 ; Pérès *et al.*, 2000
Incidence phoma au collet	Pression du phoma à l'échelle de l'exploitation	Niveau de risque dans le département			Les ascospores de phoma se disséminent sur de longues distances et l'inoculum primaire est principalement issu des parcelles avoisinantes (allo-inoculum). La présence d'inoculum sur une parcelle dépend plus des conditions de voisinage que de l'historique de celle-ci. La composition de l'assolement à l'échelle territoriale prime sur la succession culturale à l'échelle de la parcelle.	Gulya *et al.*, 1984 ; Donald *et al.*, 1987 ; Maric *et al.*, 1988 ; Penaud *et al.*, 1998 ; Pérès and Poisson, 1999 ; Pérès and Poisson, 2000 ; Pérès *et al.*, 2000 ; Jouffret, 2005
		Enfouissement des résidus après antéprécédent (si tournesol) à l'année n			La présence de résidus de tournesols infectés à la surface du sol est considérée comme la principale source d'inoculum primaire. Le broyage et l'enfouissement des résidus conditionnent la durée de survie de l'inoculum et sa production.	Masirevic and Gulya, 1992 ; Délos *et al.*, 1998 ; Poisson-Bammé et Pérès, 2000 ; Pinochet, 2003 ; Taverne, 2005 ; Jouffret, 2005 ; Remidi, 2013
	Sol et climat	Climat pour l'infection au collet			Une pluviométrie de 0,5 mm ou une forte humidité relative et une température supérieure à 6°C sont favorables à la production d'inoculum.	Guerin, 1997 ; Délos *et al.*, 1998 ; Roustaee *et al.*, 2000 ; Weeraratne and Priyantha, 2003 ; Descorps, 2010 ; Bordat *et al.*, 2011 ; Seassau *et al.*, 2012
	Pratiques culturales	Lutte chimique			Trois produits sont aujourd'hui autorisés contre le phoma du tournesol. Ils sont également utilisés contre le phomopsis. La recommandation d'une protection fongicide est liée au risque phomopsis, déterminé par le BSV (Bulletin de Santé du Végétal). La lutte chimique contre le phoma sur collet est considérée comme peu efficace.	Pinochet, 2003 ; Duroueix et Penaud, 2013
		Esquive (date de semis)			Un semis tardif (mai plutôt qu'avril) augmente la proportion de plantes infectées. Un semis précoce (début avril) permet d'éviter le stress hydrique en fin de cycle favorable au dessèchement précoce.	Pérès and Lefol, 1996 ; Debaeke *et al.*, 2001 ; Debaeke et Pérès, 2003 ; Taverne, 2005 ; Mestries *et al.*, 2011 ; Desanlis, 2013
Conditions pour l'expression du dessèchement précoce	Sol et climat	Conditions hydrique pour l'apparition du dessèchement précoce			L'absence d'irrigation postfloraison favorise la sévérité de la maladie ; une irrigation postfloraison limite le symptôme de dessèchement précoce.	Debaeke, 2002 ; Jouffret, 2005 ; Haefele *et al.*, 2008 ; Seassau, 2010
	Pratiques culturales	Contrôle cultural	Atténuation par état du couvert	Densité de peuplement	Une densité de peuplement élevée est favorable à la maladie. De 1,5 à 5 plantes/m^2, pas de changement prononcé. Pour une densité de peuplement > à 6 plantes/m^2, augmentation du pourcentage de pieds secs.	Massey, 1971 ; Stanojevic, 1985 ; Huber and Gillespie, 1992 ; Debaeke, 2002 ; Debaeke et Pérès, 2003 ; Taverne, 2003 ; Sedghi *et al.*, 2008 ; Haefele *et al.*, 2008 ; Debaeke *et al.*, 2010 ; Seassau, 2010 ; Seassau *et al.*, 2010 (a, b) ; Mestries *et al.*, 2011 ; Seassau *et al.*, 2012 ; Desanlis, 2013
				Statut azoté du tournesol	Augmentation du pourcentage de pieds secs avec l'augmentation de la fertilisation azotée.	
		Contrôle génétique			Il existe des variétés avec différents comportements ; mais à l'heure actuelle, il n'existe pas de classement officiel de sensibilité des variétés.	Jouffret, 2005

Tableau IV.18 – Synthèse des connaissances sur le phoma au collet et le dessèchement précoce, le sens de l'effet et les sources bibliographiques

2.2 Description du convertisseur

Le convertisseur pour le modèle IPSIM - Tournesol - Dessèchement précoce porte sur les variables décrivant les pratiques culturales (décrites Tableau IV.20) et la pluviométrie.

Variable d'entrée d'IPSIM	Valeurs des variables d'entrée d'IPSIM	Valeurs des variables d'entrée du convertisseur	Unités
Niveau de risque dans le département	Fréquente, Moyennement fréquente Peu fréquente	Département de production	Nom
Climat pour l'infection au collet	Défavorable, Favorable	Irrigation (date et quantité)	jour et mm
		Pluviométrie journalière	mm
		Date de semis	jour
		Variété de tournesol	Nom

Conditions hydrique pour l'apparition du dessèchement précoce	Défavorable, Favorable	Pluviométrie journalière	mm
		Irrigation (date et quantité)	mm
		Evapo-transpiration potentielle	mm
		Date de semis	jour
		Variété de tournesol	nom
Statut azoté du tournesol	Défavorable, Favorable	Reliquats azotés	u
		Fertilisation azotée	u
		profondeur du sol	mm

Tableau IV.20 – Variables d'entrée d'IPSIM, valeurs des variables d'entrée d'IPSIM, Variable(s) d'entrée utilisée(s) dans le convertisseur, valeurs des variables d'entrée du convertisseur et leurs unités

Niveau de risque dans le département

Une carte de la fréquence du phoma au collet en France est régulièrement mise à jour par Terres Inovia à partir d'études de pratiques culturales (Tableau I.1, partie I). À chaque département est associé une fréquence. Le phoma est considéré comme étant « Fréquent » dans 34 départements, « Modérément fréquent » dans 17 départements, « Peu fréquent » dans 51 départements. Les cas avec l'adjectif « Indéterminé » sur la carte mise au point par Terres Inovia sont inclus dans la catégorie « Peu fréquent ». Le phoma touche toutes les régions cultivant le tournesol et la faible fréquence du phoma est liée à la faible présence du tournesol dans le département. L'utilisateur doit renseigner le nom de son département dans le convertisseur qui lui retourne le niveau de risque pour IPSIM - Tournesol - Dessèchement précoce.

Effets du climat et des pratiques culturales sur l'infection au collet

Cette partie du convertisseur simule l'effet de la pluviométrie sur l'infection du tournesol par le phoma. Les variables d'entrée sont la pluviométrie journalière pendant un an, l'irrigation, la date de semis et la variété utilisée. Le cycle de vie du tournesol est simulé en calculant des temps thermiques (base = 6°C) à partir de la date de semis. Chaque étape clé du tournesol est associée à un temps thermique, tableau 5 du Chapitre 1 de cette partie) et un jour julien. Les précocités de floraison et de maturation sont dépendantes de la variété. Les équations suivantes définissent la valeur attribuée.

$$\begin{cases} I = \acute{}Defavorable\check{} \text{ si } \sum_{d=S}^{d=E} \delta_{rs}(j) \geq D_{min} \\ \quad sinon \\ I = \acute{}Favorable\check{} \end{cases} \tag{2.1}$$

Où S est le jour julien du semis, E est le jour du semis + 40 jours ; d est l'indice de sommation journalier ; D_{min} = 10 jours est le nombre minimum de jours où la pluviométrie atteint un certain seuil ; $\delta_{rs}(j)$ est le symbole de Kronecker :

$$\begin{cases} r = s \text{ si } R(j) > R_1^{min} \\ \quad et \\ r \neq s \text{ sinon} \end{cases} \tag{2.2}$$

Où R_{IPSIM1}^{min} est la valeur minimum de pluviométrie au dessus de laquelle le contact des asco-

sporess avec les plantes est possible. Sa valeur a été arbitrairement fixée à 0,5 mm (en raison de la ressemblance épidémiologique avec le phomopsis (Moinard et Eychenne, 1998).

Le convertisseur renvoie la valeur « Défavorable » (au tournesol) ou « Favorable » (au tournesol) pour l'entrée climatique d'IPSIM.

Réserve utile du sol

Cette partie du convertisseur mesure le stress hydrique subi par la culture au cours de la période post-floraison. Les variables d'entrée sont la pluviométrie journalière, l'irrigation (dates et quantités), l'évapo-transpiration potentielle, la date de semis et la variété utilisée. Pour cela, il calcule le bilan hydrique de la culture selon les équations suivantes :

$$ETR_j/ETP_j = min(RH_j, RU)/RU \tag{2.3}$$

$$RH_j = RH_{j-1} + R_{j-1} + IRR_{j-1} - k_c * ETP_j \tag{2.4}$$

Avec ETR l'évapo-transpiration réelle (en mm) : il s'agit de la quantité d'eau évaporée par le sol et la culture. L'ETP, l'évapo-transpiration potentielle (en mm), il s'agit de la quantité d'eau maximale que pourrait transpirer le sol et la culture ; RH_j (en mm) la Réserve Hydrique dans le sol accessible au tournesol au jour j ; $RU = RH_{max}$, la quantité en eau maximale que la plante peut utiliser dans un horizon de sol donné. Par simplicité, le choix a été fait de considérer que la RU est la même sur tout l'horizon, et de fixer sa valeur à 150 mm (pour un horizon de sol de 180 mm). Lorsque la réserve hydrique du sol dépasse la RU, le lessivage a lieu. R_{j-1} est la pluviométrie (en mm) au jour précédent ; IRR_{j-1} est l'irrigation (en mm) au jour précédent. Ces deux variables alimentent la réserve en eau du sol ; k_c est le coefficient cultural qui corrige l'ETP en fonction de la culture et de son stade phénologique. Ce dernier est décrit selon le tableau IV.21. L'ETP représente l'évapo-transpiration.

Période phénologique	Valeur
Semis - Phase végétative	0,2
Phase végétative - Bouton floral	0,5
Bouton floral - Floraison	0,7
Floraison - Maturité	1,1
Maturité - Récolte	0,2

Tableau IV.21 – Valeurs de k_c pour chaque période phénologique du tournesol (Debaeke, communication personnelle)

Lorsque $ETR_j/ETP_j <= 0,4$ au cours de la floraison, le tournesol est en déficit hydrique. L'attribut « Conditions hydrique pour l'apparition du dessèchement précoce » prend alors la valeur « Défavorable » (pour le tournesol), sinon il prend la valeur « Favorable » (pour le tournesol).

Esquive du phoma

L'esquive (ou évitement) du phoma se base sur le choix de la date de semis. Il a été arbitrairement considéré qu'un semis effectué avant le 1er mai est « Précoce », sinon il est « Tardif ».

Statut azoté du tournesol

Cette rubrique renvoie le niveau de stress azoté subi par la culture. Il dépend de la quantité d'azote initiale contenue dans le sol, de la profondeur du sol et de la fertilisation. Des enquêtes sur les pratiques culturales ont permis de mettre en place des recommandations de fertilisation `https://www.terresinovia.fr/-/privilegier-une-fertilisation-azotee-en-vegetation-plutot-qu-au-semis-du-tournesol?p_r_p_categoryId=130439&p_r_p_tag=69903&p_r_p_tags=417285`. Le convertisseur se base sur ces recommandations (Tableau IV.22). Par exemple, si la quantité d'azote dans un sol peu profond au semis est inférieure ou égale à 45 u, et si la fertilisation apportée est inférieure à 60 u, alors le statut azoté de la culture reste « Favorable » (au tournesol).

		Profondeur du sol			
		Inférieure ou égale à 100 cm	**Supérieure à 100 cm**		
N initial	**Inférieur ou égal à 45 u**	Favorable	Favorable	**Inférieure à 60 u**	**Fertilisation**
		Défavorable	Défavorable	**Supérieure ou égale à 60 u**	
	Supérieur à 45 u et inférieur à 75 u	Favorable	Favorable	**Inférieure à 60 u**	
		Défavorable	Défavorable	**Supérieure ou égale à 60 u**	
	Supérieur à 75 u	Favorable	Favorable	**Égale à 0 u**	
		Défavorable	Favorable	**Supérieure à 0 et inférieure à 60 u**	
		Défavorable	Défavorable	**Supérieure ou égale à 60 u**	

Tableau IV.22 – Convertisseur définissant le statut azoté du tournesol (en unité, *u*) en fonction du reliquat azoté (en unité, *u*), la profondeur du sol et la fertilisation azotée

2.3 Évaluation future du modèle IPSIM - Tournesol - Dessèchement précoce

Le modèle IPSIM - Tournesol - Dessèchement précoce n'a pas pu faire l'objet d'une évaluation. Un jeu de données est actuellement en cours de constitution (Terres Inovia et INRAE). Ce modèle pourra être prochainement évalué avec un jeu de données provenant de Charente, Charente-Maritime, Cher, Côte-d'Or, Deux-Sèvres, Gers, Haute-Garonne, Jura, Loiret, Lot-et-Garonne, Marne, Meurthe-et-Moselle, Nièvre, Rhône, Saône-et-Loire, Vendée et Vienne de 2002 à 2018.

3. IPSIM - Tournesol - Orobanche

3.1 Construction sur le modèle IPSIM - Tournesol - Orobanche

Le modèle préliminaire pour l'orobanche a été construit avec les experts suivants : Emmanuelle Mestries, Chrsitophe Jestin (ingénieur spécialiste de l'orobanche du tournesol, Terres Inovia) et Stéphane Muños (chercheur spécialiste de l'orobanche du tournesol, INRAE) au cours de 3 ateliers.

La structure hiérarchique d'IPSIM - Tournesol - *Orobanche* est présentée Figure IV.2 à gauche et les valeurs des attributs sont exposées à droite. Il possède sept attributs agrégés (en gras) et huit attributs basiques (police normale). Les poids des attributs sont exposés Tableau IV.23 et les tables d'agrégation sont exposées Tableaux IV.24 à IV.30. L'attribut de sortie d'IPSIM - Sunflower - *Orobanche* est l'attribut « Degré d'infestation d'orobanche » qui est le nombre moyen d'orobanches par pied de tournesol. La représentation des pratiques culturales est structurée en deux attributs : le contrôle cultural et le contrôle génétique. Le niveau de contrôle cultural est défini par la gestion du stock grainier et la phase post-fixation du cycle d'*O. cumana*. Deux attributs ont été identifiés pour la gestion du stock grainier : la rotation des cultures et l'utilisation de plantes pièges. Le niveau de contrôle génétique est une variable d'entrée du modèle.

Attribute	Scale
Incidence de l'orobanche	Classe 4; Classe 3; Classe 2; Classe 1; *Classe 0*
Pratiques culturales	Défavorable; Moyennement favorable; *Favorable*
Contrôle cultural	Défavorable; Moyennement favorable; *Favorable*
Gestion du stock grainier	Défavorable; Moyennement favorable; *Favorable*
Rotation de cultures	2-3 ans; 4-6 ans; *10 ans*
Plantes pièges	Non; *Oui*
Développement de l'orobanche post-fixation	Inefficace; *Efficace*
Herbicide	Traitement inefficace; *Traitement efficace*
Contrôle génétique	Sensible; Moyen; *Peu sensible à résistante*
Climat	Défavorable; Moyennement favorable; *Favorable*
Climat pour la germination	Défavorable; Moyennement favorable; *Favorable*
Climat post-fixation	Défavorable; Moyennement favorable; *Favorable*
Risque au niveau régional	Défavorable; Moyennement favorable; *Favorable*
Proximité de la parcelle dans une zone à orobanche	Défavorable; *Favorable*
Part du tournesol dans l'assolement régional	Importante; Moyenne; *Faible*

Figure IV.2 – Structure d'IPSIM-*O. cumana*. À gauche, la structure de l'arbre avec en gras, les attributs agrégés et en non-gras, les attributs basiques. À droite, les valeurs des attributs

Attribut	**Niveau local 1**	**Niveau local 2**	**Niveau local 3**	**Niveau local 4**	**Niveau global 1**	**Niveau global 2**	**Niveau global 3**	**Niveau global 4**
Pratiques culturales	**57**				**57**			
Contrôle cultural		**37**				**21**		
Gestion du stock grainier			36				8	
Rotation de cultures				82				6
Plantes pièges				18				1
Développement de l'orobanche post-fixation			**64**				**14**	
Herbicide				100				14
Contrôle génétique		63				36		
Climat	**31**				**31**			
Climat pour la germination		43						13
Climat post-fixation		57						18
Risque au niveau régional	**12**				**12**			
Proximité de la parcelle dans une zone à orobanche				64				8
Part du tournesol dans l'assolement régional				36				4

Tableau IV.23 – Table des poids des attributs du modèle d'IPSIM-*O. cumana*. De gauche à droite : les poids locaux normalisés et globaux normalisés.

Les six tables d'agrégation sont présentées Tables IV.24 à IV.30.

	Pratiques culturales	Climat	Risque au niveau régional	Incidence de l'orobanche
1	Défavorable	Défavorable	Défavorable	**Classe 4**
2	Défavorable	Défavorable	Moyennement favorable	**Classe 4**
3	Défavorable	Défavorable	Favorable	**Classe 4**
4	Défavorable	Moyennement favorable	Défavorable	**Classe 3**
5	Défavorable	Moyennement favorable	Moyennement favorable	**Classe 3**
6	Défavorable	Moyennement favorable	Favorable	**Classe 3**
7	Défavorable	Favorable	Défavorable	**Classe 2**
8	Défavorable	Favorable	Moyennement favorable	**Classe 1**
9	Défavorable	Favorable	Favorable	**Classe 1**
10	Moyennement favorable	Défavorable	Défavorable	**Classe 3**
11	Moyennement favorable	Défavorable	Moyennement favorable	**Classe 3**
12	Moyennement favorable	Défavorable	Favorable	**Classe 2**
13	Moyennement favorable	Moyennement favorable	Défavorable	**Classe 2**
14	Moyennement favorable	Moyennement favorable	Moyennement favorable	**Classe 2**
15	Moyennement favorable	Moyennement favorable	Favorable	**Classe 1**
16	Moyennement favorable	Favorable	Défavorable	**Classe 2**
17	Moyennement favorable	Favorable	Moyennement favorable	**Classe 1**
18	Moyennement favorable	Favorable	Favorable	**Classe 1**
19	Favorable	Défavorable	Défavorable	**Classe 1**
20	Favorable	Défavorable	Moyennement favorable	**Classe 0**
21	Favorable	Défavorable	Favorable	**Classe 0**
22	Favorable	Moyennement favorable	Défavorable	**Classe 0**
23	Favorable	Moyennement favorable	Moyennement favorable	**Classe 0**
24	Favorable	Moyennement favorable	Favorable	**Classe 0**
25	Favorable	Favorable	Défavorable	**Classe 0**
26	Favorable	Favorable	Moyennement favorable	**Classe 0**
27	Favorable	Favorable	Favorable	**Classe 0**

Tableau IV.24 – Table d'agrégation du degré d'infestation du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

	Contrôle cultural	Contrôle génétique	Pratiques culturales
1	Défavorable	Sensible	**Défavorable**
2	Défavorable	Moyen	**Défavorable**
3	Défavorable	Peu sensible à résistante	**Favorable**
4	Moyennement favorable	Sensible	**Défavorable**
5	Moyennement favorable	Moyen	**Moyennement favorable**
6	Moyennement favorable	Peu sensible à résistante	**Favorable**
7	Favorable	Sensible	**Moyennement favorable**
8	Favorable	Moyen	**Favorable**
9	Favorable	Peu sensible à résistante	**Favorable**

Tableau IV.25 – Table d'agrégation des pratiques culturales du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

	Gestion du stock grainier	Développement de l'orobanche post-fixation	Contrôle cultural
1	Défavorable	Inefficace	**Défavorable**
2	Défavorable	Efficace	**Moyennement favorable**
3	Moyennement favorable	Inefficace	**Défavorable**
4	Moyennement favorable	Efficace	**Favorable**
5	Favorable	Inefficace	**Moyennement favorable**
6	Favorable	Efficace	**Favorable**

Tableau IV.26 – Table d'agrégation du contrôle cultural du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

	Rotation de cultures	Plantes pièges	Gestion du stock grainier
1	2-3 ans	Non	**Défavorable**
2	2-3 ans	Oui	**Défavorable**
3	4-6 ans	Non	**Moyennement favorable**
4	4-6 ans	Oui	**Favorable**
5	10 ans	Non	**Favorable**
6	10 ans	Oui	**Favorable**

Tableau IV.27 – Table d'agrégation de la gestion du stock grainier du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche

	Herbicide	Développement de l'orobanche post-fixation
1	Traitement inefficace	**Inefficace**
2	Traitement efficace	**Efficace**

Tableau IV.28 – Table d'agrégation du développement de l'orobanche post-fixation du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

	Climat pour la germination	Climat post-fixation	Climat
1	Défavorable	Défavorable	**Défavorable**
2	Défavorable	Moyennement favorable	**Moyennement favorable**
3	Défavorable	Favorable	**Favorable**
4	Moyennement favorable	Défavorable	**Défavorable**
5	Moyennement favorable	Moyennement favorable	**Moyennement favorable**
6	Moyennement favorable	Favorable	**Favorable**
7	Favorable	Défavorable	**Favorable**
8	Favorable	Moyennement favorable	**Favorable**
9	Favorable	Favorable	**Favorable**

Tableau IV.29 – Table d'agrégation de l'effet du climat du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

	Proximité de la parcelle dans une zone à orobanche	Part du tournesol dans l'assolement régional	Risque au niveau régional
1	Défavorable	Importante	**Défavorable**
2	Défavorable	Moyenne	**Défavorable**
3	Défavorable	Faible	**Moyennement favorable**
4	Favorable	Importante	**Moyennement favorable**
5	Favorable	Moyenne	**Favorable**
6	Favorable	Faible	**Favorable**

Tableau IV.30 – Table d'agrégation de l'effet du risque régional à l'échelle de la parcelle du modèle préliminaire IPSIM-*O. cumana* telle que présentée sous DEXi. En rouge, les valeurs sont défavorables au tournesol (favorables à l'orobanche) ; en noir, les valeurs sont intermédiaires ; en vert, elles sont favorables au tournesol (défavorables à l'orobanche).

L'attribut « Degré d'infestation de l'orobanche » est défini par 27 règles d'agrégation. L'attribut « Risque au niveau régional » est défini par 6 règles d'agrégation. L'attribut « climat » est défini par 9 règles d'agrégation. L'attribut « Pratiques culturales » est défini par 9 règles d'agrégation. L'attribut « Contrôle cultural » est défini par 6 règles d'agrégation. L'attribut « Gestion du stock grainier » est défini par 6 règles d'agrégation. Il y a au total 3*2*2*3*3*3*2*2 = 1296 vecteurs de variables d'entrée possible.

3.2 Commentaires sur le modèle, synthèse bibliographique

Les effets du système de culture et de la situation de production sur l'*O. cumana* ont été peu étudiés en France, du fait de son apparition récente. Les seules études disponibles ont porté essentiellement sur la résistance génétique du tournesol face à l'orobanche. De nombreuses études ont été réalisées sur l'*Orobanche cernua* du tournesol, qui peut être confondue avec l'*O. cumana*. Elles ont longtemps été considérées comme une seule espèce, cependant leur lien phylogénétique a révélé qu'il s'agissait d'espèces différentes (Román *et al.*, 2003), quoique très proches. Nous retenons au sein du modèle IPSIM, les variables de pratiques culturales et de la situation de production dont les effets sur *O. cernua* ou *O. cumana* ont fait l'objet de plusieurs études consensuelles.

Ce modèle est basé sur des variables extraites de la synthèse bibliographique (de la littérature française et étrangère) réalisée par Terres Inovia (Jestin, 2016) sur l'orobanche du tournesol (*O. cumana*, *O. cernua*).

Orobanche				Sens de l'effet	Sources
Incidence				La présence d'1 orobanche n'a pas d'impact, de 2 à 5 orobanches par tournesol entraîne une perte de 10%, de 6 à 15, les pertes vont de 10 à 30% et au-delà, perte de 100%	Baeyaert, 2012
Pratiques culturales	**Contrôle cultural**	**Gestion du stock grainier**	**Rotation des cultures**	Une longue rotation de 9 ans pourrait montrer une efficacité notable	Grenz *et al.*, 2005b
			Plantes pièges	Intégrer dans les systèmes de culture des espèces faux hôtes vis-à-vis de l'orobanche, favorise la diminution du stock grainier de ce parasite dans le sol, de même que l'utilisation d'hôtes sensibles en tant que plante piège. Ces dernières peuvent être utilisées à condition que la culture soit détruite avant fructification de l'orobanche.	Rodriguez-Ojeda *et al.*, 2001b ; Fernandez-Aparicio *et al.*, 2009 ; Zhang *et al.*, 2013a ; Ma *et al.*, 2013 ; Antonova *et al.*, 2015 ; An *et al.*, 2015
		Développement de l' orobanche post-fixation	**Herbicide**	Il y a actuellement deux grands types d'herbicides inhibiteurs des AcétoLactate Synthase (ALS), les IMI (imidazlinones) et les SU (sulfonylurée). Les herbicides SU sont moins efficaces que les IMI. Les premières variétés hybrides tolérantes aux IMI sont arrivées sur le marché comme les variétés Clearfield® et Expressun® introduites en France, en 2010 ; ces variétés représentent 30% des surfaces (comm. pers.). Les variétés résistantes SU seraient moins efficaces que celles résistantes à IMI pour le contrôle de l'orobanche. Un nouveau caractère de tolérance à 4 herbicides incluant IMI et SU est en cours de développement à partir de population de tournesol sauvage en Argentine, connu sous l'appellation AIR.	Semirci and Kaya, 2009 ; Beckert *et al.*, 2011 ; Sala and Bulos, 2011 ; Malidza *et al.*, 2012
	Contrôle génétique			Cinq races d'orobanche (A à E) ont été identifiées. Des résistances ont été identifiées chez le tournesol vis-à-vis de l'orobanche, avec à la fois une composante monogénique (gènes dominant Or1 à Or5) et polygénique selon les populations d'orobanches, conférant jusqu'à une résistance totale de la plante. L'exploitation de ces résistances chez le tournesol a favorisé l'émergence de populations d'orobanches plus agressives, nécessitant l'exploitation de nouveaux gènes de résistance.	Vranceanu *et al.*, 1980 ; Pérez-Vich *et al.*, 2013 ; Fernandez-Martinez *et al.*, 2015 ; Molinero-ruiz *et al.*, 2015
Climat	**Climat pour la germination**			La pluviométrie contribue à la levée de la dormance. La température n'est pas une variable limitante	Chabrolin, 1938, Vallance, 1950, Logan and Stewert, 1992 ; Kebreab and Murdoch, 1999 ; Ephrath and Eizenberg, 2010 ; Murdoch and Kebreab, 2013
	Climat post-fixation			l'orobanche émerge généralement lors de phase de floraison du tournesol	Eizenberg *et al.*, 2009
Production de semences d'orobanche au niveau du territoire sur X dernières années	**Proximité de la parcelle dans une zone à orobanche**			Terres Inovia propose de participer à un suivi des zones à orobanche pour étudier sa progression	https://www.terresinovia.fr/-/participer-au-suivi-des-zones-a-orobanche-cumana?p_r_p_categoryId=130442&p_r_p_tag=182818&p_r_p_tags=434136
	Part du tournesol dans l'assolement régional			Elle est régulièrement renseignée dans les enquêtes régionales de Terres-Inovia	www.terresInovia.fr

Tableau IV.31 – Synthèse des connaissances sur l'orobanche avec les attributs, le sens de l'effet et les sources bibliographiques

3.2.1 Degré d'infestation de l'orobanche

L'évaluation de la nuisibilité de l'orobanche est considérée comme difficile car des comparaisons entre parcelles infestées et non infestées, disposant de contextes culturaux similaires doivent être réalisées et les parcelles infestées sont en cours de détection. Cependant, une étude espagnole et une étude française ont montré un fort impact de l'orobanche sur le tournesol avec des pertes pouvant aller de 50 à 100% pour une parcelle fortement infestée et composée de variétés sensibles (Dominguez, 1996 ; Jestin *et al.*, 2014). Une autre étude française a estimé des pertes entre 16 et 91%, en comparant des parcelles infestées et non infestées en France (Baeyaert, 2012). Cette dernière étude a mis en évidence une première relation entre le rendement et le nombre d'orobanches par pied de tournesol à la récolte. Elle montre qu'une orobanche par tournesol n'a pas beaucoup d'impact sur le rendement (<5%) ; en revanche entre 2 et 5 orobanches, des pertes allant jusqu'à 10% sont prédites ; de 6 à 15 orobanches sur un tournesol provoqueraient des pertes de rendement allant de 10 à 30%, et au-delà, les pertes de rendement pourraient être totales. Ces seuils définissent les cinq niveaux appelés classe 0 à 4 (classement défini par Terres Inovia) qui seront utilisés dans le modèle IPSIM-Tournesol-Orobanche. La classe 0, idéale pour le tournesol, indique l'absence d'orobanche. La classe 1, neutre pour le tournesol correspond à la présence d'une orobanche par pied. Les classes 2 à 4, indiquent respectivement de 2 à 5, 6 à 15 et plus de 15 orobanches émergées par pied. Cet attribut est déterminé par trois attributs principaux : le « Risque au niveau régional », le « Climat » et les « Pratiques culturales ».

3.2.2 Risque au niveau régional

Cet attribut distingue les zones avec du tournesol atteint par l'orobanche et les zones saines. Les graines d'orobanche sont facilement disséminées par des facteurs abiotiques (vent, eau), les animaux et les activités humaines (Jestin, 2019). Ainsi, la proximité de la parcelle concernée avec une parcelle infestée est considérée. Cette dernière variable est en cours d'étude car l'orobanche est encore mal identifiée par les agriculteurs et fait l'objet d'enquêtes afin de déterminer les zones d'infestation en France (enquête accessible à `https://survey123.arcgis.com/share/df617b4e01674aacb0c2e52c5758501a`).

3.2.3 Climat

Cet attribut décrit l'effet du climat sur les phases souterraine et aérienne de l'orobanche. La pluviométrie contribue à créer des conditions optimales pour la levée de dormance en réhydratant les graines d'orobanche (Chabrolin, 1938 ; Vallance, 1950 ; Logan et Stewart, 1992 ; Murdoch

et Kebreab, 2013). La température a aussi un impact, la germination a lieu entre 5 et 32°C pour l'*O. cumana*. Le taux de germination serait le plus rapide à un régime de température de 29/21°C jour/nuit (Kebreab et Murdoch, 1999). En revanche, les graines sont vulnérables à de fortes températures (30h à 45°C, 6 minutes à 60°C, Murdoch et Kebreab, 2013). Néanmoins, l'occurrence de telles conditions ne semble pas réaliste au champ.

En ce qui concerne l'attribut « climat post-fixation », l'orobanche émerge généralement alors que le tournesol est en phase de floraison (Eizenberg *et al.*, 2009). Nous supposons que la disponibilité en eau (pluviométrie et irrigation) est susceptible d'impacter de manière indirecte l'orobanche (achlorophylle) à travers le stress hydrique subi par le tournesol. Nous supposons que le climat (la pluie, la grêle, le vent) pourrait impacter l'état des hampes florales (casse ou altération du développement) ainsi que la dissémination des graines au sein de la parcelle.

3.2.4 Pratiques culturales

Le contrôle génétique

La résistance variétale du tournesol à l'orobanche est la principale méthode de lutte, elle a donc le plus de poids (36% du poids total).

O. cumana est caractérisée par l'existence de différents *pools* génétiques. Plusieurs races se distinguent par leur niveau de virulence. Le système d'identification de ces races est basé sur le profil de résistance monogénique (résistance impliquant un seul gène) des variétés de tournesol (Molinero-Ruiz *et al.*, 2015). À ce jour, cinq races d'orobanches ont été identifiées (notées de E, F, F+, G et H) en France (Jestin *et al.*, 2014 ; Bazerque, 2015 ; Coque *et al.*, 2016). Mais en raison de son émergence récente, la composition en race(s) des populations d'orobanche infestant les parcelles est encore mal connue (Jestin, 2016). La résistance est sous le contrôle de cinq gènes dominants (noté Or1 à Or5) (Vrânceanu *et al.*, 1980 ; Ish-Shalom-Gordon *et al.*, 1993 ; Lu *et al.*, 2000 ; Perez-Vich *et al.*, 2004 ; Sukno *et al.*, 1999). La nature de ces gènes est encore mal connue et est en cours d'étude. La plupart des études se sont focalisées sur l'interaction entre le gène Or5 et la race E. La résistance génétique peut être totale, mais elle est rapidement contournée. Une stratégie est de cumuler plusieurs mécanismes de défense(s) pour rendre difficile le contournement de la résistance. Ces résistances affectent plusieurs stades du cycle de vie d'*O. cumana*. Par exemple, elles peuvent inhiber la pénétration de l'haustorium du parasite dans les racines de l'hôte ou provoquer l'accumulation de composés toxiques ralentissant le développement des connexions vasculaires entre le tournesol et l'orobanche (Labrousse, 2002 ; Perez-de Luque *et al.*, 2008, 2009).

L'utilisation d'herbicide

Le contrôle chimique vise à réduire l'infestation de l'orobanche après fixation à l'hôte. C'est la deuxième principale méthode de lutte. Un poids de 14% (Tableau IV.23) lui a été attribué. Actuellement, Terres Inovia préconise l'utilisation de **V**ariété **T**olérante aux **H**erbicides (VTH) notamment les herbicides **IMI**dazlinones (IMI) comme Clearfield® qui est arrivé sur le marché en France en 2010 (Beckert *et al.*, 2011). Ces variétés hybrides tolérantes représentent 30% des surfaces (Jestin, 2016). La stratégie « VTH » est notamment employée pour lutter efficacement contre l'*O. cumana* (indépendamment de la race). Néanmoins, si cette possibilité peut apparaître attrayante à première vue, les recommandations pour une efficacité optimale du traitement contre l'orobanche peuvent être différentes de celles utilisées pour lutter contre les autres adventices. La lutte herbicide doit être appliquée au bon moment à la bonne dose. La connaissance de la phénologie du parasite est donc primordiale. Or, pour l'orobanche, les applications chimiques doivent être souvent répétées car les graines d'orobanches peuvent germer à plusieurs périodes du cycle de développement de la plante hôte (faits détectés chez *P. ramosa* sur colza à l'automne et à la reprise de végétation), limitant l'efficacité de cette stratégie. Cette lutte n'est pas à privilégier car outre les problèmes environnementaux, elle peut entraîner l'apparition d'adventices résistantes à des herbicides. Cette problématique existe pour de nombreuses adventices, pour lesquelles des résistances sont déjà apparues (Beckert *et al.*, 2011 ; Green, 2014). À ce jour, il n'existe pas de remontée officielle sur la présence d'orobanches tolérantes aux herbicides en France.

Cette lutte est généralement combinée avec le levier génétique (Molinero-Ruiz *et al.*, 2015) en France et à l'Etranger (*e.g.* Turquie, Espagne).

Gestion du stock grainier

— **Allongement de la durée de retour du tournesol**. Allonger la durée entre le semis de deux tournesols dans la succession des cultures est bénéfique. Mais en raison de la durée de survie des graines d'orobanche, une rotation d'au moins 9 ans serait nécessaire pour être efficace (Grenz *et al.*, 2005). Son efficacité est moins importante que les pratiques culturales précédentes, cet attribut a un poids de 6% (Tableau IV.23).

— **Cultures intermédiaires**. Les cultures intermédiaires peuvent être constituées d'espèces dites faux-hôtes. Elles agissent comme un piège sur l'orobanche en n'induisant que sa germination. L'orobanche ne peut pas se fixer pour se développer et meurt. Il existe différents degrés de stimulation. Bien que le tournesol soit sa plante hôte, *O. cumana* pourrait être stimulée de manière décroissante par le coton, l'avoine, le blé, le sorgho, le millet perlé, le pois chiche, et dans une moindre mesure le colza et le triticale (Fernandez-Aparicio *et al.*,

2009), mais aussi le panic érigé (An *et al.*, 2015), le sorgho du Soudan (Antonova *et al.*, 2015), le maïs (Ma *et al.*, 2013) et le soja (Zhang *et al.*, 2013). *O. cernua* serait stimulée de manière décroissante par le blé, le coton, le riz, l'aubergine, et le choux-fleurs (Rodriguez-Ojeda *et al.*, 2001). Cette méthode étant en cours d'étude et son impact mal quantifié, elle a un faible poids dans le modèle (1%, Tableau IV.23).

3.3 Variables à développer

Les effets de variables de pratiques culturales et du contexte pédoclimatique sur l'orobanche en conditions françaises sont mal connus, en revanche, leurs effets dans d'autres pays atteints par différentes espèces d'orobanche ont été décrits.

— **L'irrigation** aurait des impacts négatifs directs et indirects sur *O.cumana*. Plusieurs études réalisées en Roumanie et en Israël (Grenz *et al.*, 2008 ; Goldwasser et Rodenburg, 2013) ont montré qu'en conditions très sèches, l'irrigation permettrait de diminuer la concentration des exsudats racinaires stimulant les graines d'orobanche (en conditions de faible et forte pression parasitaire) et de renforcer la vigueur de la plante. Cet effet a été montré avec la technique du goutte-à-goutte sur l'orobanche rameuse (Karkanis *et al.*, 2007). Cependant, l'apport d'eau en conditions très sèches permettrait aussi d'humidifier les graines de sorte qu'elles deviennent capable de percevoir les stimuli de germination. L'effet du stress hydrique sur le stock grainier et le développement de l'orobanche post-fixation pourrait être étudié en France de deux manières :

1. en conditions non contrôlées en réalisant le suivi de l'humidité du sol. L'INRAE dispose de sondes de type CS650 et CS655 (marque Campbell Scientific) permettant d'enregistrer automatiquement l'humidité volumétrique, la conductivité électrique et la température sur un horizon de sol (ici, à 0-10 cm par exemple).
2. en conditions partiellement contrôlées en appliquant une irrigation ajustée en fonction de la pluviométrie de manière à créer différents statuts hydriques, en utilisant plusieurs techniques d'irrigation (goutte-à-goutte et par aspersion), avant et après émergence de l'orobanche dans différentes régions de France.

— **Le travail du sol** a pour but de maintenir les graines d'orobanche à une profondeur incompatible avec la présence de la rhizosphère de l'hôte (Vasyura, 1975 ; Ghersa et Martinez-Ghersa, 2000 ; Rubiales *et al.*, 2009 ; Calchei, 2011) et de diminuer la viabilité des graines en les soumettant à différents stress (*e.g.* manque d'oxygène). Une stratégie consisterait à réaliser un

labour profond une année à forte infestation à une profondeur incompatible avec le système racinaire, puis de labourer l'année suivante uniquement à moins de 20 cm (ou de ne pas labourer) pour éviter de faire remonter les graines à la surface. Si cette stratégie s'est avérée bénéfique contre l'orobanche rameuse sur colza (diminution de son impact de 80 à 90%) en Europe de l'Est (Parker et Riches, 1993), des expériences sont à mener en France pour la valider. En effet, le labour profond pourrait maintenir les graines dormantes et donc polluerait le sol à long terme, l'absence de labour favoriserait la dissémination des graines par le vent. Enfin la densité du sol n'aurait pas d'impact (Baeyaert, 2012).

— **L'utilisation de faux-hôtes** devra aussi être confirmée en champs afin de procéder à une hiérarchisation robuste des différents hôtes évoqués et de leur efficacité relative.

— **La date de semis** aurait un impact neutre sur l'orobanche selon une étude menée en France (Baeyaert, 2012). Plusieurs dates de semis (précoce, moyenne et tardive) seraient donc à tester pour confirmer cette absence de tendance.

— **La fertilisation** agirait de 3 manières différentes :

1. L'augmentation de la fertilisation azotée (couplée à l'irrigation) favoriserait l'infestation d'*O. cumana* selon une étude roumaine (Grenz *et al.*, 2008). Au contraire, d'autres études montrent que la fertilisation azotée (sous forme de fumure organique, d'ammonium ou d'urée) inhiberait le développement de l'orobanche (Rubiales *et al.*, 2009 ; Fernandez-Aparicio *et al.*, 2016). Les ions ammoniums agiraient soit directement en inhibant la germination ou l'élongation du procaulome de l'orobanche, soit via une modification du métabolisme de l'azote, les processus de détoxification deviendraient peu efficaces (Jain et Foy, 1992 ; Abu-Irmaileh, 1994 ; Van Hezewijk et Verkleij, 1996 ; Westwood et Foy, 1999 ; Fernandez-Aparicio *et al.*, 2016). L'urée serait aussi préjudiciable car elle serait probablement transformée en ammonium par l'uréase de l'orobanche (Fernandez-Aparicio *et al.*, 2016). D'autres études menées sur *O.crenata* sur fève (Jestin, 2016), sur *P. ramosa* sur tomate (Mariam et Suwanketnikom, 2004), sur *O. ramosa* sur différentes cultures (Abu-Irmaileh, 1994), confirment la tendance de la fertilisation azotée à pénaliser l'orobanche.
2. Un excès d'azote entraîne un couvert très développé, favorable au phoma et au phomopsis, il diminuerait le développement racinaire du tournesol et donc la probabilité que les graines d'orobanches rentrent en contact avec les exsudats racinaires.
3. De manière générale, les apports azotés, phosphatés et/ou soufrés pourraient modifier

la production et la nature des exsudats racinaires du tournesol, et donc la capacité de l'orobanche parasite à germer (Cechin et Press, 1993 ; Yoneyama *et al.*, 2007, 2012 ; Gaudin, 2013). Une étude menée en laboratoire a montré une tendance du phosphate à défavoriser l'infestation (Louarn, 2013).

La quantité d'autres éléments dans les échantillons de terre contenant des graines d'orobanche pourrait également être déterminée (phosphate, potassium, bore, magnésium sodium...). L'analyse chimique d'échantillons de terre fortement et faiblement infestée par l'orobanche pourrait être réalisée afin de déterminer sa composition en matière organique, son rapport C/N et son bilan humique et mettre ces données en relation avec le degré d'infestation.

Parmi les précédentes variables, l'humidité, la masse volumique, la température, la composition du sol ont fait l'objet de suivis au cours de cette thèse sur cinq parcelles (à Grisolles, Bourret, Montbécqui, Mondonville, Surgères en France et Belciugatelor en Roumanie) en 2017. Un protocole y a été mis en place. Mais en raison d'un trop faible nombre de parcelles suivies, des aléas rencontrés durant cette étude préliminaire (ravageurs du tournesol, difficultés d'utilisation du matériel en conditions estivales) et de l'hétérogénéité dans le recueil des données, les résultats n'ont pu faire l'objet d'une exploitation statistique pouvant mettre en évidence l'impact d'une variable en particulier. Cependant, le protocole mis en place (et traduit en anglais) pourra être de nouveau mobilisé pour les années à venir.

3.4 Convertisseur à développer

Si l'effet du climat sur l'orobanche n'a pas pu être étudié durant cette thèse, nous supposons qu'il existe et qu'il pourrait faire l'objet d'un modèle. SimMat ne pourra pas être adapté pour l'orobanche, bien que le processus de levée de dormance des graines dépende également de la pluviométrie et de la température. Les cycles de vie de la plante et des deux champignons (phoma et phomopsis) sont trop différents (la plante est souterraine et les champignons sont aériens) et mal connus (pour l'orobanche). En revanche, plusieurs études ont permis la mise au point d'un modèle de temps thermique prédisant les cinétiques d'émergence d'*O. cumana* (Ephrath et Eizenberg, 2010 ; Eizenberg *et al.*, 2012). Les cinétiques prédites par ce modèle ont été validées au champ sur cinq années en Israël (Eizenberg *et al.*, 2012). Le modèle a été évalué à l'aide d'expérimentations sur un horizon de sol de 0 à 100 cm. Ce modèle devrait être repris pour être testé dans des conditions de culture françaises, avec des profondeurs de sol et d'irrigation différentes de celles expérimentées en Israël. L'étude des cinétiques d'émergence devrait être déterminée en utilisant des sondes de

température dans le sol et en déterrant régulièrement des plantes sur des essais expérimentaux. La constitution d'un tel jeu de données permettrait ensuite de calibrer le modèle composé de 3 paramètres (Eizenberg *et al.*, 2012). Une fois ce modèle calibré pour des conditions françaises, il sera destiné à aider au raisonnement de la lutte herbicide qui peut être appliquée à une période optimale. Sous les conditions Israéliennes, une application de l'herbicide à 980 GDD (Growing Degree Days) a diminué de près de 90% les infections contre un effet nul à 620 GDD (Eizenberg *et al.*, 2012). Ceci n'est guère étonnant puisque l'herbicide a été appliqué sur le tournesol à un stade où l'orobanche ne l'avait pas parasité. Il sera donc important d'établir un seuil à partir duquel l'utilisation d'herbicide est efficace. Une application prématurée sera inefficace et une application trop tardive ne sera pas forcément possible en raison de la LPT et elle ne devra pas attendre que l'orobanche ait développé ses graines.

Pour connaître la répartition de l'orobanche sur le territoire, des avancées sont attendues puisqu'une carte de France est en cours de construction pour établir les zones d'infestation de l'*O. cumana*. L'analyse de ces enquêtes devrait permettre l'identification de nouvelles parcelles à suivre chaque année et d'établir le risque de contamination dans un département, voire une commune et de faire le lien avec les pratiques enregistrées par les agriculteurs.

Enfin, la lutte biologique n'a pu être intégrée bien que des organismes soient connus pour leurs effets défavorables sur l'orobanche et soient déjà exploités à l'étranger. Il semblerait que de nombreux champignons appartenant à plus de 30 genres (Boari et Vurro, 2004) et particulièrement le genre *Fusarium spp.*, (Thomas *et al.*, 1998) et des champignons mycorhiziens à arbuscules (Louarn, 2013) aient un impact sur la phase souterraine de l'orobanche (élongation de la radicule, diminution de la fixation). Des analyses de communautés microbiennes pourraient être réalisées dans des zones faiblement infestées et fortement infestées pour étudier l'impact de l'activité de champignons de ce genre en France. La lutte avec des insectes herbivores (notamment *Phytomyza orobanchia Kaltenbach* Klein et Kroschel, 2002) a également été explorée dans des pays de l'Est (Parker et Riches, 1993 ; Klein et Kroschel, 2002). Mais les conditions de réussite de ces méthodes étant aléatoires, elles ne sont pas envisagées en France.

Le modèle IPSIM conçu pour l'orobanche du tournesol est un modèle préliminaire. Des études devront être menées durant plusieurs années pour déboucher sur un modèle IPSIM fonctionnel, disposant d'un convertisseur *ad hoc*, et pour pouvoir l'évaluer.

V Discussion et perspectives

1 Rappel des principaux résultats 265

2 Limites des modèles mis en place ... 267

2.1 Des modèles évolutifs

2.2 Des modèles encore davantage tournés vers la PIC que la PAEC

3 Avancées méthodologiques de la thèse 269

3.1 Généricité de la méthode d'optimisation de la qualité de prédiction de modèle quantitatif de faible complexité

3.2 Les modèles IPSIM : améliorations méthodologiques pour leur construction

4 Perspectives pour la modélisation pour le tournesol et la PAEC 273

4.1 Avenir et utilisation future des modèles mis en place

4.2 Apport des modèles IPSIM existants

4.3 Réflexions sur la mise en place et l'optimisation de futurs modèles IPSIM

4.4 IPSIM - Tournesol : prise en compte des interactions entre bioagresseurs

4.5 Couplage des modèles SimMat, IPSIM et Sunflower-Pest pour la prédiction d'une partie du cycle épidémique, du profil de dégâts et des dommages associés

5 Conclusion 285

1. Rappel des principaux résultats

Cette thèse a permis deux grands types d'avancées : celles contribuant à l'amélioration de la protection du tournesol contre ses principaux bioagresseurs et celles relevant de la méthodologie de la modélisation.

1. Les avancées de type agronomiques correspondent aux trois modèles IPSIM pour le phomopsis, le phoma et l'orobanche. Ces modèles permettent la prédiction de l'incidence en fonction des pratiques culturales, du contexte pédoclimatique et de l'environnement de la parcelle. Ces modèles se basent sur des synthèses de la littérature scientifique et technique et l'expertise de spécialistes en phytopathologie, en malherbologie, en agronomie réunis dans des ateliers de co-construction des modèles. Trois modèles IPSIM ont été conçus : un modèle pour le phomopsis, un modèle préliminaire (car non évalué) pour le phoma et un modèle préliminaire pour l'orobanche. Les modèles IPSIM - Tournesol - Phomopsis et IPSIM - Tournesol - Phoma disposent de convertisseurs. Ces convertisseurs transforment les variables de différentes natures généralement quantitatives (*e.g.* densité de semis, variables climatiques) ou nominales (*e.g.* le nom de la variété utilisée) en variables qualitatives correspondant aux valeurs possibles des échelles des variables d'entrée des modèles IPSIM. La qualité de prédiction du modèle IPSIM - Tournesol - Phomopsis après optimisation est caractérisée par une valeur de $\kappa = 0{,}451$. L'accord entre observations et prédictions est considéré comme « assez bon » selon la classification proposée par Landis et Koch (1977). Le modèle préliminaire pour le phoma sera évalué à court terme. Les données nécessaires à son évaluation sont actuellement en cours de mise en forme (Terres Inovia et INRAE). Le modèle préliminaire pour l'orobanche n'a pas été évalué car nous ne disposions que

d'un jeu de données très insuffisant et incomplet. Les ateliers de conception du modèle ont néanmoins permis de cibler les connaissances manquantes et prioritaires. Plusieurs années d'observations en parcelles agricoles, ou d'expérimentations seront nécessaires pour combler ces trous de connaissance.

2. Des avancées méthodologiques significatives pour la modélisation ont été réalisées avec le développement de deux algorithmes. Le premier algorithme a permis d'adapter un modèle dynamique de faible complexité SimMat (conçu pour le phoma du colza avec un faible nombre de paramètres et de variables d'entrée) à deux bioagresseurs (phomopsis et phoma du tournesol). Ce modèle vise à prédire la maturation de fructifications sexuées et les éjections d'ascospores uniquement en fonction de la température moyenne et de la pluviométrie journalière. Le modèle SimMat-*Diaporthe* a permis d'expliquer 68% de la variabilité du jeu de données ($Efficience$ = 0,683) après optimisation, tandis que le modèle SimMat-*L. lindquistii* n'a permis d'expliquer que 33% de la variabilité du jeu de données disponible ($Efficience$ = 0,327). Le modèle SimMat étant un modèle mécaniste, générique, l'adaptation est peut-être à réaliser en utilisant d'autres variables climatiques. Dans ce cas, de nouvelles variables climatiques susceptibles d'intervenir dans le processus de maturation de *L. lindquistii* doivent être identifiées. Le deuxième algorithme mis en place a permis d'améliorer la qualité prédictive de modèles qualitatifs IPSIM. Il s'agit d'un algorithme générique validé informatiquement sur un modèle préliminaire IPSIM - Tournesol - Phoma sur tige. Ce modèle est passé d'une qualité de prédiction modérée (κ = 0,352) à assez bonne (κ = 0,406). L'amélioration de la qualité de prédiction est similaire pour IPSIM - Tournesol - Phomopsis (augmentation du κ de 0,366 à 0,451). Ces deux algorithmes peuvent désormais être mobilisés pour l'amélioration de la qualité de prédiction d'autres modèles, en prenant garde toutefois à la cohérence des améliorations proposées.

2. Limites des modèles mis en place

2.1 Des modèles évolutifs

Le modèle IPSIM - Tournesol - Phomopsis est le premier modèle IPSIM sur tournesol (qualité de prédiction finale, κ= 0,451) disposant de convertisseurs.

L'arbre agrégatif pourra être modifié en ajoutant de nouveaux attributs lorsque de nouvelles connaissances seront disponibles. Par exemple, le déséquilibre de la fertilisation du tournesol (N, P, et particulièrement un fort apport en K) aurait un effet défavorable sur le phomopsis (Iliescu et Baicu, 1984). Cet effet n'a pas été introduit dans l'arbre en raison du manque de précision des études réalisées et de l'incertitude quant à son action sur l'incidence de la maladie. En effet, l'étude de Iliescu et Baicu fait état de l'action de la fertilisation sans mentionner le contexte pédologique qui a une forte influence sur l'équilibre en nutriment. Le modèle développé intègre donc toutes les variables pour lesquelles nous disposons de connaissances fiables. Toutefois, ce modèle nécessite d'être évalué à nouveau car le modèle SimMat-*Diaporthe* utilisé par le convertisseur a fait l'objet de modifications après la première évaluation réalisée.

De même que pour le modèle construit pour le phomopsis, le modèle IPSIM - Tournesol - Phoma intègre toutes les informations à notre disposition. Ce modèle dispose d'un convertisseur plus simple que celui du phomopsis puisqu'il ne mobilise pas de modèles quantitatifs. Sa qualité de prédiction reste à évaluer.

Contrairement au phomopsis et au phoma qui ont déjà fait l'objet de nombreuses études, les effets des principaux facteurs (pratiques culturales, pédoclimat et environnement de la parcelle) sur

l'orobanche n'ont été que peu étudiés. Les variables qui ont fait l'objet d'études internationales et qui ont montré un intérêt ont été présentées partie IV. Différentes études devraient être réalisées dans les années à venir pour évaluer les principaux déterminants des infestations et les hiérarchiser, ce qui permettra de faire évoluer ce modèle, notamment les attributs relevant des pratiques culturales et du contexte pédoclimatique.

2.2 Des modèles encore davantage tournés vers la PIC que la PAEC

Les trois modèles mis en place relèvent pour le moment davantage de la PIC que de la PAEC. Les synthèses de la littérature scientifique et technique actuelles nous ont montré qu'il n'était pas possible pour le moment d'intégrer de(s) méthode(s) favorisant la biodiversité locale et les régulations trophiques à l'échelle du paysage qui soient défavorables au phomopsis et/ou au phoma au sein de nos modèles car nous ne disposons pas d'informations à ce sujet. Pour l'orobanche, la lutte biologique a fait l'objet de nombreuses études à l'étranger qui ne sont pas appliquées par les agriculteurs en France (Habimana *et al.*, 2013). Notamment l'impact de micro-organismes (champignons) n'ont pas pu être intégrés au sein des modèles car leurs études sont en cours en France et leurs applications n'ont pas encore fait leurs preuves et ne sont donc pas envisagées pour le moment.

3. Avancées méthodologiques de la thèse

3.1 Généricité de la méthode d'optimisation de la qualité de prédiction de modèle quantitatif de faible complexité

Les algorithmes améliorant la qualité prédictive de modèle quantitatif de faible complexité ont prouvé leur efficacité pour optimiser la qualité de prédiction du modèle SimMat. Ils sont parvenus à améliorer la qualité prédictive du modèle sur les deux bioagresseurs (qualité de prédiction finale de $MSE^{P_{cv1}}_{SM-Diaporthe} = 0{,}036$ contre $MSE^{P_i}_{SM-Diaporthe} = 0{,}054$ pour le phomopsis et $MSE^{P_{cv1}}_{SM-L.\,l.} = 0{,}064$ contre $MSE^{P_i}_{SM-L.\,l.} = 0{,}099$ pour le phoma du tournesol).

Les algorithmes mis en place ont le double avantage de pouvoir s'appliquer à des variables quantitatives (dynamiques de taux de maturation) et qualitatives (classes ordinales de précocité des dynamiques de maturation). Ils ont été appliqués au modèle SimMat mais il existe un grand nombre de modèles épidémiologiques simples dans la littérature qui pourraient être améliorés avec ces algorithmes ; un aperçu de ces derniers a été donné partie I. Le niveau de généricité des algorithmes mis en place devra cependant être amélioré à moyen terme afin qu'ils puissent être diffusés via une librairie R et puissent être appliqués dans un avenir proche à tout type de modèles ayant un faible nombre de paramètres à optimiser. Le nombre maximal de paramètres que l'algorithme sera en capacité d'optimiser n'est pas défini. Seul le temps de calcul est limitant. Ce temps de calcul dépend du nombre de paramètres, des intervalles de valeurs explorées de la taille et de la structure (nombre de sites-années) du jeu de données.

3.2 Les modèles IPSIM : améliorations méthodologiques pour leur construction

3.2.1 IPSIM : des modèles statiques mais non figés

Un modèle IPSIM peut-être construit et modifié de deux manières.

1. La première manière d'améliorer la qualité d'un modèle IPSIM est d'identifier les points de faiblesse du modèle au cours de sa construction. Cette phase permet d'identifier les incertitudes (les trous de connaissance ou les non-consensus) sur l'effet d'une pratique agricole, du paysage, ou du contexte pédoclimatique sur le développement d'un bioagresseur. La méthode de modélisation IPSIM permet de facilement prendre en compte l'évolution des connaissances. L'absence d'une information ne constitue donc pas un frein à la conception du modèle. Une variable d'intérêt dont on soupçonne une influence sur un bioagresseur sans qu'elle n'ait encore été précisément étudiée peut-être dotée d'un poids nul dans le modèle, en attendant la mise en place d'expérimentations idoines pour préciser son influence. Ceci a été réalisé pour la lutte génétique du phoma du tournesol. Nous aurions pu adopter ce choix pour le modèle de l'orobanche pour lequel de nombreux nouveaux attributs sont attendus, mais par clarté au sein du modèle, nous avons fait le choix de ne pas les y intégrer pour le moment. De même pour les modèles IPSIM sur le phomopsis et le phoma, la proposition d'une solution de biocontrôle, d'une nouvelle résistance génétique, ou d'un nouveau traitement de semences pourra être intégrée dans les modèles.
2. La deuxième manière d'améliorer la qualité prédictive d'un modèle IPSIM est la modification de ses règle(s) de décision. Lorsque le bioagresseur considéré a déjà fait l'objet de nombreuses expérimentations et qu'aucun attribut ne peut-être ajouté. L'algorithme d'optimisation de la qualité de prédiction des modèles IPSIM permet d'identifier la ou les règle(s) d'une ou plusieurs tables à modifier et la(les) valeur(s) de remplacement. Cet algorithme a été appliqué aux modèles IPSIM - Tournesol - Phoma tige (modèle préliminaire) et IPSIM - Tournesol - Phomopsis. Il est parvenu à en améliorer la qualité de prédiction en modifiant plusieurs règles (2 et 3 pour le phoma sur tige et 3, 4 et 5 pour le phomopsis). Ce protocole codé sous C++ par G. Quesnel (UR MIAT) a prouvé son efficacité. Il fera l'objet d'une librairie R afin de pouvoir être utilisé sur tous les modèles IPSIM. Cet algorithme est actuellement accessible à l'adresse suivante : `https://github.com/quesnel/efyj`.

3.2.2 Compromis entre résolution et complexité

Un équilibre doit être trouvé entre la résolution (caractérisée par le nombre de situations que le modèle peut différencier) d'une part, et le nombre d'attributs et les échelles associées d'autre part.

Un grand nombre d'attributs et un grand nombre de valeurs par attribut augmentent rapidement la complexité du modèle et la difficulté à le mettre au point car le nombre de règles de décision (matérialisées par les lignes dans les tables d'agrégation) à renseigner augmente. L'augmentation de complexité du modèle requiert des connaissances plus fines qui ne sont pas toujours disponibles pour renseigner les règles d'agrégation. En revanche, une résolution plus élevée de l'attribut de sortie (*i.e.* un nombre de valeurs possible plus élevé) permet de mieux discriminer les différentes situations lors de l'agrégation finale. Dans le cas des modèles présentés dans cette thèse, le nombre maximum de niveaux était de 3 pour les attributs de sortie. Le nombre de règles à renseigner pour l'attribut de sortie était de 27 pour le phomopsis et l'orobanche et de 9 pour le phoma.

3.2.3 Pertinence des convertisseurs

Les convertisseurs permettent de répondre aux besoins des utilisateurs des modèles IPSIM de transformer des données facilement renseignables et ne nécessitant aucune expertise en valeurs ordinales correspondant aux niveaux des échelles des attributs d'entrée. Ils permettent de simplifier l'utilisation du modèle. Le convertisseur mis en place pour le phomopsis au cours de cette thèse intègre des modèles quantitatifs, ce qui constitue un couplage original de modèles quantitatifs au modèle qualitatif IPSIM - Tournesol - Phomopsis. Les convertisseurs intègrent notamment le modèle SimMat-*Diaporthe* pour représenter les effets du climat et le modèle CART-LAI pour représenter les effets des états de peuplement. Le modèle IPSIM - Tournesol - Dessèchement précoce en revanche n'intègre pas de modèle quantitatif au sein de son convertisseur. Le climat est géré de manière plus simple par un seuil de pluviométrie. Bien que la gestion du peuplement ait des effets similaires sur le phomopsis et le phoma, cette partie a été abordée de manières différentes pour les deux bioagresseurs :

— Pour le phomopsis (premier modèle construit), le choix a été fait de traiter la question de l'état du peuplement en amont du modèle, en créant un sous-modèle de type CART au sein du convertisseur. Ce choix a été permis, car nous disposions d'un jeu de données de grande taille (issu du modèle SUNFLO) et des variables adaptées pour établir un modèle CART. La construction finale a été validée de manière statistique par sa qualité d'ajustement (sur le jeu qui a servi à sa construction) caractérisée par $\kappa = 0{,}74$ et par les experts. Les variables considérées sont quantitatives et qualitatives : les quantités d'azote initial (reliquats) et d'azote apporté, la densité de peuplement, le nom de la variété de tournesol et des fichiers climatiques. Ces données très variées sont traitées par le convertisseur qui retourne une donnée qualitative exploitable par le modèle IPSIM. En dépit du grand jeu de données dont

nous disposions, ce module a été la source d'erreurs importantes lors de l'évaluation du modèle IPSIM - Tournesol - Phomopsis. Sa qualité d'ajustement était très bonne, mais le modèle CART-LAI ne s'est pas révélé performant pour prédire les valeurs de *LAI* dont nous disposions dans le jeu de données servant à l'évaluation d'IPSIM (κ = 0,129). Cependant, il est la meilleure méthode dont nous disposons à l'heure actuelle. Une tentative de construction d'un modèle CART sur les données LAI observées au lieu des données de LAI issues du modèle SUNFLO n'a pas donné une meilleure qualité de prédiction (données non montrées).

— Dans le cas du phoma, les variables liées à l'état du peuplement ont été intégrées directement au modèle IPSIM. Cette façon de faire a alourdi la structure de l'arbre mais a simplifié la représentation de la gestion du couvert. Ce modèle devra être évalué afin de voir si ce choix se révèle judicieux ou non.

4. Perspectives pour la modélisation pour le tournesol et la PAEC

4.1 Avenir et utilisation future des modèles mis en place

4.1.1 SimMat-*Diaporthe*

SimMat-*Diaporthe* a prouvé son efficacité pour prédire les dynamiques de maturation de pseudothèces et les projections d'ascospores (classes de précocité des dynamiques) sur le jeu de données à notre disposition. La finalité de ce modèle est d'être utilisé au sein d'un OAD permettant de prédire la présence du risque d'infection du tournesol par le phomopsis en fonction de variables climatiques (pluviométrie et températures moyennes journalières). Ce modèle, codé sous différents langages (R, C++ et VBA), dispose de variables d'entrée simples à renseigner. Mais avant d'être utilisé pour le BSV, sa fiabilité doit être éprouvée par les acteurs de Terres Inovia. SimMat-*Diaporthe* sera donc utilisé (sans servir pour le BSV dans un premier temps) en même temps qu'ASPHODEL pour comparer leurs prédictions durant plusieurs années dans les régions à risque. Si leurs prédictions sont similaires, il remplacera le modèle ASPHODEL pour prévenir du risque phomopsis. Si les prédictions sont différentes, les prédictions de SimMat-*Diaporthe* devront être examinées plus attentivement et être comparées directement avec des observations dans différentes zones à risques.

Enfin, SimMat-*Diaporthe* pourra être utilisé pour réaliser une analyse fréquentielle dans tous les départements touchés par le phomopsis en France. Cette analyse, basée sur un jeu de données le plus grand possible, permettra de déterminer les périodes des premières projections significatives d'ascospores en fonction des scénarios climatiques rencontrés, afin de proposer des dates de semis permettant de mettre en place des stratégies d'esquive des projections d'ascospores. Cette démarche

contribuera à reconcevoir des itinéraires techniques moins à risque.

4.1.2 IPSIM - Tournesol - Phomopsis

Ce modèle devra être ré-évalué. La fiabilité de ce modèle devra être éprouvée par les acteurs de Terres Inovia pendant plusieurs années pour différents systèmes de culture et dans des situations de production différentes. Une fois éprouvé, il pourra être utilisé en routine comme OAD pédagogique par les ingénieurs de développement de Terres Inovia auprès des techniciens du conseil (coopératives, chambres d'agriculture,...). Il aura deux usages :

— un usage *ex ante* (avant la mise en culture) pour concevoir ou reconcevoir des itinéraires techniques.

— un usage *ex post* (après la culture) qui permettra de réaliser des diagnostics agronomiques.

4.1.3 IPSIM - Tournesol - Phoma

Le modèle IPSIM - Tournesol - Dessèchement précoce doit faire l'objet d'une évaluation de sa qualité de prédiction. Près de 200 observations d'incidences ont été identifiées et doivent faire l'objet d'un traitement pour composer un jeu de données idoine permettant l'évaluation du modèle. Lorsque sa qualité de prédiction sera satisfaisante, sa fiabilité devra être éprouvée par les acteurs de Terres Inovia de la même manière que pour le modèle IPSIM - Tournesol - Phomopsis, avant de l'utiliser comme OAD pédagogique auprès des acteurs du conseil.

Le modèle IPSIM sur le dessèchement précoce ne prend en compte que les symptômes sur collet qui représentent la forme la plus préjudiciable de la maladie. Un module pourrait être réalisé avec le phoma sur tige pour renseigner la constitution du stock d'inoculum primaire à l'échelle de l'exploitation, notamment à travers l'attribut d'entrée "Risque de phoma à l'échelle de l'exploitation" du modèle IPSIM - Tournesol - Dessèchement précoce. Ce module, préliminaire, doit être révisé et évalué. Ces modèles devront suivre les mêmes étapes de validation d'évaluation que pour le modèle IPSIM - Tournesol - Phomopsis.

4.2 Apport des modèles IPSIM existants

Depuis que les premiers modèles IPSIM ont été construits, la méthode a été diffusée, partagée et utilisée dans différentes formations. Ainsi, de nombreuses constructions de premières versions de modèles simples ont été réalisées par de futurs ingénieurs au sein d'écoles d'Agronomie, notamment l'EI Purpan et l'ENSAT à Toulouse. Cette diffusion est facilitée par le fait que cette approche de la modélisation est pédagogique, très souple, très facile d'accès, très facilement appropriable sans

connaissance mathématique pointue. IPSIM constitue un support idéal pour une première expérience de la modélisation, pour les néophytes. De plus, sa structure générique est particulièrement bien adaptée à l'intégration de différentes méthodes de contrôle existantes qui sont suffisamment décrites et qui font consensus au sein de la communauté scientifique, notamment celles relevant de la PAEC. Cette démarche de modélisation commence à être utilisée dans différentes zones géographiques à travers le monde pour différentes cultures. Une description des modèles en cours de développement est présentée Tableau V.1. Les agroécosystèmes différents sur lesquels ont porté ces modèles IPSIM montrent que l'outil de modélisation IPSIM est en cours d'application à différentes cultures, différents bioagresseurs et dans des zones géographiques très différentes. Outre les deux articles fondateurs (Aubertot et Robin, 2013 ; Robin *et al.*, 2013), différents sites internet permettent la diffusion et l'utilisation des modèles IPSIM.

— Un site web de l'INRAE présente les étapes de construction de ce modèle à l'adresse `https://www6.inra.fr/ipsim_fre/`

— Une interface a été réalisée pour réaliser des multi-simulations avec la plateforme DEXi (`http://wiki.inra.fr/wiki/deximasc/viewrev/Interface+DEXi-MASC/Accueil?rev=18.1`). Elle peut-être utilisée pour les modèles IPSIM.

— Le CIRAD a conçu un modèle pour prédire les dégâts causés par plusieurs espèces de mouches sur la chayotte (Deguine *et al.*, en prep.) et a mis en place une interface spécifique (`https://pvbmt-apps.cirad.fr/shinyproxy/chouchou/`) pour l'utilisation directe de ce modèle par les acteurs de terrain (agriculteurs et conseillers).

Modèle IPSIM	Culture	Bioagresseur(s)	Nombre d'attributs d'entrées	Nombre d'attributs agrégés	Nombre d'experts	Nombre de références	Domaine de validité	Taille du jeu de données	Valeur de Kappa / Précision / Efficience	Référence du modèle
Coffee - Orange rust	Caféier (*Coffea arabica*)	*Hemileia vastatrix*	14	12	4	15	Amérique Centrale et probablement toutes les zones touchées par la rouille	2 ans de suivi sur plusieurs centaines de parcelle pour 4 pays	NA	Motisi *et al.*, en construction http://142.93.27.89/dashboard/
Chayotte - Fruit fly	Chayotte (*Sechium edule*)	Mouches des cucurbitacées (*Bactrocera cucurbitae, Dacus ciliatus, Dacus demmerezi*)	15	7	7	3	La Réunion	50	k = 0,79	Deguine *et al.*, en préparation https://ipsim-chayote.shinyapps.io/25-05-2019/
Pommier - Carpocapse	Pommier (*Malus sp.*)	Carpocapse (*Cydia pomonella*)	NA	NA	11	62	France métropolitaine	6	k = [-0,233 - 0,37]	Nicolas Wallet-Errani, 2015 (modèle préliminaire)
Pommier - Puceron	Pommier (*Malus sp.*)	Puceron cendré (*Dysaphis plantaginea*)	NA	NA	11	48	France métropolitaine	6	k = [0,267 - 0,444]	Nicolas Wallet-Errani, 2015 (modèle préliminaire)
Pommier - Tavelure	Pommier (*Malus sp.*)	Tavelure (*Venturia inaequalis*)	21	13	11	55	France métropolitaine	6	k = [0,0 - 0,667]	Nicolas Wallet-Errani, 2015 (modèle préliminaire)
Sunflower - Phomopsis stem canker	Tournesol (*Helianthus annuus*)	Phomopsis (*Diaporthe helianthi*)	8	5	6	43	France métropolitaine	391	k = [0,289 - 0,426]	Vedy-Zecchini *et al.*, en préparation
Tournesol - Desséchement précoce	Tournesol (*Helianthus annuus*)	Phoma au collet (*Leptosphaeria lindquistii*)	9	10	4	19	France métropolitaine	200	k = 0,112	Vedy-Zecchini, 2019 (modèle préliminaire)
Tournesol - Phoma tige	Tournesol (*Helianthus annuus*)	Phoma sur tige (*Leptosphaeria lindquistii*)	14	7	5	24	France métropolitaine	157	k = 0,48	Arnaud Allègre, 2015 (modèle préliminaire)
Tournesol -Orobanche	Tournesol (*Helianthus annuus*)	Orobanche (*Orobanche cumana*)	8	7	3	23	France métropolitaine	NA	NA	Vedy-Zecchini, 2019 (modèle préliminaire)
Wheat-Brown Rust	Blé (*Triticum sp.*)	Rouille brune (*Puccinia recondita*)	11	7	9	37	France métropolitaine	1788	k = 0,66	Robin *et al., 2017*
Wheat-Yellow Rust	Blé (*Triticum sp.*)	Rouille jaune (*Puccinia striiformis*)	17	8	NA	NA	France métropolitaine	NA	NA	Robin *et al.*, en préparation
Blé - Septoriose	Blé (*Triticum sp.*)	Septoriose (*Septoria tritici*)	14	8	NA	43	France métropolitaine	NA	NA	Robin *et al.*, 2013
Wheat - Eyespot	Blé (*Triticum sp.*)	Piétin-Verse (*Oculimacula yallundae*)	13	7	7	41	France métropolitaine	526	Efficience = 0,64	Robin *et al.*, 2013
Blé - Fusariose	Blé (*Triticum sp.*)	Fusariose (*Fusarium sp.*)	7	4	NA	52	France métropolitaine	NA	Précision = 0,42	Robin *et al.*, 2013
Blé - Oïdium	Blé (*Triticum sp.*)	Oïdium (*Erysiphe graminis*)	NA	NA	NA	NA	France métropolitaine	NA	NA	Robin *et al.*, 2013
Wheat - Aphid	Blé (*Triticum sp.*)	Puceron (*Sitobion avenae*)	12	6	NA	27	France métropolitaine	NA	NA	Robin *et al.*, 2013
Wheat - Weeds	Blé (*Triticum sp.*)	15 espèces différentes	NA	NA	NA	NA	France métropolitaine	NA	NA	Robin *et al.*, 2013
DEXi_limace	Grandes cultures	Limace grise	19	12	13	32	France métropolitaine	NA	NA	Chabert *et al.*, 2018

Tableau V.1 – Caractéristiques des modèles IPSIM construits pour le blé, le tournesol, le pommier, la chayotte et le caféier

4.3 Réflexions sur la mise en place et l'optimisation de futurs modèles IPSIM

Une étude pourrait être réalisée pour déterminer les déterminants de la qualité de prédiction des modèles IPSIM. Cette analyse permettrait d'identifier les situations pour lesquelles l'approche IPSIM aurait de bonne chance d'aboutir à des modèles ayant une qualité de prédiction satisfaisante.. Plusieurs facteurs susceptibles d'influencer la qualité de prédiction des modèles IPSIM ont été identifiés : le nombre d'attributs d'entrée d'un modèle, la résolution du modèle (*c.f.* chapitre 3.2.2), le nombre et la fonction des experts mobilisés, la taille du jeu de données, et sa structure. Une telle analyse pourrait être réalisée grâce à une **A**nalyse en **C**omposantes **P**rincipales (ACP) [1] qui permettrait d'analyser les effets de ces facteurs sur la qualité prédictive des modèles. À ce jour, il n'est pas possible de réaliser ce type d'analyse en raison du faible nombre de modèle IPSIM existants et ayant été évalués.

Quelques observations peuvent d'ores et déjà être réalisées, mais elles nécessitent d'être étayées par de futurs modèles IPSIM pour être considérées comme génériques.

Le nombre de références

Si les modèles IPSIM mis au point pour le phomopsis et le phoma sur tournesol sont basés sur des synthèses bien documentées, la méthode IPSIM a été utilisée pour des bioagresseurs plus ou moins bien étudiés (de 3 références mobilisées à 62 ; Tableau V.1).

La taille du jeu de données

Nous pouvons poser l'hypothèse qu'un grand jeu de données est essentiel pour aboutir à un modèle satisfaisant. L'impact de la taille du jeu de données sur la qualité de prédiction n'a cependant pas été mis en évidence pour le moment. En effet, les modèles pour le blé évalués sur des jeux de données de plus de 500 situations ont une bonne qualité de prédiction ($\kappa \geq 0,6$; Tableau V.1). En revanche, le modèle pour le tournesol et le phomopsis a une moins bonne qualité de prédiction (caractérisée par $\kappa = 0,451$) alors que le jeu est composé de 1073 situations tandis que le modèle portant sur la chayotte présente la meilleure qualité de prédiction de tous les modèles IPSIM, avec un jeu de données comportant moins de 50 situations. Il s'agit de cas particuliers, sur lesquels la hiérarchisation des attributs, les échelles ou le manque d'expertise ont aussi un impact. Ces cas particuliers ne permettent pas d'établir une relation entre taille du jeu de données et qualité de prédiction. Au-delà de la taille, la structure est aussi importante.

1. Méthode d'analyse multivariée qui transforme des variables corrélées entre elles en de nouvelles variables décorrélées les unes des autres (Pearson, 1901). Ces nouvelles variables sont des « composantes principales ». Cette méthode permet de réduire le nombre de variables et de rendre l'information moins redondante.

La structure du jeu de données

La méthode d'optimisation de la qualité prédictive des modèles IPSIM mis en place dépend étroitement de la taille et de la richesse des situations contenues dans le jeu de données. Une règle d'agrégation non optimale ne pourra être corrigée par l'algorithme que si elle a un impact dans l'évaluation, c'est-à-dire que l'on dispose d'observations la faisant intervenir. À titre d'exemple, pour les modèles sur tournesol, toutes les règles d'agrégation des modèles n'ont pas été mobilisées car elles n'avaient pas d'impact sur la qualité de prédiction estimée avec le jeu de données utilisées : 94 combinaisons ont été mobilisées sur les 864 combinaisons possibles pour IPSIM - Tournesol - Phomopsis et 25 sur 1.152 combinaisons pour IPSIM - Tournesol - Phoma sur tige. Ainsi l'algorithme d'optimisation de la qualité de prédiction des modèles IPSIM présente la limite d'être davantage adapté aux agro-systèmes pour lesquels nous pouvons disposer de bases de données riches en combinaisons. La taille et la qualité des jeux de données constitueront toujours une limite à l'évaluation de la qualité de prédiction et à l'optimisation du modèle car les jeux de données ne sont pas toujours disponibles. Les jeux de données utilisés sont souvent parcellaires, puisque qu'ils n'ont pas été construits pour balayer systématiquement l'ensemble des combinaisons de valeurs possible des attributs d'entrée des modèles IPSIM.

4.4 IPSIM - Tournesol : prise en compte des interactions entre bioagresseurs

Lorsqu'un ensemble de modèles aura été réalisé pour chacun des principaux bioagresseurs, ces derniers pourront être couplés afin de donner le modèle IPSIM - Tournesol. Ce modèle rendra compte du profil de dégâts engendrés par les bioagresseurs majeurs du tournesol. Un modèle générique a été proposé par Aubertot et Robin en 2013 (Figure V.1) pour représenter les interactions entre bioagresseurs. Chaque bioagresseur d'un système est tout d'abord considéré individuellement (Figure V.1). L'utilisateur devra renseigner le type d'interaction qu'entretient chaque bioagresseur avec les autres (réduction, sans interaction, ou facilitation).

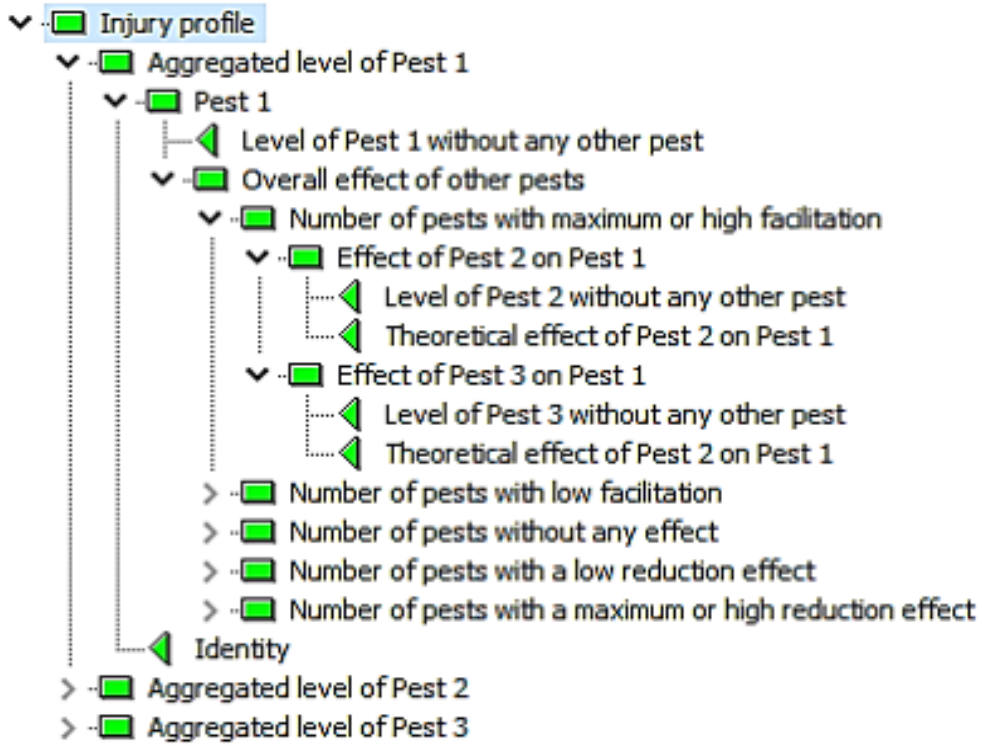


Figure V.1 – Modèle générique IPSIM - X - Interactions sous DEXi (Aubertot et Robin, 2013)

Un tel couplage nécessite une connaissance fine des mécanismes à la base des interactions entre les différents bioagresseurs. Dans le cas du tournesol, la gestion simultanée de populations de plusieurs bioagresseurs n'a pu être abordée par modélisation. Néanmoins la littérature technique fait état d'interactions entre bioagresseurs.

— L'exemple du Phoma et Phomopsis : le phoma est connu pour être capable d'infecter tous les organes du tournesol (McDonald, 1964). Il est, contrairement au phomopsis, capable d'infecter directement la tige, à travers l'auget formé par le pétiole à l'insertion de la tige. Cet auget se remplit d'eau lors de pluie (ou d'irrigation), créant des conditions favorables pour la germination d'ascospores de phoma. Son installation sur le pétiole et la nécrose de celui-ci bloque la progression du phomopsis vers la tige. Il s'agit donc d'une compétition dans laquelle le phomopsis peut-être perdant.

— L'exemple du mildiou : le mildiou peut entraîner la mort du tournesol au stade « plantule ». Dans ce cas, l'impact d'autres bioagresseurs est négligeable. En revanche , lorsque les plantes infectées survivent (plantes nanifiées), elles peuvent être touchées par l'orobanche. Cette interaction, rarement observée, n'a pas été étudiée (C. Jestin, communication personnelle).

— L'exemple de l'orobanche et du sclérotinia : le sclérotinia n'est pas spécifique du tournesol et il peut aussi attaquer l'orobanche (Ding *et al.*, 2012).

— De manière générale, la présence de l'orobanche et celles des maladies principales du tournesol ne sont pas antagonistes (C. Jestin, comm. pers.). Il n'est cependant pas possible de quantifier le nombre de cas avec orobanche et autres maladies associées pour les raisons évoquées précédemment.

4.5 Couplage des modèles SimMat, IPSIM et Sunflower-Pest pour la prédiction d'une partie du cycle épidémique, du profil de dégâts et des dommages associés

Une des limites de la méthode IPSIM est qu'elle s'arrête au profil de dégâts et ne prend pas en compte les dommages. Tous les dégâts ne provoquent pas des dommages. Ceux-ci résultent d'interactions complexes entre les bioagresseurs présents et le fonctionnement écophysiologique de la culture. Les modèles IPSIM mis en place dans cette thèse cherchent à prévoir uniquement des dégâts engendrant des dommages (par exemple pour le phomopsis la présence de taches encerclantes). D'autres modèles ont été créés pour prédire les dommages et leur couplage avec IPSIM pourrait permettre d'aller jusqu'à la prédiction des dommages via les dégâts. La plateforme informatique ModelBuilder (`xpest.inra.fr`) a été conçue dans le cadre du projet européens PURE et du projet français PROMOSOL (Aubertot *et al.*, 2016a,b). Cette plateforme permet aux chercheurs de développer librement des modèles quantitatifs dynamiques faisant intervenir les variables de leur choix. En particulier les modèles X-Pest ont été conçus sous cette plateforme. Ces modèles reposent sur un modèle générique écophysiologique simple de croissance des cultures. Sa finalité est de prédire les dommages en fonction des stress biotiques et abiotiques subis par la culture. Il se base sur l'équation de Monteith (Monteith et Moss, 1977) :

$$RG(t) = RUE.(1 - e^{-k.LAI(t)}).\varepsilon_c.RAD.dt \tag{4.1}$$

Où

— RG est le taux de croissance de la culture au cours du temps (en g.m^{-2}.j^{-1}). C'est la quantité de biomasse créée à chaque pas de temps.

— RUE (Radiation Use Efficiency) ou efficience biologique est une variable dynamique représentant la biomasse créée par unité de radiation reçue (g.MJ^{-1}) par la culture. La RUE est une variable synthétisant les processus physiologiques de la plante. Elle varie avec le stade phénologique de la culture (Trapani *et al.*, 1992 ; Lecoeur *et al.*, 2011 ; Garofalo et

Rinaldi, 2015), le génotype (Lecoeur *et al.*, 2011), la conduite de culture (l'apport en azote, l'irrigation ; Massignam *et al.*, 2009 ; Garofalo et Rinaldi, 2015).

— k le coefficient d'interception de la lumière (s.u.).

— LAI (Leaf Area Index) est le ratio de la surface totale des feuilles à la surface du sol sur laquelle la végétation se développe (s.u.).

— ε_c la part du rayonnement photo-synthétiquement actif

— RAD, la radiation globale ($MJ.m^{-2}.j^{-1}$).

— dt, le pas de temps en jour

Le modèle générique basé sur l'équation de Monteith a servi de base pour créer Sunflower-Pest, un modèle permettant de modéliser de manière quantitative et dynamique l'évolution de la biomasse du tournesol. Sunflower-Pest fonctionne à l'aide de compartiments décrivant le climat (la température rythme la croissance depuis l'émergence), le stade de développement de la culture (DVS), la biomasse de la culture (calculée avec RG), les différents organes de la culture (racines, tiges, feuilles, capitules puis graines). Chaque compartiment est caractérisé par une ou plusieurs variables, des paramètres constants ou dynamiques et les liens qu'ils ont avec les autres compartiments (Figure V.2). Ces liens sont matérialisés par des flèches (Figure V.2).

Ces liens permettent de générer des flux de données quantitatives entre les compartiments à chaque pas de temps. Ceci permet de faire évoluer le stade de développement de la culture, le pool de biomasse créée et la répartir dans chaque organe grâce à des coefficients de partition dynamiques.

La variable de LAI de l'équation de Monteith est simulée par les feuilles grâce à l'équation suivante.

$$LAI(t) = BM_{leaves}(t).SLA \tag{4.2}$$

Où la SLA (en $m^2.g^{-1}$) est un paramètre dynamique représentant la biomasse volumique de la feuille au cours de la croissance de la culture. BM_{leaves} est la biomasse des feuilles simulée de la manière suivante :

$$BM_{leaves}(t) = BM_{leaves}(t-1) + (CPL.RG(t-1) - RSENL).dt \tag{4.3}$$

Avec

— BM_{leaves} est la quantité de biomasse des feuilles créée.

— CPL est le taux de croissance de la culture au cours du temps (en g.m^{-2}.j^{-1}).

— RG la biomasse générale créée.

— $RSENL$ est le coefficient dynamique de sénescence des feuilles.

— dt, le pas de temps en jour

Les simulations démarrent au stade de la plantule. Une valeur initiale est donc assignée à la biomasse des feuilles et au pool de biomasse.

Sunflower-Pest a servi de base pour développer un modèle de nuisibilité pour le phomopsis, le phoma, le mildiou, le sclérotinia du collet, l'orobanche et le verticillium pour le tournesol (A. Costa, 2017). Ces processus sont complexes à modéliser car il existe plusieurs types de pertes (nombre de graines, taille des graines, teneur en huile). Ce sont les pertes engendrées par les symptômes les plus nuisibles qui ont été modélisées. En particulier, c'est la perte de biomasse des graines qui a été ciblée dans le modèle. Les stress biotiques et abiotiques peuvent agir à différents niveaux en fonction des organes touchés. Ces stress vont être modélisés par des fonctions de dommages et formalisés au travers d'équations faisant intervenir des coefficients réducteurs des variables décrivant le fonctionnement écophysiologique de la culture. Par exemple, lorsqu'un bioagresseur attaque une plante, il peut diminuer le LAI, la RUE, puiser dans la biomasse journalière créée par la plante (non accumulée dans les organes), ou dans les organes.

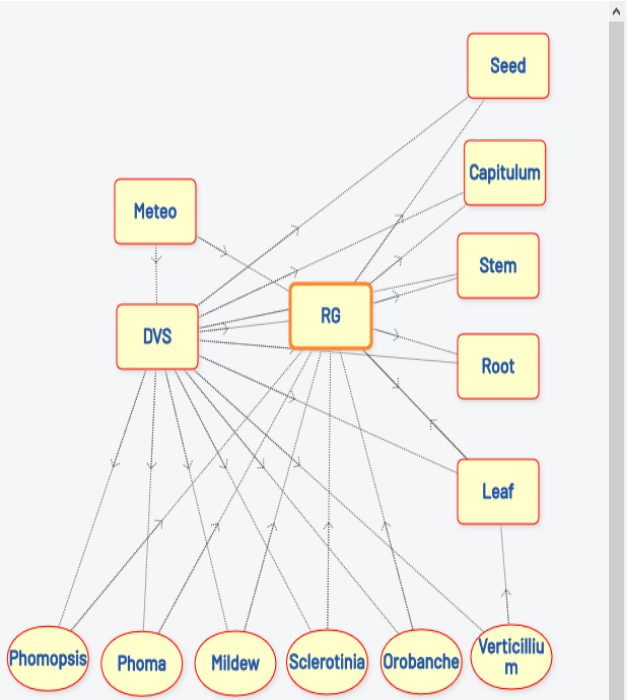
Des dynamiques génériques de bioagresseurs, associées à un profil de dégât donné simulé par IPSIM, pourraient être utilisées pour réaliser le couplage.

Cependant, le couplage IPSIM-tournesol et Sunflower-Pest se heurte à plusieurs difficultés :

— Les modèles IPSIM sur tournesol pour le mildiou, le sclérotinia et la verticilliose doivent être créés (réalisation ou mise à jour de synthèses bibliographiques et/ou de l'arbre). Les modèles pour le phoma et l'orobanche doivent être évalués et/ou améliorés.

— Le manque de jeu de données. Pour le moment, Sunflower-Pest n'intègre que le phomopsis ($Biais$ = 0,31%, MSE = 6,86% et EF= 0,99), le sclérotinia du collet ($Biais$ = -0,90%, MSE = 3,66% et EF = 0,91) et l'orobanche ($Biais$ = -0,12%, MSE = 3,53% et EF = 0,64). Les évaluations des modèles de nuisibilité de chaque bioagresseur pris individuellement ont été réalisées sur la base de très petits jeux de données : 4 observations pour le phomopsis par exemple (étude préliminaire d'A. Costa de l'EI Purpan). Des jeux de données doivent être crées pour évaluer les modèles de nuisibilité de manière robuste. Sunflower-Pest a déjà fait l'objet de communications pendant lesquelles les acteurs de la filière tournesol (agronomes, semenciers, modélisateurs...) ont été invités à mettre en place, à recueillir et à partager des

jeux de données.



Expressions de variables

Variable	Expression	Init.	Unité
RG	=(ec*RAD*RUE*RFphomopsis*RFphoma*RFsclero_collet*RFsclero_capitule*RFmildew*RForobanche*(1-exp(-k*LAI)))*dt	1	g.m-2

Paramètres

Paramètre	Valeur	Unité
dt	= 1	day
k	= 0.948	Dimensionless
ec	= 0.48	Dimensionless

Paramètres et variables dynamiques

Variable	Temps	Fichier	Unité
RUE	= f(DVS)	RUEm.txt	g.MJ-1

Liens

Variable	Modèle	Mode	Unité
DVS	ModelDVS	synchrone	Dimensionless
LAI	ModelLeaf	synchrone	m2.m-2
RAD	ModelMeteo	synchrone	MJ.m-2.day-1
RFphomopsis	Phomopsis	synchrone	Dimensionless
RFphoma	Phoma	synchrone	Dimensionless
RForobanche	Orobanche	synchrone	Dimensionless
RFmildew	Mildew	synchrone	Dimensionless
RFsclero_collet	Sclerotinia	synchrone	Dimensionless

Figure V.2 – Capture d'écran du modèle Sunflower-Pest sous la plateforme ModelBuilder (`xpest.inra.fr`). À gauche, le modèle Sunflower-Pest avec des rectangles représentants les variables météorologiques, le stade de développement de la culture, le pool de biomasse créée et les compartiments de la culture. Les ovales représentent les bioagresseurs du tournesol. Les flèches en pointillés représentent les processus biologiques (effets des variables climatiques sur la physiologie de la culture et les transferts de biomasse). À droite, le tableau de bord permet de définir les variables, les paramètres avec leurs équations, leurs valeurs et leurs unités. Les captures d'écran sont issues des travaux de M.A. Vedy-Zecchini (2016) et A. Costa (2017).

5. Conclusion

Cette thèse a permis diverses avancées pour la protection du tournesol, et des avancées méthodologiques pour la modélisation en général. Les avancées agronomiques correspondent à un premier jalon pour le développement du modèle IPSIM-Tournesol. Ce modèle permettra d'aider la gestion concomitante d'un ensemble de bioagresseurs (intégration horizontale de la PIC), en mobilisant un ensemble de techniques (intégration verticale de la PIC). Ce modèle pourra servir d'Outil d'Aide à la Décision pour les agriculteurs et leurs conseillers. Ceux-ci pourront étudier les effets *ex post* de leurs pratiques sur les stress biotiques lors de diagnostics ou bien mobiliser le modèle *ex ante*, pour les aider à concevoir des stratégies de conduite du tournesol *ad hoc*. D'ores et déjà, les modèles développés peuvent être utilisés par les collègues de Terres Inovia pour les aider dans leur(s) conseil(s) aux agriculteurs avec les modèles dont la qualité de prédiction est jugée satisfaisante (IPSIM-Tournesol-Phomopsis et SimMat-*Diaporthe* pour la prévision des classes de précocité). Une phase de coexistence entre SimMat-*Diaporthe* et le modèle historique ASPHODEL est prévue pour la rédaction des BSV. Ceci permettra que les utilisateurs s'approprient ce nouveau modèle, non seulement en terme d'utilisation technique, mais aussi pour leur laisser le temps de prendre confiance vis-à-vis de sa qualité de prédiction. Les autres modèles présentés dans la thèse peuvent également jouer le rôle de support d'animation lors de réunions techniques, en attendant que de nouvelles données permettent de les améliorer. L'approche de modélisation IPSIM présente l'avantage d'être adaptée aux enjeux de la PAEC, ainsi qu'à sa dimension participative. En effet, elle permet d'intégrer un très grand nombre d'informations aussi bien issues de la littérature, de jeux de données, de modèles préexistants, que de l'expertise de chercheurs, de techniciens, voire d'agriculteurs. Elle peut-être mobilisée pour tout bioagresseur et commence à être utilisée par

différentes équipes de par le monde. Les avancées méthodologiques présentées dans la thèse sont désormais mises à disposition des équipes développant des modèles IPSIM et disposant de jeux de données *ad hoc*.

Bibliographie

Bibliographie

A

Abu-Irmaileh, B. (1994). Nitrogen reduces branched broomrape (*Orobanche ramosa*) seed germination. *Weed Science*, 42:57–60.

Aggarwal, P., Kalra, N., Chander, S., and Pathak, H. (2006). InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. *Agricultural Systems*, 89:1–25.

Agreste (2019). Pratiques culturales en grandes cultures 2017. IFT et nombre de traitements. *Pratiques culturales en grandes cultures. Agreste Chiffres et Données*, 2019-3:1–30.

Agresti, A. and Kateri, M. (2010). *Analysis of ordinal categorical data. Matched-pairs data with ordered categories*. John Wiley & Sons, New Jersey, 2nd edition.

Agrios, G. (1936). *Plant pathology*. Science, Elsevier, 5th edition.

Ahmad, R., Waraich, E., Ashraf, M., Ahmad, S., and Aziz, T. (2014). Does nitrogen fertilization enhance drought tolerance in sunflower? A review. *Journal of Plant Nutrition*, 37:942–963.

Ainsworth, E., Leakey, A., Ort, D., and Long, S. (2008). FACE-ing the facts: inconsistencies and interdependence among field, chamber and modeling studies of elevated (CO_2) impacts on crop yield and food supply. *New Phytologist*, 179:5–9.

Alignan, M., Hewezi, T., Petitprez, M., Dechamp-Guillaume, G., and Gentzbittel, L. (2006). A cDNA microarray approach to decipher sunflower (*Helianthus annuus*) responses to the necrotrophic fungus *Phoma macdonaldii*. *New Phytologist*, 170:523–536.

An, Y., Ma, Y., Shui, J., and Zhong, W. (2015). Switchgrass (*Panicum virgatum L.*) has ability to induce germination of *Orobanche cumana*. *Journal of Plant Interactions*, 10(1):142–151.

Andrianasolo, F., Casadebaig, P., Champolivier, L., Maza, E., Maury, P., and Debaeke, P. (2014). Prediction of sunflower grain oil concentration as a function of variety, crop management and environment by the means of statistical models. *European Journal of Agronomy*, 54:84–96.

Andrianasolo, F., Debaeke, P., Champolivier, L., and Maury, P. (2016). Analysis and modeling of the factors controlling seed oil content in sunflower: a review. *OCL - Oleagineux Corps Gras Lipides*, 23(2):D206.

Anonyme (2014). Loi n 2014-1170 du 13 octobre 2014 d'avenir pour l'agriculture, l'alimentation et la forêt.

Anonyme (2015). *Plan ECOPHYTO II.* Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt et de l'Energie. Ministère de l'Ecologie, du Developpement Durable.

Antonova, T., Alonso, L., Strl'nikov, E., and Araslanova, N. (2015). Stimulating effect of the root exudates of sorghum, millet, and Sudan grass ont the seed germination of broomrape (*Orobanche cumana Wallr.*) infesting sunflowers in Russia. *Russian agricultural sciences*, 41(5):347–351.

Attoumani-Ronceux, A., Aubertot, J.-N., Guichard, L., Jouy, L., Mischler, P., Ornon, B., Petit, M., Pleyber, E., Reau, R., and Seiler, A. (2010). *Guide pratique pour la conception de systèmes de culture plus économes en produits phytosanitaires. Application aux systèmes de polyculture.* chargés de l'agriculture et de L'environnement, Ministères, rmt sdci edition.

Aubertot, J., Barbier, J., Carpentier, A., Gril, J., Guichard, L., L.as, P., Savary, S., Savini, I., and Voltz, M. (2005). *Expertise scientifique collective INRA - Cemagref. Réduire l'utilisation des pesticides et en limiter les impacts environnementaux Rapport d'expertise*. Aubertot, J.N. and Barbier, J.M. and Carpentier, A. and Gril, J.J. and Guichard, L. and L.as, P. and Savary, S. and Savini, I. and Voltz, M.

Aubertot, J., Guichard, L., Jouy, L., Mischler, P., Omon, B., Petit, M., Pleyber, É., Reau, R., and Seiler, A. (2010). *Guide pratique pour la conception de systèmes de culture plus économes en produits phytosanitaires. Application aux systèmes de polyculture*. Ministère de l'écologie, du développement durable, des transports et du logement et Ministère de l'agriculture, de l'alimentation, de la pêche, de la ruralité et de l'aménagement du territoire.

Aubertot, J., Robaldo, G., Vedy-Zecchini, M., and Mestries, E. (10 février 2016a). Projet SUNFLOWER-PEST.

Aubertot, J., Robaldo, G., Vedy-Zecchini, M., and Mestries, E. (26 may 2016b). XPEST : plateforme de modélisation en ligne pour la création et le développement de modèles de nuisibilité. In *Journée d'animation du réseau des utilisateurs de RECORD*, Avignon, France.

Aubertot, J. and Robin, M. (2013). Injury Profile SIMulator, a qualitative aggregative modelling framework to predict crop injury profile as a function of cropping practices, and the abiotic and biotic environment. I. Conceptual bases. *PLoS ONE*, 8(9):e73202.

Aubertot, J., Salam, M., Diggle, A., Dakowska, S., and Jedryczka, M. (2006). SimMat, a new dynamic module of blackleg sporacle for the prediction of pseudothecial maturation of *L. maculans*/ *L. biglobosa* species complex. Parameterisation and evaluation under Polish conditions. *Integrated Control in Oilseed Crops IOBC/wprs Bulletin*, 29(7):277–285.

Aubry, C., Papy, F., and Capillon, A. (1998). Modeling decision-making processes for annual crop management. *Agricultural Systems*, 56(1):45–65.

Aujas, P., Lacroix, A., Lemarié, S., and Reau, R. (2011). Réduire l'usage des pesticides. Un défi pour le conseil aux agriculteurs. *Économie rurale*, 324:18–33.

B

Baeyaert, A. (2012). Étude sur l'extension et la lutte contre l'*Orobanche cumana* en France. Memoire de fin d'étude. Technical report, École d'ingénieurs de Purpan, Toulouse.

Bazerque, J. (2015). Rapport de stage : Caractérisation de la diversité de populations françaises d'*Orobanche cumana*. Technical report, INRA-LIPM, Auzeville-Tolosane.

Beckert, M., Dessaux, Y., Charlier, C., Darmency, H., Richard, C., Savini, I., and Tibi, A. (2011). Les variétés vegetales tolérantes aux herbicides. Effets agronomiques, environnementaux, socio-economiques. Expertise scientifique collective, rapport. Technical report, CNRS-INRA, France.

Bergez, J., Debaeke, P., Deumier, J., Lacroix, B., Leenhardt, D., Leroy, P., and Wallach, D. (2001). MODERATO: an object-oriented decision tool for designing maize irrigation schedules. *Ecological Modelling*, 137(1):43–60.

BIO IS (2010). Analyses de cycle de vie appliquées aux biocarburants de première génération consommes en France. Rapport final. Technical report, Coordination ADEME.

Blackwell, M., Vilgalys, R., James, T., and Taylor, J. (2012). Fungi. Eumycota: mushrooms, sac fungi, yeast, molds, rusts, smuts, etc...
`http://tolweb.org/Fungi/2377/2012.01.30`. Version 30 January 2012.

Blanchet, R. and Merrien, A. (1990). *Le tournesol et l'eau : adaptation à la sécheresse, réponse à l'irrigation*. Blanchet, R. and Merrien, A., Paris, cetiom edition.

Boari, A. and Vurro, M. (2004). Evaluation of *Fusarium spp.* and other fungi as biological control agents of broomrape (*Orobanche ramosa*). *Biological control*, 30(2):212–219.

Bohanec, M. (2015). *DEXi: program for multi-attribute decision making user's manual*. Ljubljana, Slovenija.

Bohanec, M., Žnidaržic, M., Rajkovic, V., Bratko, I., and Zupan, B. (2013). DEX methodology: Three decades of qualitative multi-attribute modeling. *Informatica*, 37(1):49–54.

Boote, K., Jones, J., and Pickering, N. (1996). Potential uses and limitations of crop models. *Agronomy Journal*, 88(5):704–716.

Bordat, A., Debaeke, P., Dechamp-Guillaume, G., Mestries, E., Seassau, C., and Vincourt, P. (2011). *Phoma et Dessèchement précoce du tournesol*. CETIOM, Paris, cetiom edition.

Bouveris, N. (1980). Étude du cycle primaire de la tavelure du pommier dans la région "midi-pyrénées" en vue de la mise au point d'un modèle devant servir aux avertissements agricoles. *Avertissements agricoles et lute intégrée. Service de la protection des végétaux n° 8*, pages 61–68.

Breiman, L., Friedman, J., Stone, C., and Olshen, R. (1984). *Classification and Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M., Ruget, F., Bernard, N., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., H., M., Jeuffroy, F., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J., Meynard, J., and Delecolle, R. (1998). STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie*, 18(5):311–346.

C

Caffi, T., Rossi, V., Cossu, A., and Fronteddu, F. (2007). Empirical vs. mechanistic models for primary infections of *Plasmopara viticola*. *Bulletin OEPP/EPPO*, Bulletin 3:247–254.

Calchei, E. (2011). Methods of combating broomrape in tobacco plantations. In *International symposium on broomrape (Orobanche spp.) in sunflower*, page 15, Chisinau, Moldavie.

Carson, M. (1991). Relationship between phoma black stem severity and yield losses in hybrid sunflower. *Plant disease*, 75(11):1150–1153.

Casadebaig, P., Guilioni, L., Lecoeur, J., Christophe, A., Champolivier, L., and Debaeke, P. (2011). SUNFLO, a model to simulate genotype-specific performance of the sunflower crop in contrasting environments. *Agricultural and Forest Meteorology*, 151(2):163–178.

Cechin, C. and Press, M. (1993). Nitrogen relations of the Sorghum–*Striga hermonthica* host–parasite association: germination, attachment and early growth. *The New Phytologist*, 124:681–687.

Cerrutti, N. and Pontet, C. (2016). Differential attractiveness of sunflower cultivars to the honeybee *Apis mellifera L. Oilseeds & fats Crops and Lipids*, 23(2):7.

Chabrolin (1938). La germination des graines de l'*Orobanche speciosa*. Comptes rendus de Biologie.

Champolivier, L., Debaeke, P., and Merrien, A. (2011). Pourquoi irriguer le tournesol, une culture réputée tolérante a la sécheresse ? TT - Why to irrigate sunflower, a drought tolerant crop? *Innovations Agronomiques*, 14:151–164.

Cheng, W., Sims, D., Luo, Y., Coleman, J., and Johnson, D. (2000). Photosynthesis, respiration, and net primary production ofsunflower stands in ambient and elevated atmospheric CO2 concentrations: an invariant NPP:GPP ratio? *Global Change Biology*, 6:931–941.

Chimenti, C., Hall, A., and López, M. (2001). Embryo-growth rate and duration in sunflower as affected by temperature. *Field Crops Research*, 69:81–88.

Chimenti, C., Pearson, J., and Hall, A. (2002). Osmotic adjustment and yield maintenance under drought in sunflower. *Field Crops Research*, 75:235–246.

Chollet, D., Estragnat, A., Jouffret, P., Lagarde, F., Mestries, E., Penaud, A., and Pinochet, X. (2002). *Les maladies du tournesol*. CETIOM, Paris.

Connor, D. and Hall, A. (1997). Sunflower physiology. In *Sunflower technology and production. Agronomy monograph 35*, pages 113–182. Schneiter, A.A.

Coque, M., André, T., Gimenez, R., Archipiano, M., Polovynko, L., Tardin, M., Jestin, C., and Grezes-Besset, B. (2016). Study of *Orobanche cumana* genetic diversity. In *19th International Sunflower Conference*, page 759, Edirne, Turkey.

Corre, C., Gourdain, E., Grignon, G., Héritier, E., Manstretta, V., and Ross, V. (2015). Un modèle épidémiologique pour piloter le traitement contre *F. graminearum* sur céréales à paille : analyse de sensibilités sur les paramètres climatiques. In *AFPP – Onzième Conférence Internationale sur les Maladies des Plantes – 7 au 9 décembre*, pages 1–10, Tours, France.

Cox, P. (1996). Some issues in the design of agricultural decision support systems. *Agricultural Systems*, 52(2-3):355–381.

D

Debaeke, P. (2002). Des pistes pour prévoir les effets de la conduite du tournesol sur les maladies. *Oléoscope*, 67:23–25.

Debaeke, P., Bedoussac, L., Bonnet, C., Bret-Mestries, E., Seassau, C., Gavaland, A., Raffaillac, D., Tribouillois, H., VE.el, G., and Justes, E. (2017a). Sunflower crop: environmental-friendly and agroecological. *OCL*, 24(3)(D304).

Debaeke, P., Casadebaig, P., Flenet, F., and Langlade, N. (2017b). Sunflower crop and climate change: vulnerability, adaptation, and mitigation potential from case-studies in Europe. *OCL - Oilseeds and fats, Crops and Lipids*, 24(D102):15.

Debaeke, P., Délos, M., and Moinard, J. (2001). Simulating the influence of sowing date on disease occurrence in sunflower. In *Proceedings of the 2nd International ESA Symposium Modelling Cropping Systems*, pages 123–124, Florence, Italie.

Debaeke, P., Délos, M., Moinard, J., Berault, S., and Lambert, R. (2000). Prise en compte du couvert de tournesol dans la simulation des épidémies de *Diaporthe helianthi* par le modèle ASPHODEL. In *Proceedings of the Sixth International Conference on Plant Diseases, 6–8 décembre 2000*, pages 251–258, Tours, France.

Debaeke, P. and Estragnat, A. (2003). A simple model to interpret the effects of sunflower crop management on the occurrence and severity of a major fungal disease: Phomopsis stem canker. *Field Crops Research*, 83(2):139–155.

Debaeke, P. and Estragnat, A. (2009). Crop canopy indicators for the early prediction of Phomopsis stem canker (*Diaporthe helianthi*) in sunflower. *Crop Protection*, 28(9):792–801.

Debaeke, P., Estragnat, A., and Reau, R. (2003). Influence of crop management on sunflower stem canker (*Diaporthe helianthi*). *Agronomy*, 23(7):581–592.

Debaeke, P. and Moinard, J. (2010). Effect of crop management on epidemics of phomopsis stem canker (*Diaporthe helianthi*) for susceptible and tolerant sunflower cultivars. *Field Crops Research*, 115(1):50–60.

Debaeke, P. and Pérès, A. (2003). Influence of sunflower (*Helianthus annuus L.*) crop management on Phoma black stem (*Phoma macdonaldii Boerema*). *Crop Protection*, 22:741–752.

Debaeke, P. and Raffaillac, D. (1996). Light interception as an indicator of leaf area index and risk of diseases in sunflower. *Helia*, 19:1–15.

Deguine, J., Gloanec, C., L., P., Ratnadass, A., and Aubertot, J. (2016). *Décliner les principes de l'agroecologie a la protection des cultures: les contours de l'agroécologie.*, page 288. Versailles, France, quae edition.

Del Grosso, S., Parton, W., Keough, C., and Reyes-Fox, M. (2011). Special features of the daycent modeling package and additional procedures for parameterization, calibration, validation, and applications. *Methods of Introducing System Models into Agricultural Research*, pages 155–176.

Delaplane, K. and Mayer, D. (2000). *Crop pollination by bees*. CABI Publishing, New York, USA.

Desanlis, M. (2013). *Analyse et modelisation des effets de la conduite de culture sur deux maladies cryptogamiques majeures du tournesol causees par Phoma macdonaldii et Phomopsis helianthi*. PhD thesis, Institut National Polytechnique de Toulouse.

Desanlis, M., Aubertot, J., Mestries, E., and Debaeke, P. (2013). Analysis of the influence of a sunflower canopy on *Phomopsis helianthi* epidemics as a function of cropping practices. *Field Crops Research*, 149:63–75.

Descorps, C. (2010). Phoma du collet et symptomes "pieds secs" du tournesol : etiologie, épidémiologie et mise au point de méthodes. Memoire Master Bioprotection et Valorisation Non-alimentaire de Agrosystèmes. Technical report, Universite de Pau, Pau.

Ding, L., Zhao, S., Zhang, X., Yao, Z., and Zhang, J. (2012). Sclerotinia rot of broomrape (*Orobanche cumana*) caused by *Sclerotinia sclerotiorum* in China. *Plant disease*, 96(6):196.

Dominguez, J. (1996). R-41, a sunflower restorer inbred line, carrying two genes for résistance against a highly virulent spanish population of *Orobanche cernua*. *Plant breeding*, 115:203–204.

Donald, P., Hartman, C., Venette, J., and Secor, G. (1987). Response of sunflowers to culture filtrates from *Phoma macdonaldii*. *Phytopathology*, 75:466–468.

Dubrulle, P., Dupont, A., Publicol, M., Rousse, N., Baratte, C., Sohbi, Y., and Charron-Moirez, M. (2014). Rapport Outils d 'Aide à la Décision. Technical report, INRA.

Duroueix, F. and Penaud, A. (2013). Les fongicides en renfort contre phoma et phomopsis. *Perspectives Agricoles*, 400:18–21.

Délos, M. and Moinard, J. (1997). ASPHODEL : modèle de simulation des épidémies de phomopsis du tournesol (*Diaporthe helianthi*). In *Vme Conférence Internationale sur les maladies des plantes - ANPP*, pages 413–420, Tours, France.

Délos, M., Moinard, J., and Mühlberger, E. (1998). Le phoma du tournesol. Mise au point d'une méthode d'évaluation des risques de contaminations par les ascospores. *Phytoma - La Défense des Végétaux*, 504:26–29.

E

Edminster, T. (1978). *Concepts for using modeling as a research tool*. U.S. Department of Agriculture, Beltsville, MD, USA.

Eizenberg, H., Hershenhorn, J., Achdari, G., and Ephrath, J. (2012). A thermal time model for predicting parasitism of *Orobanche cumana* in irrigated sunflower—Field validation. *Field crops research*, 137:49–55.

Eizenberg, H., Hershenhorn, J., and Ephrat, J. (2009). Factors influencing the efficacy of *Orobanche cumana* chemical control in sunflower. *Weed research*, 49:308–315.

Ephrath, J. and Eizenberg, H. (2010). Quantification of the dynamics of *Orobanche cumana* and *Phelipanche aegyptiaca* parasitism in confectionery sunflower. *Weed Research*, 50:140–152.

Estragnat, A. (1992). Phomopsis du tournesol. *Oléoscope*, 8:8–10.

F

Fabregue, C. (1992). Phomopsis du tournesol. *Oléoscope*, 8:11.

Fannader, R. and Leroux, H. (1999). *UML. Principes de modelisation (Informatiques)*. Dunod.

Fayret, J. and Assemat, P. (1987). Évolution du *Diaporthe helianthi* (*Phomopsis helianthi* Munt.-Cvet. *et al.*) et différenciation des organes reproducteurs sur les plants de tournesol après la période de végétation. *Informations Techniques CETIOM*, 98:2–11.

Fernandez-Aparicio, M., Flores, F., and Rubiales, D. (2009). Recognition of root exudates by seeds of broomrape (*Orobanche* and *Phelipanche*) species. *Annals of Botany*, 103:423–431.

Fernandez-Aparicio, M., Reboud, X., and Gibot-Leclerc, S. (2016). Broomrape weeds. Underground mechanisms of parasitism and associated stratégies for their control: a review. *Frontiers in Plant Science*, page 7:135.

Fleiss, J. and Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33:613–619.

G

Gadoury, D. M. and MacHardy, W. E. (1982). A Model to Estimate the Maturity of Ascospores of *Venturia inaequalis*. *Phytopathology*, 72(7):901–904.

Garofalo, P. and Rinaldi, M. (2015). Leaf gas exchange and radiation use efficiency of sunflower (*Helianthus annuus L.*) in response to different deficit irrigation Stratégies: From solar radiation to plant growth analysis. *European Journal of Agronomy*, 64:88–97.

Gaudin, Z. (2013). *Place dans l'azote dans l'interaction plante- plante parasite: Brassica napus L.-Phelipanche ramosa (L.) Pomel*. PhD thesis, Universite de Nantes.

Gaunt, J., Riley, J., Stein, A., and Penning de Vries, F. (1997). Requirements for effective modelling strategies. *Agricultural Systems*, 54(2):153–168.

Geiger, F., Bengsson, J., Berendse, F., Weisser, W., Emmerson, M., Morales, M., Ceryngier, P., Liira, J., Tscharntke, T., Winqvist, C., Eggers, S., BomM.o, R., Pärt, T., Bretagnolle, V., Plantegenest, M., Clement, L., Dennis, C., Palmer, C., Onate, J., Guerrero, I., Hawro, V., Aavik, T., Thies, C., Flohre, A., Hänke, S., Fischer, C., Goedhart, P., and Inchausti, P. (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic Applied Ecology*, 11:97–105.

Ghersa, C. and Martinez-Ghersa, M. (2000). Ecological of weed seed size and persistence in the soil under different tilling systems: implications for weed management. *Field Crops Research*, 67:141–148.

Gini, C. (1921). Measurement of inequality of income. *Economic Journal*, 31:22–43.

Girard, N. and Hubert, B. (1999). Modelling expert knowledge with knowledge-based systems to design decision aids - The example of a knowledge-based model on grazing management. *Agricultural Systems*, 59(2):123–144.

Goldwasser, Y. and Rodenburg, J. (2013). *Integrated management of parasitic weed seed banks*, pages 393–413. D.M., Joel.

Gorry, A. and Scott-Morton, M. (1971). A Framework for Information Systems. *Sloan Management Review*, 13(1):56–79.

Green, J. (2014). Current state of herbicides in herbicide-resistant crops. *Pest Management Science*, 70:1351—1357.

Grenz, J., Istoc, V., Manschadi, A., and Sauerborn, J. (2008). Interactions of sunflower (*Helianthus annuus*) and sunflower broomrape (*Orobanche cumana*) as affected by sowing date, resource supply and infestation level. *Field crops research*, 107:170–179.

Grenz, J., Manschadi, A., de Voil, P., Meinke, H., and Sauerborn, J. (2005). Assessing stratégies for *Orobanche sp.* control using a combined seed bank and competition model. *Agronomy journal*, 97:1551–1559.

Guerin, O. (1997). Contribution a l'Étude biologique, epidemiologique de *Leptosphaeria lindquistii Frezzi* (*Phoma macdonaldii Boerema*), champignon parasite du tournesol. Memoire DAA Biotechnologies Amelioration et Protection des Plantes.

Gulya, T., Charlet, L., Donald, P., and Venette, J. (1984). What do we know about premature ripening? *The sunflower (USA)*, 10:20–21.

Gulya, T., Rashid, K., and Masirevic, S. (1997). *Sunflower Technology and Production*. Schneiter A.A., Madison, Wisconsin, USA.

H

Habimana, S., Murthy, K., Hatti, V., and Nduwumuremyi, A. (2013). Management of orobanche in field crops – a review. *Scientific journal of crop science*, 2(11)(11):144–158.

Haefele, S., Jabbar, S., Siopongco, J., Tirol-Padre, A., Amarante, S., Cruz, P., and Cosico, W. (2008). Nitrogen use efficiency in selected rice (*Oryza sativa L.*) genotypes under different water regimes and nitrogen levels. *Field Crops Research*, 107:137–146.

Heemskerk, M., Wilson, K., and Pavao Zuckerman, M. (2003). Conceptual models as tools for communication across disciplines. *Conservation Ecology*, 7(38).

Huber, D. and Gillespie, T. (1992). Modelling leaf wetness in relation to plant disease epidemiology. *Annual Review of Phytopathology*, 30:553–577.

Huyer, W. and Neumaier, A. (1999). Global optimization by multilevel coordinate search. *Journal of Global Optimization*, 14:331–355.

I

Iliescu, H. and Baicu, T. (1984). Mesures de lutte intégrées contre les agents pathogènes du tournesol. *Helia*, 7:39–42.

Ish-Shalom-Gordon, N., Jacobson, R., and Cohen, Y. (1993). Inheritance of resistance to *Orobanche cumana* in sunflower. *Phytopathology*, 83:1250–1252.

J

Jain, R. and Foy, C. (1992). Nutrient effects on parasitism and germination of Egyptian broomrape (*Orobanche aegyptiaca*). *Weed Technology*, 6:269–275.

Jestin, C. (2016). Orobanche cumana : synthèse bibliographique. Document interne Terres Inovia. 97 p.

Jestin, C., Lecomte, V., and Duroueix, F. (2014). Current situation of sunflower broomrape in France. In *Proceeding 3rd International Symposium On Broomrape (Orobanche spp.) in Sunflower*, pages 28–31, Córdoba, Spain.

Jestin, C. and Martin-Monjaret, C. (2016). *Orobanche cumana* : opter pour la bonne stratégie de lutte. *Perspectives agricoles*, 430:28–31.

Johnson-Laird, P. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge University Press, Cambridge.

Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., and Ritchie, J. (2003). The DSSAT Cropping System Model. *European Journal of Agronomy*, 18:235–265.

Jouffret, P. (2005). Stratégie de lutte durable contre phomopsis et phoma. *Oléoscope n°80*, pages 28–30.

Jouffret, P., Labalette, F., and Thibierge, J. (2011). Atouts et besoins en innovations du tournesol pour une agriculture durable. *Innovations agronomiques*, 14:1–17.

K

Karkanis, A., Bilalis, D., and Efthimiadou, A. (2007). Tobacco (*Nicotiana tabaccum*) infection by branched broomrape (*Orobanche ramosa*) as influenced by irrigation system and fertilization, under east mediterranean conditions. *Journal of Agronomy*, 6(3):397–402.

Kebreab, E. and Murdoch, A. (1999). A quantitative model for loss of primary dormancy and induction of secondary in imbibed seeds of *Orobanche spp*. *Journal of Experimental Botany*, 50:211–219.

Klein, O. and Kroschel, J. (2002). Biological control of *Orobanche spp.* with *Phytomyza orobanchia*, a review. *BioControl*, 47(3):245–277.

Krupp, A., Rücker, E., Heller, A., and O., S. (2015). Seed Structure Characteristics of *Orobanche cumana* Populations. *Helia*, 38(62):1–14.

L

Labrousse, P. (2002). *Contribution a l'Étude de la résistance de differents genotypes d'Helianthus (Asteracees) a Orobanche cumana Wallr. (Orobanchacees)*. PhD thesis, Université de Nantes.

Lamarque, C., Tourvieille, D., Regnault, T. Y., and Pierre, J. C. (1985). Les problèmes posés par la lutte chimique contre le *Sclerotinia sclerotiorum* sur capitule de tournesol. *ANPP - Premières Journées d'études sur les maladies des plantes.*, 2:359–367.

Lamichhane, J., Dachbrodt-Saaydeh, S., Kudsk, P., and Messean, A. (2016). Toward a reduced reliance on conventional pesticides in european agriculture. *Plant Disease*, 100(1):24.

Lamichhane, J., Messean, A., and Ricci, P. (2019). Research and innovation priorities as defined by the Ecophyto plan to address current crop protection transformation challenges in France. *Advances in Agronomy*, 154:81–152.

Lancashire, P., Bleiholder, H., Tvd., B., P., L., Stauss, R., and *et al.* (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 119:561–601.

Landis, J. and Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33:159 –174.

Lecoeur, J., Poire-Lassus, R., Christophe, A., Pallas, B., Casadebaig, P., Debaeke, P., Vear, F., and Guilioni, L. (2011). Quantifying physiological determinants of genetic variation for yield potential in sunflower. SUNFLO: A model-based analysis. *Functional Plant Biology*, 38(3):246–59.

Lecomte, V. and Nolot, J. (2011). Place du tournesol dans le système de culture TT - Sunflower in cropping system. *Innovations Agronomiques*, 14:59–76.

Limasset, P. and Darpoux, H. (1950). *Principes de pathologie végétale*. INRA, Paris, dunod edition.

Logan, D. and Stewart, G. (1992). Germination of the seeds of parasitic angiosperms. *Seed Science research*, 2:179–190.

Louarn, J. (2013). *Exploitation des champignons mycorhiziens a arbuscules pour la protection du tournesol contre Orobanche cumana*. PhD thesis, Université Toulouse 3 Paul Sabatier.

Lovell, D., Powers, S., Welham, S., and Parker, S. (2004). A perspective on the measurement of time in plant disease epidemiology. *Plant Pathology*, 53:705–712.

Loyce, C. and Wéry, J. (2006). Les outils des agronomes pour l'évaluation et la conception des systèmes de culture. *Agronomy*, pages 77–95.

Lu, Y., Melero-Vara, J., Garcia-Tejada, J., and Blanchard, P. (2000). Development of SCAR markers linked to the gene Or5 conferring résistance to broomrape (*Orobanche cumana Wallr.*) in sunflower. *Theoretical and Applied Genetics*, 100:625–632.

Lynch, T., Gregor, S., and Midmore, D. (2000). Intelligent support systems in agriculture: how can we do better? *Australian Journal of Agricultural Research*, 40:609–620.

M

Ma, Y., Jia, J., An, Y., Wang, Z., and Mao, J. (2013). Potential of some hybrid maize lines to induce germination of sunflower broomrape. *Crop science*, 53:260–270.

Mariam, E. and Suwanketnikom, R. (2004). Effect of nitrogen fertilizers on branched broomrape (*Orobanche ramosa L.*) in tomato (*Lycopersicon esculentum Mill.*). *Nat Sci*, 38:311–319.

Maric, A., Camprag, D., and Masirevic, S. (1988). Sunflower diseases and pests and their control. pages 37–45.

Masirevic, S. and Gulya, T. (1992). Sclerotinia and Phomopsis — two devastating sunflower pathogens. *Field Crops Research*, 30(3-4):271–300.

Massey, J. (1971). Effects of nitrogen rates and plant spacing on sunflower-D seed yields and other characteristics. *Oléoscope*, 63:137–138.

Massignam, A., Chapman, S., Hammer, G., and Fukai, S. (2009). Physiological determinants of maize and sunflower grain yield as affected by nitrogen supply. *Field Crops Research*, 113(3):256–67.

Mathew, F., Harveson, R., Gulya, T., Thompson, S., Block, C., and Markell, S. (2018). Phomopsis stem canker of sunflower. *Plant Health Instructor*.

McDonald, W. (1964). Phoma black stem of suflower. *Phytopathology*, 54:492–493.

Mead, R. and Nelder, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

Meliala, C., Tourvieille de Labrouhe, D., and Pilorgé, E. (1999). Mildiou du tournesol : des symptômes très variés. *Oléoscope - Bulletin du CETIOM*, 50:14–16.

Meliala, C., Vear, F., and Tourvieille de Labrouhe, D. (2000). Secondary infection of sunflower downy mildew: symptomatology, epidemiology and its economic risks. In *Proceeding 15th International Sunflower Conference*, pages 1–6, Toulouse, France.

Merrien, A. and Grandin, G. (1990). Comportement hydrique du tournesol : synthèse des essais 'irrigation' 1983-88. In *Le tournesol et l'eau : adaptation à la sécheresse, réponse à l'irrigation*, pages 75–90. Cetiom, Paris, France.

Merrien, A. and Milan, M. (1992). *Physiologie du tournesol*. CETIOM, Paris.

Mestries, E., Desanlis, M., Seassau, C., Moinard, J., Debaeke, P., and Dechamp-Guillaume, G. (2011). Impact de la conduite de culture sur les maladies du tournesol. *Innovations Agronomiques*, 14:91–108.

Mihaljcevic, M., Muntanola-Cvetkovic, M., Vukojevic, J., and Petrov, M. (1985). Source of infection of sunflower plants by *Diaporthe helianthi* in Yugoslavia. *Journal of Phytopathology*, 113(4):334–342.

Moinard, J. and Eychenne, N. (1998). Le phomopsis. Guide "Grandes Cultures". Dossier support des Avertissements Agricoles routés et télécopiés. Technical report, DGAL-Service Régional de la Protection des Végétaux Midi-Pyrénées, dossier support des Avertissements Agricoles routes et télécopies.

Molinero-Ruiz, L., Delavault, P., Perez-Vich, B., Pacureanu-Joita, M., Bulos, M., Altieri, E., and Domínguez, J. (2015). History of the race structure of *Orobanche cumana* and the breeding of sunflower for resistance to this parasitic weed: A review. *Spanish Journal of Agricultural Research*, 13(4):19.

Monteith, J. and Moss, C. (1977). Climate and the efficiency of crop production in Britain [and Discussion]. Philosophical transactions of the royal society of London B. *Biological Sciences*, 281(980):277–94.

Moore, A. (2001). Cross-validation for detecting and preventing overfitting.

Morgan, J. (2014). Classification and Regression Tree Analysis.

Muhlberger, E. (1996). Phoma et phomopsis du tournesol : Synthèse des connaissances acquises. Contribution a l'étude biologique et epidemiologique. Rapport de fin d'études.

Muntanola-Cvetkovic, M., Vukojevic, M., and Mihaljcevic, M. (1988). Pathohistology of sunflower stems attacked by *Diaporthe helianthi*. *Canadian Journal of Botany*, 67:1119–1125.

Murdoch, A. and Kebreab, E. (2013). *Parasitic Orobanchaceae: Parasitic Mechanisms and Control Strategies*. Springer-Verlag Berlin Heidelberg.

N

Nasraoui, B. (2003). *Les champignons phytopathogènes, [Pathogenic fungi].*, pages 111–143. De boeck edition.

Nolot, J. and Debaeke, P. (2003). Principes et outils de conception, conduite et évaluation de systèmes de culture. *Cahiers Agricultures*, 12:387–400.

P

Panneton, B., Vincent, C., and Fleurat-Lessard, F. (2000). *La lutte physique en phytoprotection*. INRA Editions, Paris.

Parker, C. and Riches, C. (1993). *Parasitic weeds of the world: Biology and control*. Centre for Agricultural Bioscience International, Wallingford, UK.

Parts, E., Bazzalo, M., Leon, A., and Jorrin, J. (2003). Accumulation of soluble phenolic compounds in sunflower capitula correlates with resistance to *Sclerotinia sclerotiorum*. *Euphytica*, 132:321–329.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572.

Penaud, A., Maisonneuve, C., Pérès, A., and Pilorge, E. (1998). Tournesol : phomopsis et phoma : 1995 - 1996 - 1997. *Dossier Technique CETIOM*, page 62.

Penaud, A., Pérès, A., and Regnault, Y. (1994). Rappels sur la maladie. *Oléoscope*, 14:15–26.

Perez-de Luque, A., Fondevilla, S., Perez-Vich, B., Aly, R., Thoiron, S., Simier, P., Castillejo, M., Jorrin, J., Rubiales, D., and Delavault, P. (2009). Understanding *Orobanche* and *Phelipanche*-host plant interactions and developing résistance. *Weed research*, 49:8–22.

Perez-de Luque, A., Moreno, M., and Rubiales, D. (2008). Host plant resistance against broomrapes (*Orobanche spp.*) : defence reactions and mechanisms of résistance. *Annals of Applied Biology*, 152:131–141.

Perez-Vich, B., Akhtouch, B., Mateos, A., Velasco, L., Jan, C., Fernandez, J., Dominguez, J., and Fernandez-Martinez, J. (2004). Dominance relationships for genes conferring resistance to sunflower broomrape (*Orobanche cumana Wallr.*). *Helia*, 40:183–192.

Perny, R. (1992). L'évolution du phomopsis en France. *Oléoscope*, 8:7.

Petrov, M., Muntanola-Cvetkovic, M., and Mihaljcevic, M. (1981). New findings on the disease of sunflower plants caused by *Phomopsis helianthi Munt.-Cvet. et al.* (*Diaporthe helianthi* Munt.-Cvet *et al.*). *Archives des Sciences Biologiques (Beograd)*, 33:13–19.

Pilorgé, E. (2016). Sunflower: some examples of current research - Introduction. *Oilseeds & fats Crops and Lipids*, 23(2):1.

Pinochet, X. (1994). Le Phomopsis du tournesol. Bilan des travaux du CETIOM de 1993 à 1995. *Oléosope Numéro Special N14*, pages 27–35.

Pinochet, X. (2003). Phoma du tournesol. *Rencontres techniques régionales du Cetiom-Tournesol*, pages 30–40.

Poisson, A. (2015). GIS Fruits - Les Outils d'Aide à là Décision. Technical report, INRA.

Poisson-Bammé, B. and Pérès, A. (6-8 Décembre 2000). Survie du phoma du tournesol (*Leptosphaeria lindquistii*) sur les résidus de récolte. In *Annales 6me Conférence Internationale. Maladies des Plantes, AFPP*, pages 331–338.

Purdy, L. (1979). *Sclerotinia sclerotiorum* : history, diseases and symptomatology, host range, geographic distribution and impact. *Phytopathology*, 69:875–880.

Pérès, A. and Lefol, C. (1996). *Phoma macdonaldii Boerema* : éléments de biologie et mise au point d'une méthode de contamination artificielle en conditions contrôlées. *Annales 14me Conférence Internationale du Tournesol*, pages 687–693.

Pérès, A. and Pilorge, E. (1997). Tournesol : phomopsis et phoma. *Oléoscope*, 38:12.

Pérès, A., Poisson, B., and Drolon, G. (2000). Le syndrôme "pieds secs" du tournesol : étude des causes et approche de la nuisibilité. In *Annales 15ème Conférence Internationale du Tournesol ISA - 12-15 Juin*, volume 2(1), pages 17–21, Toulouse, France.

Pérès, A. and Poisson-Bammé, B. (1999). Pied sec du tournesol : les connaissances progressent. *Oléoscope*, 52:27–30.

Pérès, A. and Poisson-Bammé, B. (2000). The early withering of sunflower ("dry plant"): the part played by Phoma and study of the contamination (in French). In *Annales 6me Conférence Internationale. Maladies des Plantes, AFPP - 6-8 December.*, pages 323–329.

R

Rapilly, F. (1991). *L'épidémiologie en pathologie végétale - Mycoses aériennes*. INRA, Paris, France, inra edition.

Remidi, B. (2013). Phoma du tournesol : Caractérisation de l'impact du travail du sol sur le réservoir d'inoculum primaire de *Leptosphaeria lindquistii*. Technical report, Memoire Ingenieur AgroSup, Dijon, France.

Robin, M., Colbach, N., Lucas, P., Montfort, F., Cholez, C., Debaeke, P., and Aubertot, J. (2013). Injury Profile SIMulator, a qualitative aggregative modelling framework to predict an injury profile as a function of cropping practices, and abiotic and biotic environment. II. Proof of concept: design of IPSIM-Wheat-Eyespot. *PLoS ONE*, 8(10).

Rodriguez-Ojeda, M., Alonso, L., and Fernandez-Escobar, J. (2001). Effect of different crops on the germination of *Orobanche cernua Loefl.* (*O. cumana Wallr.*) seeds. In *Proceedings of the 7th international parasitic weed symposium. 5-8 juin*, page 124, Universite de Nantes, Nantes, France.

Román, B., Alfaro, C., Torres, A., Moreno, M., Satovic, Z., Pujadas, A., and Rubiales, D. (2003). Genetic relationships among *Orobanche* species as revealed by RAPD analysis. *Annals of Botany*, 91:637—-642.

Roustaee, A., Costes, S., Dechamp-Guillaume, G., and Barrault, G. (2000). Phenotypic variability of *Leptosphaeria lindquistii* (anamorph: *Phoma macdonaldii*), a fungal pathogen of sunflower. *Plant Pathology*, 49:227–234.

Rubiales, D., Fernandez-Aparicio, M., Wegmann, K., and Joel, D. (2009). Revisiting stratégies for reducing the seedbank of *Orobanche* and *Phelipanche spp*. *Weed research*, 49:23–33.

S

Salam, M., Khangura, R., Diggle, A., and Barbetti, M. (2003). Blackleg Sporacle: an aid to understanding and managing blackleg in canola (*Brassica napus*). *Phytopathology*, pages 2002–2004.

Saltelli, A., Chan, K., and Scott, E. (2000). *Sensitivity Analysis*. Wiley, New York.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saiana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer*. John Wiley & Sons Inc., john wiley & sons edition.

Salzberg, S. L. (1994). C4.5: Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993. *Machine Learning*, 16(3):235–240.

Savary, S., Elazegui, F., Moody, K., Litsinger, J., and Teng, P. (1994). Characterization of rice cropping practices and multiple pest systems in the Philippines. *Agricultural Systems*, 46(4):385–408.

Savary, S., Teng, P., Willocquet, L., and Nutter, F. (2006). Quantification and modeling of crop losses: A review of purposes. *Annual Review of Phytopathology*, 44:89–112.

Seassau, C. (2010). *Etiologie du syndrome de Dessèchement précoce du tournesol : implication de Phoma macdonaldii et interaction avec la conduite de culture*. PhD thesis, Institut National Polytechnique de Toulouse.

Seassau, C., Dechamp-Guillaume, G., Mestries, E., and Debaeke, P. (2010). Nitrogen and water management can limit ripening of sunflower induced by *Phoma macdonaldii*. *Field Crops Research*, 115:99–106.

Seassau, C., Dechamp-Guillaume, G., Mestries, E., and Debaeke, P. (2012). Low plant density can reduce sunflower premature ripening caused by *Phoma macdonaldii*. *European Journal of Agronomy*, 43:185–193.

Seassau, C., Mestries, E., Debaeke, P., and Dechamp-Guillaume, G. (2010b). Phoma du tournesol et Dessèchement précoce : point sur les avancées de la recherche. Phoma du tournesol et Dessèchement précoce : point sur les avancées de la recherche. *OCL - Oleagineux, Corps Gras, Lipides*, 17(3):157–160.

Sebillotte, M. (1990). Système de culture: un concept opératoire pour les agronomes. In *Les systèmes de culture*, pages 165–196. Combe, L. and Picard, D., INRA, Paris.

Sedghi, M., Sharifi, R., Namvar, A., Khandan-e Bejandi, T., and Molaei, P. (2008). Responses of sunflower yield and grain filling period to plant density and weed interference. *Research Journal of Biological Sciences*, 3.

Sims, D., Cheng, W., Luo, Y., and Seemann, J. (1999). Photosynthetic acclimation to elevated CO_2 in a sunflower canopy. *Journal of Experimental Botany*, 50:645–653.

Skoric, D., Jocic, S., Hladni, N., and Vannozzi, G. (2007). An analysis of heterotic potential for agronomically important traits in sunflower (*Helianthus annuus L.*). *Helia*, 30(46):55–74.

Smits, N., Dagher, A., and Aubertot, J. (2015). A simple grapevine growth model: towards GRAPE-PEST, a X-PEST model of damage caused by multiple pests on grapevine. In *IPM Innovation in Europe - 14 January - 16 January*, page 59, Poznań, Poland.

Stanojevic, D. (1985). Relationship between stand density and microclimate and their effect on some characters of sunflower. In *Proceeding 11th International Sunflower Conference, ISA*, volume 1, pages 301–306, Mar del Plata, Argentina.

Sukno, S., Melero-Vara, J., and Fernandez-Martinez, J. (1999). Inheritance of résistance to *Orobanche cernua Loefl.* in six sunflower lines. *Crop Science*, 39:647–678.

T

Taverne, M. (2003). Dessèchement précoce du tournesol et Phoma. Technical report, Dossier technique interne CETIOM.

Taverne, M. (2005). Effet du milieu sur le phomopsis et le phoma. *Oléoscope*, 80:16–18.

Tegtmeier, E. and Duffy, M. (2004). External costs of agricultural production in the United States. *International Journal of Agricultural Sustainability*, 2:1–20.

Terekhov, V. I., Begunov, I. I., and Begunova, A. G. (2004). Role of aerial infection in sunflower phomopsis (*Phomopsis* [*Diaporthe*] *Helianthi*) epiphytology. In *Proc. 16th International Sunflower Conference*, pages 124–128, Fargo, ND USA.

Thomas, H., Sauerborn, J., Müller-Stöver, D., Ziegler, A., Bedi, J., and Kroschel, J. (1998). The potential of *Fusarium oxysporum f.sp. orthoceras* as a biological control agent for *Orobanche cumana* in sunflower. *Biol Control*, 13:41–48.

Törn, A. (1978). A search-clustering approach to global optimization. *Towards Global Optimization*, 2:49–62.

Tourvieille de Labrouhe, D., Pilorge, E., Nicolas, P., and Vear, F. (2000). *Le mildiou du tournesol.* INRA-CETIOM, Paris.

Trapani, N., Hall, A., Sadras, V., and Vilella, F. (1992). Ontogenetic changes in radiation use efficiency of sunflower (*Helianthus annuus L.*) crop. *Field Crops Research*, 29:301–316.

V

Vallance (1950). Studies on the germination of the seeds of *Striga hermonthica*. I.The influence of moisture-treatment, stimulant dilution and after-ripening on germination. *Annals of Botany*, 14:348–363.

Van Hezewijk, M. and Verkleij, J. (1996). The effect of nitrogenous compounds on in vitro germination of *Orobanche crenata Forsk. Weed Research*, 36:395–404.

Vasyura, V. (1975). Dynamics of broomrape seed germination depths in soil. *Tacbak URSS*, 4:54–55.

Vear, F. (2016). Changes in sunflower breeding over the last fifty years. *OCL - Oléagineux, Corps Gras, Lipides*, 23(D203):8.

Vear, F. and Muller, M. (2011). Progrès variétal chez le tournesol : l'apport des ressources génétiques au sein du genre Helianthus. *Innovations Agronomiques*, 14:139–150.

Vilain, M. (1989). *La maîtrise technique de la production*. Eds, Lavoisier, Paris, France, 2nd edition.

Voinov, A. and Bousquet, F. (2010). Modelling with stakeholders. *Environmental Modelling & Software*, 25:1268–1281.

Voudouris, C. (1998). Guided local search - an illustrative example in function optimisation. *BT Technology Journal*, 16(3):46–50.

Vrânceanu, A., Tudor, V., Stoenescu, F., and Pirvu, N. (1980). Virulence groups of *Orobanche cumana Wallr.*, differential hosts and resistance source genes in sunflower. In *Proceeding 9th International Sunflower Conference*, volume 2, pages 74–82, Torremolinos, Spain.

W

Wallach, D., Goffinet, B., Bergez, J., Debaeke, P., Leenhardt, D., and Aubertot, J. (2001). Parameter estimation for crop models. *Agronomy Journal*, 93(4):757.

Wallach, D., Makowski, D., and Jones, J. (2006). *Working with dynamic crop models : evaluation, analysis, parameterization, and applications*. Elsevier, 2nd edition.

Weeraratne, W. and Priyantha, M. (2003). First report of Phoma black stem of sunflower in Sri Lanka and its management. *Annals of the Sri Lanka Department of Agriculture*, 5:263–270.

Westwood, J. and Foy, C. (1999). Influence of nitrogen on germination and early development of broomrape (*Orobanche spp.*). *Weed science*, 47(1):2–7.

Willocquet, L., Aubertot, J., Lebard, S., Robert, C., Lannou, C., and Savary, S. (2008). Simulating multiple pest damage in varying winter wheat production situations. *Field Crops Research*, 107:12–28.

Willocquet, L., Savary, S., Fernandez, L., Elazegui, F., Castilla, N., Zhu, D., Tang, Q., Huang, S., Lin, X., Singh, H., and Srivastava, R. (2002). Structure and validation of RICEPEST, a production situation-driven, crop growth model simulating rice yield response to multiple pest injuries for tropical Asia. *Ecological Modelling*, 153:247–268.

Witten, I., Eibe, F., and Hall, M. (2011). *Data mining practical machine learning tools and techniques*. Elsevier, 3rd edition.

Wright, M. (1996). Direct search methods: Once scorned, now respectable. In *Numerical Analysis 1995: Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis*, pages 191–208, Addison Wesley Longman,Harlow, UK. Griffiths, D.F. and Watson, G.A.

Y

Yoneyama, K., Xie, X., Kim, H., Kisugi, T., Nomura, T., Sekimoto, H., Yokota, T., and Yoneyama, K. (2012). How do nitrogen and phosphorus deficiencies affect strigolactone production and exudation? *Planta*, 235:1197–1207.

Yoneyama, K., Xie, X., Kusumoto, D., Sekimoto, H., Sugimoto, Y., Takeuchi, Y., and Yoneyama, K. (2007). Nitrogen deficiency as well as phosphorus deficiency in sorghum promotes the production and exudation of 5-deoxystrigol, the host recognition signal for arbuscular mycorrhizal fungi and root parasites. *Planta*, 227:125–132.

Z

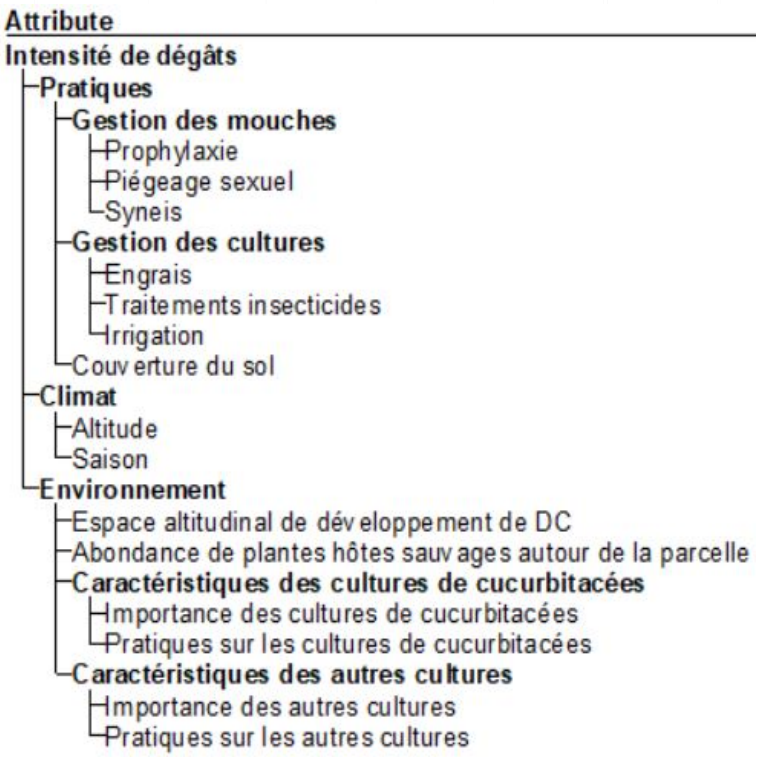
Zadoks, J. (1993). *Perspectives for rural areas in the European community.*, pages 161–170. Wageningen, Pers.

Zecchini, B. (1992). *Les serpents de France : systématique, biologie des venins, clinique et thérapeutique des morsures : expérience du Centre anti-poisons de Marseille*. PhD thesis, Université d'Aix-Marseille II - Faculté de médecine de Marseille.

Zhang, W., Ma, Y., Wang, Z., Ye, X., and Shui, J. (2013). Some soybean cultivars have ability to induce germination of sunflower broomrape. *PLoS ONE*, 8(3):e59715.

Annexes

Annexe 1 - IPSIM - Chouchou - Mouche

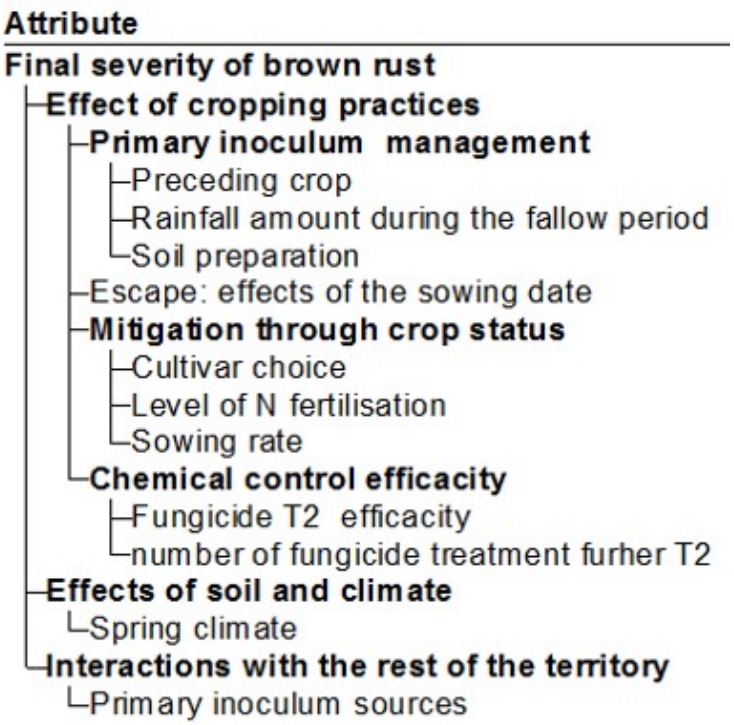


Structure du modèle IPSIM - Chayotte - Mouche

Variable (niveau de l'observation : nombre d'observations)													
Exploitation	Saison	Prophylaxie	Piégeage sexuel	Syneis	Engrais	Traitements insecticides	Couverture du sol	Abondance de plantes hôtes sauvages autour de la parcelle	Importance des cultures de cucurbitacées	Pratiques sur les cultures de cucurbitacée	Importance des autres cultures	Incidence (niv. d'inc.) : Nombre d'obs.	
												Observée	Simulée
1	été : 2 hiver : 1	non : 2 oui : 1	non : 2 oui : 1	non : 2 oui : 1	supérieur à la préconisation : 2 selon la préconisation : 1	fort : 2 nul : 1	nu : 2 couvert : 1	faible : 3	faible : 3	gamour partiel : 3	faible : 3	(1) : 4 (2) : 2	(4) : 1 (2) : 2
2	été : 2 hiver : 2	non : 2 oui : 2	non : 2 oui : 2	non : 2 oui : 2	supérieur à la préconisation : 2 selon la préconisation : 2	nul : 4	nu : 2 couvert : 2	forte : 4	faible : 4	gamour partiel : 4	moyenne à élevée : 4	(5) : 4	(5) : 3 (4) : 1
3	été : 2 hiver : 2	non : 4	non : 2 oui : 2	non : 4	selon la préconisation : 4	moyen : 2 nul : 2	couvert : 4	faible : 4	élevée : 4	gamour partiel : 4	faible : 4	(5) : 1 (4) : 2 (3) : 1	(5) : 2 (4) : 2
4	été : 1 hiver : 1	non : 2	non : 2	non : 2	selon la préconisation : 2	moyen : 2	nu : 2	faible : 2	faible : 2	conventionnelles : 2	faible : 2	(4) : 1 (3) : 1	(5) : 1 (4) : 1
5	été : 1	oui : 1	oui : 1	non : 1	selon la préconisation : 1	nul : 1	couvert : 1	faible : 1	faible : 1	conventionnelles : 1	faible : 1	(4) : 1	(3) : 1
6	été : 2 hiver : 2	non : 2 oui : 2	non : 2 oui : 2	non : 2 oui : 2	selon la préconisation : 4	fort : 2 moyen : 2	nu : 4	faible : 4	faible : 4	conventionnelles : 4	faible : 4	(5) : 1 (2) : 2 (3) : 1	(2) : 1 (3) : 1 (4) : 1 (5) : 1
7	été : 2 hiver : 2	non : 2 oui : 2	non : 2 oui : 2	non : 2 oui : 2	selon la préconisation : 4	fort : 2 moyen : 2	nu : 2 couvert : 2	faible : 4	faible : 4	conventionnelles : 4	faible : 4	(5) : 1 (2) : 2 (3) : 1	(2) : 1 (3) : 1 (4) : 1 (5) : 1
8	été : 2 hiver : 2	non : 2 oui : 2	non : 2 oui : 2	non : 2 oui : 2	selon la préconisation : 4	fort : 2 moyen : 2	couvert : 4	faible : 4	faible : 4	conventionnelles : 4	faible : 4	(5) : 1 (2) : 2 (3) : 1	(2) : 1 (3) : 1 (4) : 1 (5) : 1
9	été : 1 hiver : 1	non : 2	non : 2	non : 2	selon la préconisation : 2	moyen : 2	couvert : 2	faible : 2	faible : 2	conventionnelles : 2	faible : 2	(5) : 1 (3) : 1	(5) : 1 (4) : 1
10	été : 3 hiver : 2	oui : 5	non : 5	non : 5	peu ou pas : 5	nul : 5	couvert : 5	faible : 5	nulle : 5	gamour partiel : 5	moyenne à élevée : 5	(2) : 4 (3) : 1	(2) : 2 (3) : 3
11	été : 1 hiver : 1	non : 2	oui : 2	non : 2	selon la préconisation : 2	nul : 2	couvert : 2	faible : 2	faible : 2	conventionnelles : 2	faible : 2	(5) : 1 (3) : 1	(5) : 1 (4) : 1
12	été : 2 hiver : 1	non : 2 oui : 1	non : 2 oui : 1	non : 2 oui : 1	supérieur à la préconisation : 2 selon la préconisation : 1	fort : 2 nul : 1	nu : 2 couvert : 1	faible : 3	faible : 3	gamour partiel : 3	faible : 3	(4) : 1 (2) : 2	(4) : 1 (2) : 2
13	été : 5 hiver : 4	non : 5 oui : 4	non : 5	non : 5	supérieur à la préconisation : 5 selon la préconisation : 4	fort : 5 nul : 4	nu : 5 couvert : 4	faible : 3	faible : 9	conventionnelles : 9	faible : 3	(3) : 2 (4) : 6 (5) : 1	(2) : 2 (3) : 2 (4) : 2 (5) : 3

Description du jeu de données utilisé pour comparer les méthodes CART et IPSIM dans le cas de la mouche de la chayotte

Annexe 2 - IPSIM - Blé - Rouille brune



Structure du modèle IPSIM - Blé - Rouille brune

Variable (nombre d'observations : niveau de l'observation)												
Dép.	Nombre parcelle	Année	Hôte	Travail du sol	Date semis	Niveau sensibilité	Densité	fertilisation N	Efficacité fongicide parties aériennes	Nombre de fongicide	Niveau d'incidence	
											Observée	Simulée
35	29	2010 - 2013	Autres : 264	Travail : 264	Normale : 128 ; Précoce : 1 ; Retardée : 135	(3) : 14 ; (4) : 20 ; (5) : 41 ; (6) : 107 ; (7) : 55 ; (8) : 27	Faible : 189 ; Normale : 75	Equilibrée : 264	Bonne efficacité : 122 ; Faible efficacité : 142	(0) : 142 ; (1) : 69 ; (2) : 53	(0-5) : 244 ; (5-10) : 8 ; (10-20) : 3 ; (20-50) : 7 ; (50-100) : 2	(0-5) : 255 ; (10-20) : 5 ; (20-50) : 4
91	33	2010 - 2012	Autres : 194 ; non hôte : 2	Travail : 123 ; Non travail : 73	Normale : 192 ; Retardée : 4	(3) : 25 ; (4) : 18 ; (5) : 48 ; (6) : 63 ; (7) : 33 ; (8) : 9	Faible : 120 ; Normale : 76	Equilibrée : 196	Bonne efficacité : 121 ; Faible efficacité : 75	(0) : 121 ; (1) : 47 ; (2) : 28	(0-5) : 161 ; (5-10) : 19 ; (10-20) : 9 ; (20-50) : 7	(0-5) : 178 ; (5-10) : 13 ; (20-50) : 5
78	2	2007 - 2009	Autres : 45	Travail : 25 ; NR : 20	Normale : 41 ; Précoce : 4	(2) : 1 ; (3) : 3 ; (4) : 2 ; (5) : 10 ; (6) : 18 ; (7) : 6 ; (8) : 5	Faible : 45	Equilibrée : 45	Bonne efficacité : 25 ; Efficacité moyenne : 20	(1) : 45	(0-5) : 38 ; (5-10) : 3 ; (10-20) : 1 ; (20-50) : 2 ; (50-100) : 1	(0-5) : 42 ; (10-20) : 3
62	2	2007 - 2008	Autres : 11	Non travail : 11	Normale : 7 ; Précoce : 4	(2) : 2 ; (3) : 4 ; (6) : 3 ; (8) : 2	Faible : 11	Equilibrée : 11	Bonne efficacité : 4 ; Faible efficacit : 7	(0) : 7 ; (1) : 4	(0-5) : 6 ; (5-10) : 1 ; (10-20) : 1 ; (50-100) : 3	(0-5) : 4 ; (10-20) : 4 ; (50-100) : 3
64	1	2008 - 2008	Autres : 3	Travail : 3	Normale : 3	(6) : 1 ; (8) : 2	Faible : 3	Equilibrée : 3	Faible efficacité : 3	(0) : 3	(0-5) : 3	(0-5) : 3
37	3	2005 - 2007	Autres : 25	Travail : 6 ; Non travail : 19	Normale : 19 ; Précoce : 5 ; Retardée : 1	(3) : 4 ; (4) : 5 ; (4.5) : 1 ; (5) : 3 ; (6) : 8 ; (8) : 4	Faible : 18 ; Normale : 7	Equilibrée : 25	Faible efficacité : 25	(0) : 25	(0-5) : 14 ; (5-10) : 6 ; (10-20) : 4 ; (20-50) : 1	(0-5) : 21 ; (5-10) : 3 ; (10-20) : 1
49	2	2007 - 2012	Autres : 51	Travail : 45 ; Non travail : 6	Normale : 51	(3) : 11 ; (5) : 11 ; (6) : 11 ; (7) : 9 ; (8) : 9	Faible : 51	Equilibrée : 51	Bonne efficacité : 48 ; Faible efficacité : 3	(0) : 6 ; (1) : 45	(0-5) : 48 ; (10-20) : 2 ; (20-50) : 1	(0-5) : 50 ; (10-20) : 1
28	7	2009 - 2013	Autres : 133 ; non hôte : 2	Travail : 20 ; Non travail : 115	Normale : 7 ; Précoce : 1 ; Retardée : 127	12 : (3) ; 19 : (4) ; 32 : (5) ; 56 : (6) ; 9 : (7) ; 7 : (8)	Elevée : 8 ; Faible : 77 ; Normale : 50	Equilibrée : 135	Bonne efficacité : 100 ; Faible efficacité : 35	(0) : 104 ; (1) : 30 ; (2) : 1	(0-5) : 132 ; (5-10) : 1 ; (20-50) : 1 ; (50-100) : 1	(0-5) : 131 ; (5-10) : 2 ; (20-50) : 2
77	4	2005 - 2007	Autres : 38	Travail : 10 ; Non travail : 28	Normale : 33 ; Retardée : 5	(2) : 3 ; (3) : 9 ; (4) : 4 ; (5) : 1 ; (6) : 16 ; (7) : 1 ; (8) : 4	Faible : 38	Equilibrée : 38	Bonne efficacité : 29 ; Faible efficacité : 9	(0) : 19 ; (1) : 19	(0-5) : 19 ; (5-10) : 11 ; (10-20) : 1 ; (20-50) : 6 ; (50-100) : 1	(0-5) : 25 ; (10-20) : 9 ; (50-100) : 4
18	1	2005 - 2005	Autres : 9	Travail : 9	Normale : 8 ; Retardée : 1	(3) : 2 ; (4) : 1 ; (5) : 1 ; (6) : 2 ; (7) : 1 ; (8) : 2	Faible : 9	Equilibrée : 9	Faible efficacité : 9	(0) : 9	(0-5) : 8 ; (5-10) : 1	(0-5) : 7 ; (5-10) : 2
36	1	2005 - 2005	Autres : 9	Travail : 9	Normale : 8 ; Précoce : 1	(3) : 2 ; (4) : 1 ; (5) : 1 ; (6) : 2 ; (7) : 1 ; (8) : 2	Faible : 9	Equilibrée : 9	Faible efficacité : 9	(0) : 9	(0-5) : 5 ; (5-10) : 2 ; (10-20) : 1 ; (50-100) : 1	(0-5) : 7 ; (5-10) : 2
41	1	2005 - 2005	Autres : 8	NR : 8	Normale : 6 ; Retardée : 2	(3) : 2 ; (4) : 1 ; (5) : 1 ; (6) : 2 ; (8) : 2	Faible : 8	Equilibrée : 8	Faible efficacité : 8	(0) : 8	(0-5) : 8	(0-5) : 8
16	1	2005 - 2005	Autres : 4	Non travail : 4	Normale : 4	(3) : 1 ; (5) : 1 ; (6) : 2	Normale : 4	Equilibrée : 4	Faible efficacité : 4	(0) : 4	(0-5) : 4	(0-5) : 4
85	1	2003 - 2003	Autres : 6	Non travail : 6	Normale : 6	(3) : 1 ; (5) : 2 ; (6) : 1 ; (7) : 2	Faible : 6	Equilibrée : 6	Faible efficacité : 6	(0) : 6	(0-5) : 6	(0-5) : 6
32	1	2011 - 2011	Autres : 34	Non travail : 34	Retardée : 34	(3) : 3 ; (4) : 4 ; (5) : 14 ; (6) : 4 ; (7) : 2 ; (8) : 5 ; (9) : 2	Faible : 18 ; Normale : 16	Equilibrée : 34	Faible efficacité : 34	(0) : 34	(0-5) : 25 ; (5-10) : 2 ; (10-20) : 4 ; (20-50) : 3	(0-5) : 19 ; (5-10) : 12 ; (20-50) : 3
47	1	2011 - 2011	Autres : 17	Non travail : 17	Normale : 17	(3) : 2 ; (4) : 2 ; (5) : 6 ; (6) : 2 ; (7) : 1 ; (8) : 3 ; (9) : 1	Normale : 17	Equilibrée : 17	Faible efficacité : 17	(0) : 17	(0-5) : 11 ; (5-10) : 3 ; (20-50) : 3	(0-5) : 11 ; (5-10) : 4 ; (20-50) : 2
89	6	2011 - 2013	Autres : 64	Travail : 37 ; Non travail : 27	Normale : 42 ; Précoce : 22	(3) : 6 ; (4) : 13 ; (5) : 19 ; (6) : 21 ; (7) : 2 ; (8) : 3	Faible : 38 ; Normale : 26	Equilibrée : 63 ; Excédentaire : 1	Bonne efficacité : 26 ; Faible efficacité : 38	(0) : 51 ; (1) : 13	(0-5) : 48 ; (5-10) : 3 ; (10-20) : 6 ; (20-50) : 6 ; (50-100) : 1	(0-5) : 51 ; (5-10) : 11 ; (20-50) : 2
95	4	2012 - 2013	Autres : 88	Travail : 48 ; Non travail : 40	Normale : 48 ; Retardée : 40	(3) : 7 ; (4) : 9 ; (5) : 15 ; (6) : 33 ; (7) : 14 ; (8) : 10	Faible : 42 ; Normale : 46	Equilibrée : 88	Bonne efficacité : 88	(0) : 21 ; (1) : 41 ; (2) : 26	(0-5) : 65 ; (10-20) : 11 ; (20-50) : 4 ; (50-100) : 8	(0-5) : 76 ; (5-10) : 8 ; (10-20) : 2 ; (20-50) : 2
80	6	2011 - 2013	Autres : 366	Travail : 366	Normale : 126 ; Précoce : 96 ; Retardée : 144	(3) : 27 ; (4) : 33 ; (5) : 78 ; (6) : 129 ; (7) : 69 ; (8) : 30	Faible : 218 ; Normale : 148	Equilibrée : 366	Bonne efficacité : 242 ; Faible efficacité : 124	(0) : 198 ; (1) : 96 ; (2) : 72	(0-5) : 323 ; (5-10) : 9 ; (10-20) : 5 ; (20-50) : 14 ; (50-100) : 9	(0-5) : 351 ; (5-10) : 2 ; (10-20) : 7 ; (20-50) : 6
86	2	2012 - 2012	Autres : 102	Travail : 102	Normale : 102	(3) : 12 ; (4) : 10 ; (5) : 20 ; (6) : 40 ; (7) : 12 ; (8) : 8	Faible : 50 ; Normale : 52	Equilibrée : 102	Bonne efficacité : 52 ; Faible efficacité : 50	(0) : 50 ; (1) : 52	(0-5) : 93 ; (5-10) : 2 ; (10-20) : 1 ; (20-50) : 6	(0-5) : 92 ; (5-10) : 1 ; (10-20) : 3 ; (20-50) : 6
76	2	2013 - 2013	Autres : 104	Travail : 104	Normale : 104	(3) : 12 ; (4) : 8 ; (5) : 20 ; (6) : 32 ; (7) : 20 ; (8) : 12	Faible : 52 ; Normale : 52	Equilibrée : 104	Efficacité moyenne : 52 ; Faible efficacité : 52	(0) : 52 ; (1) : 26 ; (2) : 26	(0-5) : 104	(0-5) : 104
31	125	1997 - 2013	Autres : 3 ; hôte : 6 ; non hôte : 192	Travail : 29 ; Non travail : 172	Normale : 175 ; Précoce : 8 ; Retardée : 18	(2) : 54 ; (3) : 8 ; (4) : 16 ; (4.5) : 28 ; (5) : 9 ; (5.5) : 2 ; (6) : 12 ; (6.5) : 2 ; (7) : 36 ; (8) : 18 ; (8.5) : 2 ; (9) : 14	Elevée : 105 ; Faible : 10 ; Normale : 86	Equilibrée : 194 ; Excédentaire : 7	Bonne efficacité : 24 ; Efficacité moyenne : 29 ; Faible efficacité : 148	(0) : 136 ; (1) : 65	(0-5) : 117 ; (5-10) : 18 ; (10-20) : 10 ; (20-50) : 15 ; (50-100) : 41	(0-5) : 137 ; (5-10) : 24 ; (10-20) : 18 ; (20-50) : 22
17	1	2012 - 2012	non hôte : 1	Non travail : 1	Normale : 1	(4) : 1	Normale : 1	Equilibrée : 1	Faible efficacité : 1	(0) : 1	(5-10) : 1	(0-20) : 1
21	1	2013 - 2013	non hôte : 8	Travail : 8	Normale : 8	(6) : 4 ; (7) : 4	Elevée : 8	Equilibrée : 8	Faible efficacité : 8	(0) : 8	(0-5) : 3 ; (5-10) : 3 ; (10-20) : 2	(0-5) : 8

Description du jeu de données utilisé pour comparer les méthodes CART et IPSIM dans le cas du blé

Annexe 3 - Script de SimMat codé sous R

Le modèle générique SimMat a été codé sous différents langages : R, VBA et C++. Sous R, les sept paramètres sont notés ci-dessous :

Paramètre	Définition	Unité
tmin	Seuil de température journalière minimale permettant la maturation des fructifications	°C
tmax	Seuil de température journalière maximale permettant la maturation des fructifications	°C
rmin	Seuil pluviométrique minimal permettant la maturation des fructifications	mm
nd	Nombre de jours pour le cumul pluviométrique	jour
meanFD	Nombre attendu de jours favorables	jour
sdFD	Écart-type des jours favorables	jour
k	Paramètre lié à la proportion de pseudothèces matures qui libèrent leurs ascospores en fonction des précipitations quotidiennes	mm^{-1}

Paramètres, définitions et unités des parameters SimMat.

Le script R est présenté ci-après :

```
SimMat <- function(tmin, tmax, rmin, meanFD, sdFD, k, nd, weather,
    startingJDate)
{
  temperatures <- weather[,which(names(weather)=="Temperatures")] #
      temperatures dataset vector
  rainfall <- weather[,which(names(weather)=="Rainfall")] # rainfall dataset vector
  Date <- weather[,which(names(weather)=="Date")] # date dataset vector

  DFM <- numeric(length(temperatures)) # Day Favourable to Maturation
  CDFM <- numeric(length(temperatures)) # Cumulated Day Favourable to Maturation
  Crain <- numeric(length(temperatures)) # Cumulated Rain
  PMP <- numeric(length(temperatures)) # Proportion of Mature Perithecia
  PEP <- numeric(length(temperatures)) # Proportion of Empty Perithecia
  p_empty_pseudothecia <- numeric(length(temperatures))
  p_full_pseudothecia <- numeric(length(temperatures))
  p_perithece_relargant_jourJ <- numeric(length(temperatures))

  for (i in startingJDate:length(temperatures)){ # from the starting date to the end
      of dataset
    Crain[i] <- sum(rainfall[(i+1-nd):i],na.rm = T) # Rainfall sum on the nd-1 last
        days and the day i
    if (!is.na(temperatures[i]>tmin & temperatures[i]<tmax & Crain[i]>rmin) &
        (temperatures[i]>tmin & temperatures[i]<tmax & Crain[i]>rmin))
    {DFM[i] <- 1} # The day d is favourable to maturation
    CDFM[i] <- sum(DFM[1:i]) # cumulation of days favourables to maturation

    PMP[i] <- pnorm(CDFM[i], mean = meanFD, sd = sdFD)
    # proportion of mature perithecia that did not release ascospore until the day d

    PEP[i] <- ifelse(!is.na(1-exp(-k*rainfall[i])),1-exp(-k*rainfall[i]),0)
```

```
1       p_empty_pseudothecia[i] <- p_empty_pseudothecia[i-1] + (PMP[i]-p_empty_
            pseudothecia[i-1])*PEP[i] # taux de perithecia vides au jour d
2
3       p_full_pseudothecia[i] <- (PMP[i]-p_empty_pseudothecia[i])/(1-p_empty_
            pseudothecia[i]) # taux de maturation sur les perithecia non vides
4
5       p_perithece_relargant_jourJ[i] <- (PMP[i] - p_empty_pseudothecia[i-1])*PEP
            [i]
6       # Taux de prithces relarguant des ascopores au jour j
7     }
8     data.frame(Date,temperatures,rainfall,Crain,DFM,CDFM,PMP,PEP,
9               p_empty_pseudothecia, p_full_pseudothecia, p_perithece_relargant_
                    jourJ)
10  }
```

Annexe 4 - Formalisme d'ASPHODEL

Formalisme mathématique d'ASPHODEL

La maturation journalière des pseudothèces est associée à un indice, noté $i_{pm}.(d)$, calculé comme le produit de l'indice hygrométrique et de l'indice de température journalier, soit $i_h(d)$ et $i_t(d)$ respectivement. Ces variables sont calculées à l'aide des équations (1), (2) et (7).

$$i_{pm}(d) = i_h(d) * i_t(d) \tag{1}$$

L'indice hygrométrique journalier $i_h(d)$ est calculé par le système d'équations (2).

$$\begin{cases} i_h(d) = \frac{NHF(d) - NHF_{min}}{NHF_{max} - NHF_{min}} \text{ si } NHF(d) > NHF_{min} \\ \quad sinon \\ i_h(d) = 0 \end{cases} \tag{2}$$

où $NHF(d)$ est le nombre d'heures où les conditions sont favorables à la maturation pour le jour d ; NHF_{min} est le nombre minimal d'heures où les conditions sont favorables à la maturation pour que le jour soit considéré comme favorable à la maturation ; NHF_{max} est le nombre d'heures dans une journée. La variable $NHF(d)$ est calculée comme suit avec les variables d'entrée météorologiques

horaires :

$$\begin{cases} NHF(d) = \sum_{m=1}^{m=24} \delta_{rs}(m) \end{cases} \tag{3}$$

Où m est un indice de sommation associé aux mesures horaires ; $\delta_{rs}(m)$ est le symbole Kronecker avec :

$$\begin{cases} r = s \text{ si } RH_h(m) > RH_0 \text{ et } \theta_h(m) > \theta_{min}^{ih1} \\ \quad et \\ r \neq s \text{ sinon} \end{cases} \tag{4}$$

où RH_0 est le seuil minimum pour que l'humidité relative soit favorable à la maturation des périthèces ; θ_{min}^{ih1} est le seuil de température journalière minimum utilisé pour calculer les indices hygrométriques journaliers

Pour les variables d'entrée résumant 3 heures, $NHF(d)$ est calculé comme suit :

$$\begin{cases} NHF(d) = \sum_{m=1}^{m=8} 3\delta_{rs}(m) \end{cases} \tag{5}$$

Où m est un indice de sommation associé à trois mesures horaires ; $\delta_{rs}(m)$ est le symbole Kronecker avec

$$\begin{cases} r = s \text{ si } RH_{3h}(m) > RH_0 \text{ et } \theta_{3h}(m) > \theta_{min}^{ih3} \\ \quad et \\ r \neq s \text{ sinon} \end{cases} \tag{6}$$

où θ_{min}^{ih3} est le seuil de température journalière minimale utilisé pour calculer les indices hygrométriques journaliers.

L'indice de température journalier $i_t(d)$ est calculé par le système des équations 4.

$$\begin{cases} i_t(d) = \frac{1}{2}[1 + \cos(2\pi \frac{\bar{\theta}(d) - \theta_{opt}^{it}}{\theta_{max}^{it} - \theta_{min}^{it}})] \text{ if } \theta_{min}^{it} < \bar{\theta}(d) < \theta_{max}^{it} \\ \quad \text{sinon} \\ i_t(d) = 0 \end{cases} \tag{7}$$

où $\bar{\theta}(d)$ est la température journalière moyenne du jour d ; θ_{opt}^{it} est la température journalière optimale pour la maturation des fructifications ; θ_{max}^{it} et θ_{min}^{it} sont respectivement la température journalière maximale et minimale de maturation.

Les éjections d'ascospores sont décrites à l'aide de deux variables : une variable binaire journalière, appelée $AR1(d)$, décrit la première éjection d'ascospores avec une valeur de 1 pour une éjection importante d'ascospores survenant pendant le jour d et une valeur nulle sinon. Cette variable est calculée comme suit :

$$AR1(d) = \delta_{rs}(d) \tag{8}$$

Avec :

$$\begin{cases} r = s \text{ si } IPM_{cum}^{1}(d) \geq IPM_{min}^{1} \text{ et } R(d) > R_{min}^{ASPHODEL} \\ \quad et \\ r \neq s \text{ sinon} \end{cases} \tag{9}$$

où $IPM_{cum}^{1}(d)$ est le cumul de l'indice associé à la proportion quotidienne de maturation pseudothéciale i_{pm} de février 1st au jour d courant. IPM_{min}^{1} est la valeur minimale à atteindre par l'indice associé à la proportion quotidienne de maturation pseudothéciale pour considérer que le taux de maturation pseudothéciale est de 50% ; et $R_{min}^{ASPHODEL}$ est la valeur minimale des précipitations quotidiennes à atteindre pour permettre la libération des ascospores.

Une autre variable binaire quotidienne, appelée $ARF(d)$, décrit les libérations quotidiennes d'ascospores suivantes avec une valeur de 1 pour les libérations importantes d'ascospores survenant pendant la journée d et une valeur nulle sinon. Cette variable est calculée comme suit :

$$ARF(d) = \delta_{rs}(d) \tag{10}$$

Avec :

$$\begin{cases} r = s \text{ si } IPM_{cum}^{2}(d) \geq IPM_{min}^{2} \text{ et } R(d) > R_{min}^{ASPHODEL} \\ \quad et \\ r \neq s \text{ sinon} \end{cases} \tag{11}$$

où $IPM_{cum}^{2}(d)$ est le cumul de l'indice associé à la proportion quotidienne de maturation pseudothéciale i_{pm} de la dernière libération d'ascospores au jour courant d. IPM_{min}^{2} est la valeur minimale à atteindre par l'indice associé à la proportion quotidienne de maturation pseudothéciale.

Symbole	Définition	Type	Variable	Unité
i_{pm}	Indice de maturation pseudothéciale	Journalier	Sortie	-
$AR1$	Première éjection de l'année	Journalière	Sortie	-
ARF	Relargage d'ascospores suivantes	Journalière	Sortie	-
i_h	Indice Hygrométrique	Journalier	Etat	-
i_t	Température index	Journalier	Etat	-
NHF	Nombre d'heures au cours desquelles les conditions sont favorable à la maturation au cours de la journée	Journalier	Etat	h
$\bar{\theta}$	Température moyenne journalière	Journalier	Etat	°C
IPM^1_{cum}	Cumul de l'indice associé à la proportion quotidienne de maturation pseudothéciale i_{pm} du 1st février au jour d	Journalier	Etat	-
IPM^2_{cum}	Cumul de l'indice associé à la proportion quotidienne de maturation pseudothéciale i_{pm} de la dernière ejection au jour d	Journalier	Etat	-
RH_h	Humidité relative	Horaire	Entrée	%
RH_{3h}	Humidité relative	Horaire	Entrée	%
θ_h	Température quotidienne utilisée pour calculer les indices hygrométriques journalier	Horaire	Entrée	°C
θ_{3h}	Température quotidienne utilisée pour calculer les indices hygrométriques journalier	Horaire	Entrée	°C
R	Pluviométrie journalière	Journalier	Entrée	mm

Symboles, définitions, pas de temps, et unités des variables d'ASPHODEL utilisées dans le module du champignon pour décrire la maturation des périthèces et les émissions d'ascospores

Symbole	Définition	Valeur	Unité
NHF_{min}	Nombre minimal d'heures favorables à la maturation pseudothéciale pour qu'un jour soit favorable à la maturation	6	h
NHF_{max}	Nombre d'heures par jour	24	h
RH_0	Humidité relative minimale favorable à la maturation pseudothéciale	85	%
θ_{min}^{ih1}	Seuil de température minimale des journaliers utilisé pour calculer les indices hygrométriques des journaliers avec les variables d'entrée horaire météo	10	°C
θ_{min}^{ih3}	Seuil de température minimale des journaliers utilisé pour calculer les indices hygrométriques des journaliers à l'aide de trois variables Horaire météo	10	°C
θ_{opt}^{it}	Température optimale du journalier pour la maturation des fructifications	18.5	°C
θ_{min}^{it}	Seuil minimal de température de journalier utilisé pour calculer les indices de température journalier	9	°C
θ_{max}^{it}	Seuil maximal de température du journalier utilisé pour calculer les indices de température journalier	28	°C
IPM_{min}^{1}	Valeur minimale de l'indice de maturation pseudothéciale pour la première libération d'ascospores	3.7	-
IPM_{min}^{2}	Valeur minimale de l'indice de maturation pseudothéciale pour les dissémination ultérieures d'ascospores	0.4	-
$R_{min}^{ASPHODEL}$	Seuil minimal moyen des précipitations cumulées des journaliers pour la libération d'ascospores	0.5	mm

Symboles, définitions, valeurs et unités des paramètres d'ASPHODEL utilisés dans le sous-modèle champignon pour décrire la maturation et les émissions d'ascospores.

Formalisme informatique d'ASPHODEL

Ce module nécessite deux jeux de données : un jeu avec un pas de temps journalier ($tabdonnee.j$) et un autre avec un pas de temps horaire ($tabdonnee.h$). Le jeu de données horaire doit être composé des variables suivantes : le mois (noté $MOIS$), le jour ($JOUR$), l'heure ($HEURE$) (de nature facteur) et l'année (AN), la température moyenne (T), l'humidité (U) et la pluviométrie (RR) (de nature numérique). Le jeu de données journalier doit être composé des variables suivantes : la température minimale (TN), maximale (TX) et la pluviométrie (RR) (de nature numérique).

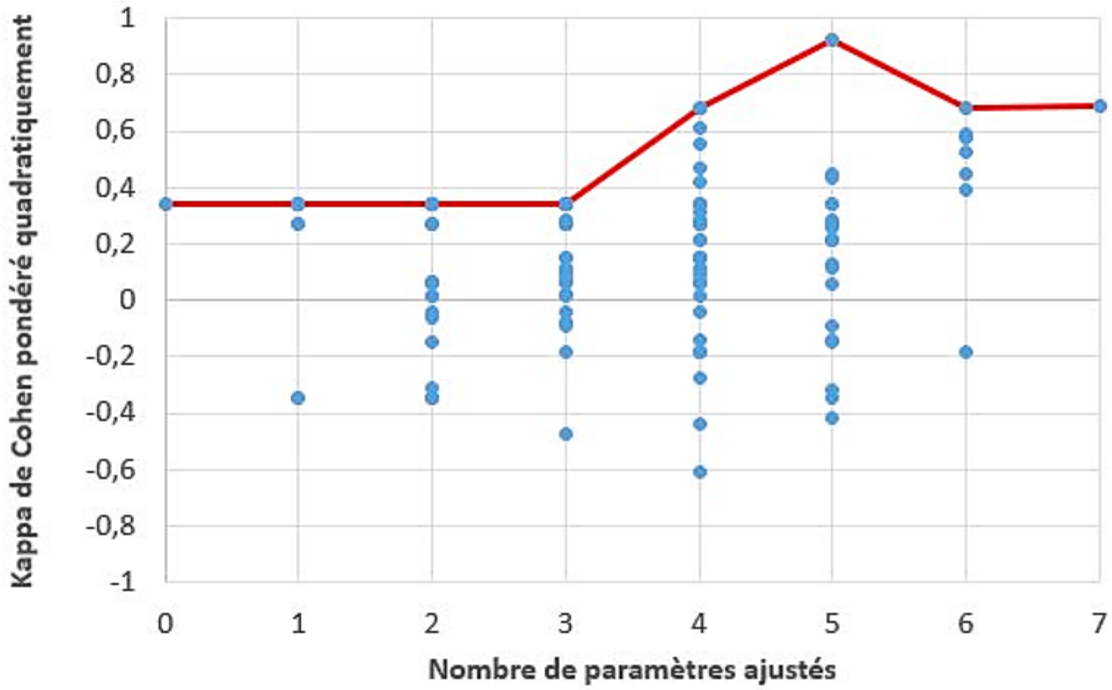
```
Asphodel=function (tabdonnee.j,tabdonnee.h){

  # Compartiment MATURATION
  # 1. Calcul du cumul journalier des periodes de 3h d'hygrometrie relative prise en
      compte dans la maturation des peritheces (hy)
  hy=NULL ; HR85=NULL
  for (i in 1:length(levels(tabdonnee.h$MOIS))){
    hy_cum_1MOIS=NULL ; HR85_cum_1MOIS=NULL

    dm=tabdonnee.h[which(tabdonnee.h$MOIS==i),]
    dm$JOUR=dm$JOUR[,drop = TRUE]
    for (j in 1:length(levels(dm$JOUR))){
      dj=dm[which(dm$JOUR==j),]

      if (length(dj$HEURE)==24){
        hy_cum_1JOUR=0 ; HR85_cum_1JOUR=0
        for (k in 1:8){
          if (dj$U[k*3] >= 85 & dj$T[k*3] > 10){
            hy_cum_1JOUR = hy_cum_1JOUR + 3 ; HR85_cum_1JOUR = HR85_cum_1JOUR +
                3
          } else {hy_cum_1JOUR = hy_cum_1JOUR ; HR85_cum_1JOUR = HR85_cum_1
              JOUR}

          if (dj$U[k*3] >= 85){HR85_cum_1JOUR=HR85_cum_1JOUR+3} else {hHR85_
              cum_1JOUR=HR85_cum_1JOUR}
        }
        hy_cum_1MOIS = c(hy_cum_1MOIS, hy_cum_1JOUR)
        HR85_cum_1MOIS = c(HR85_cum_1MOIS, HR85_cum_1JOUR)
      }
    }
    hy = c(hy, hy_cum_1MOIS)
    HR85 = c(HR85, HR85_cum_1MOIS)
  }
  hy=round(hy,2) ; HR85=round(HR85,2)
```

```
# 2. Calcul du cumul journalier des periodes de 3h d'hygrometrie relative prise en compte pour la contamination (hh)
hh = NULL
for (i in 1:length(levels(tabdonnee.h$MOIS))){
  hh_cum_1MOIS=NULL

  dm=tabdonnee.h[which(tabdonnee.h$MOIS==i),]
  dm$JOUR=dm$JOUR[,drop = TRUE]
  for (j in 1:length(levels(dm$JOUR))){
    dj=dm[which(dm$JOUR==j),]

    if (length(dj$HEURE)==24){
      hh_cum_1JOUR=0
      for (k in 1:8){
        if (dj$U[k*3] >= 90){hh_cum_1JOUR = hh_cum_1JOUR + 3}
        if (dj$U[k*3] < 90 & dj$U[k*3] >= 85){hh_cum_1JOUR = hh_cum_1JOUR +
            ((dj$U[k*3]-84)/2)
        } else {hh_cum_1JOUR = hh_cum_1JOUR}
      }
      hh_cum_1MOIS = c(hh_cum_1MOIS, hh_cum_1JOUR)
    }
  }
  hh = c(hh, hh_cum_1MOIS)
}

# 3. Calcul de l'indice temperature journalier (it)
tabdonnee.j$TM_recalc=(tabdonnee.j$TN+tabdonnee.j$TX)/2
it1=ifelse ((9 < tabdonnee.j$TM_recalc & tabdonnee.j$TM_recalc < 28), ((cos
    ((tabdonnee.j$TM_recalc - 18.5) * (pi/9.5))/2) + 0.5), 0)
it = it1[1:length(hy)]

# 4. Calcul de l'indice hygrometrique journalier (ih)
ih = ifelse (hy >6, ((hy/18)-(1/3)), 0)

# 5. Calcul de l'indice de maturation journalier (ij)
ij = it*ih
```

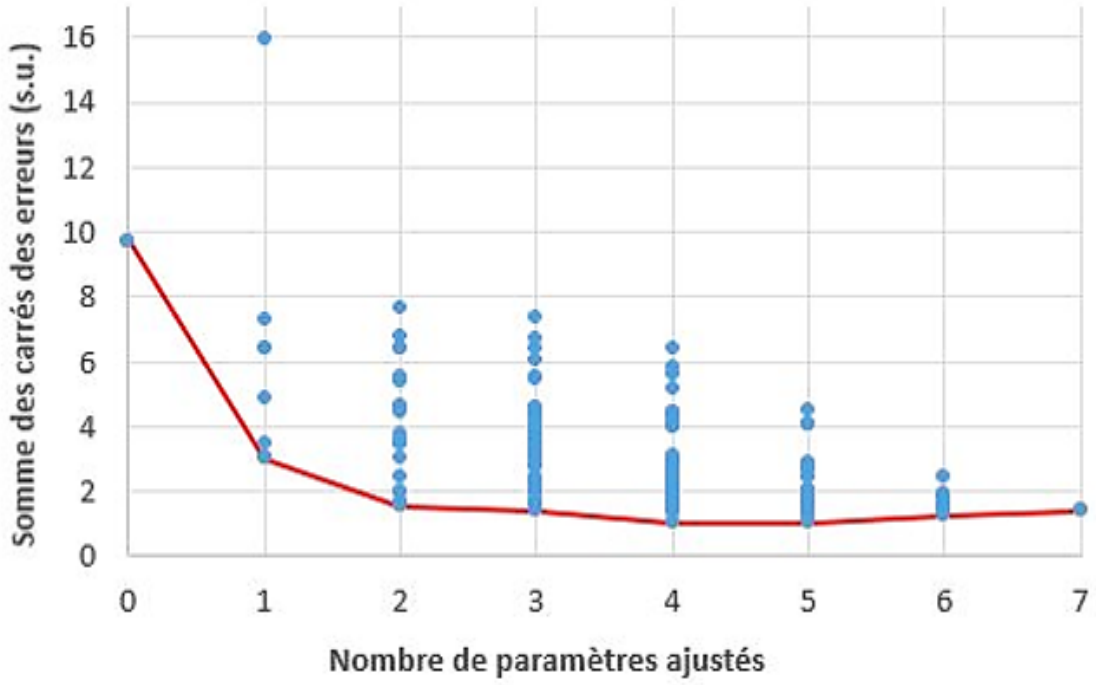
```
# 6.  Calcul de l'indice de maturation cumule (ijcum)
ijcum=0 ; mat=rep(NA, length(ij))
maturation=rep(0, length(ij))
for (i in 32:length(mat)){#
  ijcum = ijcum + ij[i]

  # Seuil de maturation :
  if (ijcum > 3.7){mat[i]= "TRUE"}
  maturation[i]=round(((ijcum*50)/3.7), 2)
}

# Compartiment PROJECTION ####
#************************
projection=rep(NA, length(ij))
# on indique dans la matrice projection la premiere projection
first.projection = which(mat=="TRUE" & tabdonnee.j$RR[1:length(ij)] > 0.5)
    [1]
projection[first.projection]="proj"

recharge=0
if (length(which(!is.na(mat))) > 0){
  for (i in (first.projection + 1):length(projection)) {
    if ((ij[i]>0.4 | recharge > 0.4) & tabdonnee.j$RR[i] > 0.5 & mat[i]=="
        TRUE"){
      projection[i]= "proj" ; recharge=0
    } else {projection[i]= "noproj" ; recharge=recharge + ij[i]}
  }
}
```

```r
    # Compartiment CONTAMINATION ####
    #***************************
    # 1.  Calcul de l'indice determinant le  risque  de contamination  (ic4)  pour le  jour  i
    ic=rep(NA, length(hh))
    for (i in 3:length(hh)){
      if (!is.na(projection[i]) & hh[i] >= 9) {
        ic[i] = round((hh[i]/11)*1.7, 2)
        } else {ic[i] = round(hh[i]/11,2)}
    }
    ic4=rep(NA, length(hh))
    ic4[32] = ic[32] ; ic4[33]= round(ic[33] + ic[32], 2)
    for (i in 34:length(ic4)){ic4[i] = round(ic[i] + ic[i-1] + ic[i-2], 2)}


    # 2.  Determination du jour  de contamination
    contamination=rep(NA, length(projection))
    if (length(which(!is.na(projection))) > 0){
      for (i in first.projection:length(contamination)){
        if ((!is.na(projection[i-1]) | !is.na(projection[i-2]) | !is.na(
            projection[i-3]) | !is.na(projection[i-4]) |
            !is.na(projection[i-5]) | !is.na(projection[i-6]) | !is.na(
                projection[i-7]) | !is.na(projection[i-8])) & ic4[i] > 4){
          contamination[i] = "Conta"}
      }
    }

    Date =as.character(tabdonnee.j$Date)
    JourJulian=c(1:length(hh))
    Tmoy=round(tabdonnee.j$TM_recalc,2)

    results1=cbind(Date=Date[1:length(hh)], JourJulian, ij, maturation,
        projection, contamination, ic, ic4, Tmoy[1:length(hh)], HR85, hy, it, ih
        )
    results=results1[32:nrow(results1),]
    return(results)
  }
```

Annexe 5 - Résultats d'optimisation de SimMat



Résultats de l'algorithme d'optimisation de la qualité prédictive du modèle SimMat-*Diaporthe* sur les précocités des dynamiques de maturation des périthèces. • = valeur de kappa pour une combinaison ; – = meilleure valeur de kappa pour chaque nombre de paramètres modifiés



Résultats de l'algorithme d'optimisation de la qualité prédictive du modèle SimMat-*Diaporthe* sur les dynamiques de maturation des périthèces. • = valeur de somme des carrés des erreurs pour une combinaison ; – = meilleure valeur de somme des carrés des erreurs pour chaque nombre de paramètres modifiés

Annexe 7. IPSIM - Tournesol - Phomopsis optimisé

	Pression_du_phomopsis_au_niveau_de_la_parcelle	Climat	Pratiques_culturales	Incidence_du_Phomopsis
1	Eleve	Defavorable	Defavorable	**Elevee**
2	Eleve	Defavorable	Moyennement_favorable	**Elevee**
3	Eleve	Defavorable	Favorable	**Moderee**
4	Eleve	Moyennement_Favorable	Defavorable	**Elevee**
5	Eleve	Moyennement_Favorable	Moyennement_favorable	**Moderee**
6	Eleve	Moyennement_Favorable	Favorable	**Moderee**
7	Eleve	Favorable	Defavorable	**Moderee**
8	Eleve	Favorable	Moyennement_favorable	**Moderee**
9	Eleve	Favorable	Favorable	**Faible**
10	Intermediaire	Defavorable	Defavorable	**Elevee**
11	Intermediaire	Defavorable	Moyennement_favorable	**Elevee**
12	Intermediaire	Defavorable	Favorable	**Moderee**
13	Intermediaire	Moyennement_Favorable	Defavorable	**Elevee**
14	Intermediaire	Moyennement_Favorable	Moyennement_favorable	**Moderee**
15	Intermediaire	Moyennement_Favorable	Favorable	**Faible**
16	Intermediaire	Favorable	Defavorable	**Moderee**
17	Intermediaire	Favorable	Moyennement_favorable	**Faible**
18	Intermediaire	Favorable	Favorable	**Faible**
19	Faible	Defavorable	Defavorable	**Moderee**
20	Faible	Defavorable	Moyennement_favorable	**Moderee**
21	Faible	Defavorable	Favorable	**Faible**
22	Faible	Moyennement_Favorable	Defavorable	**Moderee**
23	Faible	Moyennement_Favorable	Moyennement_favorable	**Faible**
24	Faible	Moyennement_Favorable	Favorable	**Faible**
25	Faible	Favorable	Defavorable	**Faible**
26	Faible	Favorable	Moyennement_favorable	**Faible**
27	Faible	Favorable	Favorable	**Faible**

FIGURE 1 – Table d'agrégation de l'Incidence du modèle - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous . En rouge les valeurs sont défavorables pour chaque variable ; en noir, les valeurs sont intermédiaires ; en vert les valeurs sont favorables.

	Niveau_de_risque_dans_le_departement	Enfouissement_des_residus_apres_anteprecedent_si_tournesol	Pression_du_phomopsis_au_niveau_de_la_parcelle
1	Eleve	Non_Retournement	**Eleve**
2	Eleve	Retournement	**Faible**
3	Intermediaire	Non_Retournement	**Faible**
4	Intermediaire	Retournement	**Faible**
5	Faible	Non_Retournement	**Faible**
6	Faible	Retournement	**Faible**

FIGURE 2 – Table d'agrégation de la Pression du modèle - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous . En rouge les valeurs défavorables pour chaque variable ; en noir, les valeurs intermédiaires ; en vert les valeurs favorables.

	Climat_pour_linfection	Climat_pour_le_passage_sur_tige	Climat
1	Defavorable	Defavorable	**Defavorable**
2	Defavorable	Moyennement_favorable	**Moyennement_Favorable**
3	Defavorable	Favorable	**Favorable**
4	Favorable	Defavorable	**Favorable**
5	Favorable	Moyennement_favorable	**Favorable**
6	Favorable	Favorable	**Favorable**

FIGURE 3 – Table d'agrégation du climat du modèle - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous . En rouge les valeurs défavorables pour chaque variable ; en noir, les valeurs intermédiaires ; en vert les valeurs favorables.

	Controle_par_esquive	Attenuation_par_letat_du_couvert	Controle_cultural
1	Inefficace	Defavorable	**Defavorable**
2	Inefficace	Moyennement_favorable	**Moyennement_favorable**
3	Inefficace	Favorable	**Moyennement_favorable**
4	Efficace	Defavorable	**Defavorable**
5	Efficace	Moyennement_favorable	**Favorable**
6	Efficace	Favorable	**Favorable**

FIGURE 4 – Table d'agrégation du contrôle cultural du modèle - Tournesol - Phomopsis après optimisation de la qualité de prédiction telles que présentées sous . En rouge les valeurs défavorables pour chaque variable ; en noir, les valeurs intermédiaires ; en vert les valeurs favorables.

	Controle_cultural	Lutte_chimique	Controle_genetique	**Pratiques_culturales**
1	Defavorable	Non_Traitement	Sensible_et_peu_sensible	**Defavorable**
2	Defavorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Moyennement_favorable**
3	Defavorable	Traitement	Sensible_et_peu_sensible	**Moyennement_favorable**
4	Defavorable	Traitement	Tres_peu_sensible_et_resistante	**Defavorable**
5	Moyennement_favorable	Non_Traitement	Sensible_et_peu_sensible	**Defavorable**
6	Moyennement_favorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Favorable**
7	Moyennement_favorable	Traitement	Sensible_et_peu_sensible	**Favorable**
8	Moyennement_favorable	Traitement	Tres_peu_sensible_et_resistante	**Favorable**
9	Favorable	Non_Traitement	Sensible_et_peu_sensible	**Favorable**
10	Favorable	Non_Traitement	Tres_peu_sensible_et_resistante	**Favorable**
11	Favorable	Traitement	Sensible_et_peu_sensible	**Favorable**
12	Favorable	Traitement	Tres_peu_sensible_et_resistante	**Favorable**

FIGURE 5 – Table d'agrégation des pratiques culturales du modèle - Tournesol - Phomopsis après optimisation des règles d'agrégation telles que présentées sous DEXi. En rouge les valeurs défavorables à l'agriculteur (favorables à la maladie) ; en noir, les valeurs intermédiaires ; en vert les valeurs favorables à l'agriculteur (défavorables à la maladie).