



HAL
open science

Modélisation des conséquences évolutives de la compétition pour la lumière dans des populations composites de blé

Emmanuelle Blanc

► **To cite this version:**

Emmanuelle Blanc. Modélisation des conséquences évolutives de la compétition pour la lumière dans des populations composites de blé. Evolution [q-bio.PE]. Université Paris-Saclay, 2021. Français. NNT : 2021UPASB013 . tel-03767519

HAL Id: tel-03767519

<https://theses.hal.science/tel-03767519v1>

Submitted on 2 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modélisation des conséquences évolutives
de la compétition pour la lumière dans des
populations composites de blé
*Modelling evolutionary outcomes of competition
for light in wheat composite populations*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°567, Sciences du végétal : du gène à l'écosystème (SEVE)
Spécialité de doctorat : Biologie
Unité de recherche : Université Paris-Saclay, INRAE, CNRS, AgroParisTech,
GQE - Le Moulon, 91190, Gif-sur-Yvette, France.
Réfèrent : Faculté des Sciences d'Orsay

**Thèse présentée et soutenue à Paris-Saclay,
le 12/03/2021, par**

Emmanuelle BLANC

Composition du Jury

Christine DILLMANN Professeure, UPSaclay, Orsay	Présidente
Evelyne COSTES Directrice de recherche, INRAE, Montpellier	Rapporteur & Examinatrice
Hélène FREVILLE Directrice de recherche, INRAE, Montpellier	Rapporteur & Examinatrice
Eric GOZE Cadre de recherche, CIRAD, Montpellier	Examineur

Direction de la thèse

Jérôme ENJALBERT Directeur de recherche, INRAE, Gif-sur-Yvette	Directeur de thèse
Pierre BARBILLON Maître de conférence, AgroParisTech, Paris	Co-Directeur de thèse

AVANT-PROPOS

Cette thèse a été réalisée au sein de l'équipe "Diversité, Evolution et Adaptation des Populations" (DEAP) de l'Unité Mixte de Recherche "Génétique Quantitative et Evolution – Le Moulon".

Ce travail est soutenu par une subvention publique de l'Agence Nationale de la Recherche (ANR) dans le cadre du programme "Investissements d'Avenir", à travers le projet "Initiative Doctorale Interdisciplinaire (IDI) 2017" financé par l'IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

REMERCIEMENTS

Tout au long de ma thèse, j'ai eu la chance d'échanger avec de nombreuses personnes qui ont participé au projet et qui m'ont soutenue et je souhaite les remercier chaleureusement.

Je tiens tout d'abord à remercier grandement Jérôme Enjalbert et Pierre Barbillon de m'avoir fait confiance en me donnant l'opportunité de faire cette thèse et de m'avoir encadrée pendant plus de 3 ans. Vous avez su vous montrer toujours disponibles pour m'accompagner, vous avez fait preuve de qualités humaines et professionnelles remarquables et j'ai beaucoup appris de ce travail avec vous. Je souhaite également remercier Christophe Lecarpentier, qui m'a le premier accueillie au Moulon et m'a fait découvrir l'univers passionnant de WALTer. Merci pour tous tes conseils et merci de m'avoir fait confiance pour travailler sur ton modèle et de m'en avoir patiemment expliqué les spécificités avec pédagogie et humour.

Je suis également profondément reconnaissante envers Christian Fournier et Christophe Pradal : j'ai été particulièrement heureuse de vous rencontrer et de travailler avec vous. Merci beaucoup pour votre accueil chaleureux lors de mes visites à Montpellier. Votre collaboration a été très précieuse pour le projet et m'a énormément appris. Un grand merci également à Franck Gauthier (Bubar) pour tout le temps que tu as consacré à travailler sur les développements de WALTer avec moi.

Parmi les personnes qui ont le plus contribué au projet, je veux remercier chaleureusement Meije Gawinowski, avec qui j'ai beaucoup aimé travailler pendant son stage, et qui m'a fait bénéficier de son expertise et de toutes ses merveilleuses qualités en développant un modèle de génétique des populations bien plus abouti que ce que j'aurais pu imaginer. J'ai également été très heureuse de travailler avec Antoine Payen, dont le travail de stage a beaucoup contribué à la thèse : merci pour ton implication, ton sérieux et ton enthousiasme. Merci également à Bilâl Derouich, dont le travail en stage, encadré par Christian Fournier et Christophe Pradal, a aidé à l'amélioration de WALTer.

Je souhaite également remercier vivement les membres de mon comité de thèse : Robert Faivre, Judith Legrand et Gaëtan Louarn, qui ont été d'une grande aide tout au long du projet. Merci beaucoup pour votre temps et pour votre accompagnement. Je garde de très bons souvenirs de mes comités de thèse, qui ont été l'occasion d'échanges agréables et particulièrement constructifs. Un grand merci également aux membres du jury : Evelyne Costes, Christine Dillmann, Hélène Fréville et Eric Gozé, pour avoir accepté d'évaluer mon travail et pour des discussions très enrichissantes.

J'aimerais aussi remercier Charlotte Prieu de m'avoir fait confiance en me donnant l'opportunité d'encadrer des TD et TP à l'UVSQ. J'ai particulièrement apprécié cette expérience, qui a été très instructive pour moi.

Un grand merci également à Timothée Flutre et Isabelle Goldringer qui m'ont chacun fait bénéficier de leurs précieux conseils, jusqu'aux répétitions de ma soutenance. Merci à toutes les autres personnes qui m'ont aidée avec ces répétitions : Pierre Barbillon et Jérôme Enjalbert bien sûr, mais aussi Robert Faivre, Meije Gawinowski, Xavier Raffoux, Adrienne Ressayre, Baptiste Rouger et Michel Turbet Delof. Merci beaucoup à Alice Beaugrand et Yannick De Oliveira de m'avoir aidée avec toutes mes questions de Python. Merci à Adrien Falce pour toute son aide et à Benoît Johannet de m'avoir accompagnée toute ma thèse jusqu'au jour de la soutenance et même après. Je souhaite aussi remercier Valérie Lespinas, Rozenn Le Guyader et les membres de l'école doctorale pour leur

travail et leur bienveillance qui ont rendu ma thèse plus facile.

Plus largement, je veux remercier tous les membres du Moulon avec qui j'ai eu plaisir à échanger : merci pour les discussions scientifiques enrichissantes et pour la bonne ambiance. Je suis particulièrement heureuse d'avoir partagé mon bureau avec Emma Forst : merci pour ta gentillesse et pour tout ce que tu m'as appris, de vive voix ou grâce à tes BD. J'ai aussi eu la joie d'avoir Meije Gawinowski comme voisine de bureau lors de son retour en thèse dans l'équipe, encore merci à toi. Merci aussi à nos voisins du 2^{ème} étage, avec qui il était toujours agréable de discuter ou de boire un thé. Merci en particulier à Baptiste Rouger et à Gaëlle van Frank pour votre soutien, votre bonne humeur et pour tous les bons moments partagés. Merci aussi aux anciens du 2^{ème} étage et en particulier à Julie Borg et à Christophe Lecarpentier ainsi qu'à Louise Thiaville pour ta positivité et pour ton travail, qui m'a beaucoup aidé pour ma thèse. Merci à tous les stagiaires et mains d'œuvre qui ont apporté leur bonne humeur lors de leur passage dans le bureau et en particulier à Michel Turbet Delof, qui est un ancien stagiaire, mais surtout un nouveau thésard dans l'équipe. Merci à tous les membres de l'équipe DEAP, avec qui j'ai eu plaisir à travailler, à échanger et à manger des gâteaux. Merci aussi aux membres de l'équipe ABI, qui partagent avec nous le bâtiment, les gâteaux et les bons moments.

Merci à toutes les personnes qui ont fait de ma thèse au Moulon une période joyeuse en s'impliquant dans l'organisation des barbecues, des soirées crêpes, des soirées jeux et de tous les clubs plus ou moins officiels : potager, gym, aquagym, tennis, et bien sûr Secret Santa, qui m'ont permis de garder la forme et le sourire pendant toutes ces années. Merci en particulier à Baptiste, Meije, Emma, Gaëlle, Michel, Christophe L., Julie, Xavier, Alice, Pierre M., Yannick, Harry, Mélanie, Lætitia, Natalia, Valérie C., Zeineb, Elise, l'équipe des Dinofort...

Pour finir, j'aimerais remercier ma famille et en particulier mes parents, Aurélie et François, pour leur soutien et leur écoute. Et un immense merci à Guillaume de m'avoir soutenue et encouragée tous les jours et d'avoir été le meilleur collègue de télétravail pendant ma dernière année de thèse.

SOMMAIRE

INTRODUCTION GÉNÉRALE	1
Remise en question du modèle agricole dominant.....	2
La transition agroécologique	3
Le cas du blé	4
L'architecture aérienne dans les peuplements hétérogènes de blé	7
Une approche par modélisation.....	10
Présentation du plan de la thèse.....	12
Références	14
CHAPITRE 1 : Adaptations du FSPM WALTer pour l'étude des peuplements hétérogènes de blé....	19
Introduction.....	20
Matériel et méthodes.....	21
Résultats	26
Discussion	29
Références.....	34
CHAPITRE 2 : Impact de l'aléatoire dans le modèle WALTer	39
Introduction.....	40
Matériel et méthodes.....	42
Résultats	44
Discussion	46
Références.....	49
CHAPITRE 3 : Analyse de sensibilité de WALTer	51
Introduction.....	52
Matériel et méthodes.....	53
Résultats	58
Discussion	62
Références.....	66

CHAPITRE 4 : Calibration of WALTer using an adaptive design and a metamodelling approach	69
Introduction.....	70
Material and methods.....	71
Results	75
Discussion	79
References.....	81
CHAPITRE 5 : Functional-structural plant modelling highlights how diversity in leaf dimensions and tillering capability could promote the efficiency of wheat cultivar mixtures	84
Introduction.....	85
Material and methods.....	86
Results	89
Discussion	96
References.....	101
CHAPITRE 6 : Étude de l'évolution de peuplements hétérogènes de blé par le couplage de WALTer et d'un modèle de génétique des populations.....	105
Introduction.....	106
Matériel et méthodes.....	107
Résultats	116
Discussion	121
Références.....	125
DISCUSSION GÉNÉRALE	129
Apports de la thèse pour la modélisation structure-fonction des peuplements hétérogènes de blé	130
Modélisation de l'évolution de peuplements hétérogènes.....	131
Apports de la thèse pour la conception et la gestion des peuplements hétérogènes de blé.....	133
Perspectives.....	135
Références.....	137
ANNEXE 1 : Ressources complémentaires du chapitre 1	140
ANNEXE 2 : Ressources complémentaires du chapitre 3	150
ANNEXE 3 : Ressources complémentaires du chapitre 5	156
ANNEXE 4 : Ressources complémentaires du chapitre 6	159

INTRODUCTION GÉNÉRALE

Remise en question du modèle agricole dominant

Le modèle agricole dominant dans les pays industrialisés aujourd'hui repose sur un ensemble de progrès techniques mis en place peu après la seconde guerre mondiale (Malézieux, 2012 ; Gliessman, 2014). Il s'appuie ainsi sur une large mécanisation des cultures, la mise en place de systèmes d'irrigation intensive et l'utilisation d'intrants de synthèse (herbicides, pesticides et engrais). La mise en place de ces différentes pratiques vise à optimiser et standardiser les conditions de culture en réduisant les pressions liées aux maladies, aux ravageurs et à la compétition des adventices et en éliminant les stress associés aux limitations en ressources (eau, azote, phosphore, etc.). Parallèlement, ce modèle agricole se concentre sur un petit nombre d'espèces et de variétés très productives qui sont adaptées à ces conditions de culture. Cette simplification des systèmes agricoles se base sur le développement de variétés modernes et conduit à une importante diminution de la diversité cultivée : environ 75% de la diversité génétique des plantes cultivées a ainsi été perdue entre 1900 et 2000 (Gliessman, 2014). Ces progrès ont permis en partie de faire face à la demande alimentaire par d'importantes augmentations de rendement et une baisse du prix des denrées (Evenson et Gollin, 2003 ; Gliessman, 2014).

Cependant, ce modèle productiviste a également montré des limites importantes, avec des impacts environnementaux, sanitaires et sociaux négatifs (Hatt *et al.*, 2016). Ce modèle est en effet peu durable car il dépend largement de ressources non renouvelables, notamment des énergies fossiles (Gliessman, 2014). L'agriculture contribue significativement aux changements climatiques, en étant notamment responsable de 25% des émissions totales de gaz à effet de serre (Hatt *et al.*, 2016), alors que des changements dans le modèle dominant permettraient de diminuer significativement ces émissions (Smith *et al.*, 2008). L'agriculture intensive est également responsable d'une importante dégradation et pollution des sols qui met en danger la sécurité alimentaire (Gomiero, 2016). De plus, elle est la première cause de pollution des eaux à l'échelle mondiale (Gliessman, 2014). Ces différentes pollutions ont un impact négatif sur la production agricole, mais également sur le fonctionnement des écosystèmes non cultivés et sur la santé humaine. En outre, l'agriculture intensive atteint ses limites, comme l'illustre la stagnation des rendements de blé en Europe depuis le milieu des années 1990 (Brisson *et al.*, 2010). Cette stagnation n'est pas un cas isolé puisque d'autres pays (comme le Mexique) et d'autres cultures (comme le riz en Asie) sont concernés. Plusieurs causes expliquent cette tendance, notamment la simplification des rotations de cultures. Le changement climatique, qui est globalement défavorable au rendement des céréales, est également en cause. De plus, on note que les systèmes agricoles "conventionnels" montrent une faible résilience face à la variabilité et aux extrêmes climatiques (Kahiluoto *et al.*, 2019), alors même que ces derniers devraient s'intensifier ces prochaines années du fait des changements climatiques (Coumou et Rahmstorf, 2012). Cela souligne la vulnérabilité de ces systèmes agricoles et le risque associé pour la sécurité alimentaire. Face à ces constats, et dans un contexte d'augmentation de la demande alimentaire mondiale, lié en partie à la population croissante (Tilman *et al.*, 2011), le modèle agricole productiviste est remis en question et un changement de paradigme paraît nécessaire (Østergård *et al.*, 2009).

La transition agroécologique

L'agroécologie est une solution possible pour une transition vers un modèle agricole plus durable, qui permettrait de surmonter les limites du système "conventionnel" (Gliessman, 2014 ; Hatt *et al.*, 2016 ; Altieri, 2018 ; Altieri et Nicholls, 2020). L'agroécologie est un concept dynamique qui a beaucoup évolué depuis son apparition en 1928 (Hatt *et al.*, 2016 ; Wezel *et al.*, 2020). Elle présente plusieurs définitions et peut être considérée à la fois comme une discipline scientifique, une pratique agricole et un mouvement social ou politique (Wezel *et al.*, 2009). Quelle que soit la définition considérée, l'agroécologie est basée sur un rapprochement entre l'écologie et l'agronomie : elle consiste à appliquer les concepts de l'écologie à la conception et à la gestion des systèmes agricoles au sens large. Les sites, ou régions, de production agricole sont ainsi considérés comme des écosystèmes : on parle d'agro-écosystèmes (Gliessman, 2014). Différents ensembles de principes, sur lesquels reposent l'agroécologie, ont été proposés (Migliorini et Wezel, 2017). Ces différents ensembles ont cependant tous en commun de considérer la mobilisation de la biodiversité comme un principe central de l'agroécologie. En effet, plusieurs études d'écologie ont montré un lien entre la diversité des écosystèmes et leur fonctionnement, avec une plus grande productivité et une meilleure stabilité temporelle des écosystèmes les plus diversifiés (Loreau *et al.* 2001 ; Hooper *et al.* 2005 ; van der Plas, 2019). Dans le cadre de la transition agroécologique, une meilleure mobilisation de la biodiversité permettrait ainsi de fournir de nombreux services écosystémiques^a, qui pourraient se substituer, au moins en partie, à l'utilisation d'intrants de synthèse (Hatt *et al.*, 2016).

La biodiversité peut être définie à différents niveaux interconnectés : elle peut se définir comme le nombre d'espèces ou de variétés d'un système, mais il est également possible de considérer la diversité génétique ou la diversité fonctionnelle. La diversité fonctionnelle repose sur la notion de traits fonctionnels, qui sont des traits qui impactent indirectement la fitness des individus de par leur effet sur leur croissance, leur reproduction ou leur survie (Violle *et al.*, 2007). Les traits fonctionnels sont liés aux stratégies des individus et à leurs fonctions dans l'écosystème (Zakharova *et al.*, 2019). Si beaucoup d'études portant sur le lien entre le fonctionnement des écosystèmes et leur diversité se sont concentrées sur la diversité en espèces, l'importance de l'analyse sous l'angle de la diversité fonctionnelle est aujourd'hui largement établie (Díaz et Cabido, 2001 ; van der Plas, 2019). Ainsi, plusieurs mécanismes, décrits initialement dans le cadre de la diversité en espèces, peuvent être à l'origine d'un impact positif de la diversité fonctionnelle sur le fonctionnement des écosystèmes (Loreau, 2000). L'effet de complémentarité de niche peut prendre place lorsque les espèces de l'écosystème ont des niches écologiques différentes, ce qui limite la compétition entre elles et leur permet donc d'exploiter au mieux les ressources de leur environnement. Un effet de facilitation peut également prendre place si la présence d'une espèce est favorable au développement d'une autre espèce dans l'écosystème. Enfin, un effet de sélection peut être observé lorsqu'une compétition asymétrique entraîne la dominance d'une espèce très productive par rapport aux autres espèces de l'écosystème. La diversité en espèces/variétés et la diversité fonctionnelle sont toutes deux liées à la diversité génétique, et l'importance de cette dernière pour le fonctionnement des écosystèmes a également été mise en avant (Hooper *et al.* 2005 ; Hughes *et al.*, 2008). Pour la transition agroécologique, il est donc intéressant de considérer une augmentation de la diversité des agro-

a Un service écosystémique est un service fourni par un écosystème et qui bénéficie aux humains.

écosystèmes à la fois en termes d'espèces/de variétés, de diversité fonctionnelle et de diversité génétique (Wezel *et al.*, 2020). Que l'on considère la diversité en espèces/variétés, la diversité fonctionnelle, ou la diversité génétique, il est possible de distinguer différents niveaux d'intégration de la biodiversité dans les agro-écosystèmes (Jackson *et al.*, 2007 ; Hatt *et al.*, 2016). En effet, l'agro-biodiversité peut être réduite à la diversité planifiée, qui correspond à la diversité introduite volontairement par le producteur dans le système. Cependant, il est également possible d'y intégrer la diversité associée, qui correspond à l'ensemble des organismes du système qui n'ont pas été introduits volontairement dans le système, qu'ils soient considérés comme bénéfiques, neutres ou nuisibles pour le producteur. Pour ces deux niveaux d'intégration, l'augmentation de la biodiversité permet la production de services écosystémiques. Si l'agroécologie s'intéresse donc à l'augmentation de la diversité dans son sens le plus large (biodiversité planifiée et associée), il est intéressant de noter que l'augmentation de la diversité planifiée peut conduire à une augmentation de la diversité associée (Altieri, 1999 ; Duru *et al.*, 2015). Enfin, dans le cadre de la transition agroécologique, la mobilisation de la biodiversité doit se faire à différentes échelles spatiales (parcelle, ferme, paysage) et prendre en compte la diversité temporelle (par la rotation des cultures par exemple) (Gliessman, 2014). À l'échelle de la parcelle, de nombreuses pratiques culturales, telles que les associations d'espèces, l'agroforesterie ou la culture de couverts, permettent de mobiliser la diversité interspécifiques. Cependant, plusieurs études ont mis en avant l'intérêt de la diversité intra-spécifique et son rôle dans le fonctionnement des écosystèmes (Bolnick *et al.*, 2011 ; Cianciaruso *et al.*, 2009 ; Violle *et al.*, 2012). La transition agroécologique pourrait ainsi également s'appuyer sur une augmentation de diversité intra-spécifique à l'échelle de la parcelle, par la culture de populations hétérogènes ou d'associations variétales par exemple.

Le cas du blé

Notre étude est focalisée sur le blé, qui est la culture la plus cultivée dans le monde, pour ce qui est de la surface récoltée, avec plus de 15% des surfaces cultivées mondiales occupées par cette céréale. Le blé est également une culture majeure en France, où elle occupe près de 38% des surfaces cultivées (FAO, 2018). En France, c'est principalement le blé tendre (*Triticum aestivum* L.) qui est cultivé.

Si, comme pour de nombreuses espèces cultivées, la domestication du blé tendre a entraîné une diminution de sa diversité génétique (Goldringer *et al.*, 2013 ; Bonnin *et al.*, 2014), son histoire évolutive, étroitement liée à l'Homme, a ensuite contribué à une diversification de ses caractéristiques agronomiques. Ainsi, jusqu'à la fin du XIX^e siècle, le blé tendre a été sélectionné localement par les paysans, conduisant à la création de variétés de pays (ou landraces). La sélection effectuée par les paysans est basée sur le phénotype des plantes (sélection massale) : les graines semées d'une année sur l'autre proviennent des plantes présentant les caractéristiques les plus favorables aux yeux des paysans. Du fait de la sélection d'individus différents cultivés en mélange, les variétés de pays sont des variétés hétérogènes, pour lesquelles toutes les plantes ne sont pas génétiquement identiques : on parle de variétés population. Cette diversité s'explique également par le régime de reproduction du blé. En effet, bien que le blé soit autogame, ce qui signifie que la reproduction se fait par autofécondation, *i.e.* par la fécondation d'un individu avec lui-même, un

faible pourcentage d'allogamie existe chez cette espèce. De ce fait, des croisements entre individus ont lieu au sein des variétés de pays, conduisant à un brassage génétique, et à une diversité de génotypes dans ces populations. En France, comme dans d'autres pays industrialisés, la sélection moderne s'est substituée à la sélection massale à la ferme, à la fin du XIX^e siècle. Cette sélection professionnalisée a conduit à la création de variétés lignées, bien plus homogènes génétiquement que les landraces, en se basant initialement sur le tri des meilleurs individus issus des variétés de pays, leur hybridation, puis l'autofécondation et la sélection des meilleurs descendants. L'utilisation de ces nouvelles variétés, appelées aujourd'hui variétés anciennes, s'est rapidement substituée à celle des variétés de pays, conduisant à une diminution de la diversité génétique cultivée. Enfin, les variétés modernes, apparues en France après la seconde guerre mondiale, sont des lignées pures : des plantes homozygotes, génétiquement plus homogènes que les variétés anciennes. Ces variétés sont très utilisées, avec un succès particulier de quelques variétés phares cultivées très largement, entraînant une diminution supplémentaire de la diversité génétique du blé cultivé en France (Bonnin *et al.*, 2014).

Parallèlement à la diminution de la diversité génétique, la domestication et la sélection du blé tendre ont également entraîné une diminution de sa diversité fonctionnelle. Les variétés modernes de blé présentent ainsi une très faible diversité pour les traits ayant été directement sélectionnés comme le poids des grains ou la hauteur de la plante (Cantarel *et al.*, 2020). En effet, la hauteur des variétés de blé a été réduite rapidement, à partir de la Révolution Verte, par l'introduction de gènes de nanismes, appelés gènes Rht pour *Reduced Height*, provenant de variétés japonaises (Gale *et al.*, 1985). La réduction de la hauteur du blé est un objectif de sélection car les plantes courtes allouent moins de ressources au développement de l'appareil végétatif et peuvent ainsi consacrer plus de ressources à la production de grains. De plus, le risque de verse, particulièrement fort avec l'usage intensif de la fertilisation azotée, est réduit pour les plantes courtes. En plus de la réduction de diversité sur les traits directement sélectionnés, Cantarel *et al.* (2020) ont également noté une très faible diversité pour certains traits fonctionnels qui n'étaient pas directement sélectionnés, comme la teneur en azote des feuilles par exemple.

Face au constat que les variétés modernes de blé sont très homogènes, plusieurs pratiques permettent d'augmenter la diversité à l'échelle de la parcelle. Ainsi, il est possible de cultiver des variétés anciennes, qui, bien que peu hétérogènes, présentent un peu de variabilité allélique (Bonnin *et al.*, 2014). La culture de landraces est également possible, car certaines sont conservées dans des banques de gènes ou chez des paysans qui les cultivent encore. D'autres variétés population peuvent également être cultivées. En particulier, il est possible de créer des populations composites par le croisement de plantes issues de différentes variétés. La diversité des populations composites peut-être très importante, en particulier si elles sont issues de croisements entre un nombre important de parents, sélectionnés pour offrir une base génétique large. La diversité de ces populations, qui sont semées de génération en génération, peut être encore augmentée par l'intégration d'un gène de stérilité mâle, qui permet d'augmenter le taux d'allogamie au sein de la parcelle. Enfin, l'augmentation de la diversité intra-parcelle peut aussi passer par la culture de mélanges de variétés : le semis de la parcelle se fait alors à partir de graines issues de plusieurs variétés différentes mélangées dans le semoir.

Cette pratique s'est beaucoup développée en France ces dernières années (Figure 1). Les données

récoltées par [FranceAgriMer \(2020\)](#) montrent ainsi une augmentation de la proportion de la surface de blé tendre en France occupée par des mélanges à partir de 2010. Cette augmentation s'est intensifiée à partir de 2015, portant à près de 12% la part des mélanges dans la surface de blé française en 2020. Des entretiens réalisés auprès de paysans par [van Frank \(2018\)](#) ont apporté un éclairage sur l'intérêt qu'ils portent à la culture de mélanges de blé. Les paysans interrogés ont ainsi évoqué les bénéfices potentiels sur la stabilité inter-annuelle des rendements et la réduction des risques liés aux aléas ; la possibilité de combiner des caractéristiques différentes qui ne pourraient pas être toutes trouvées au sein d'une seule variété ; et la facilité de gestion, en comparaison à la gestion de différents lots de semences lorsque plusieurs variétés sont cultivées en pur dans la ferme.

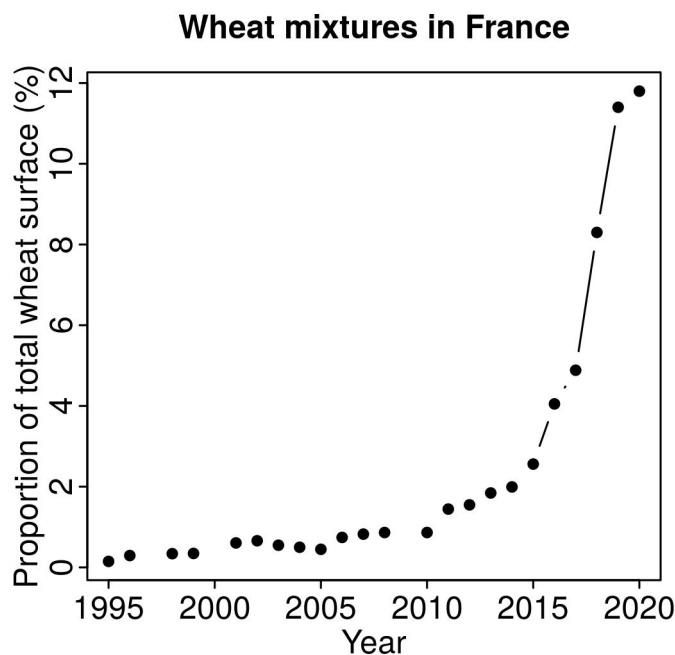


Figure 1. Évolution de la proportion des mélanges de variétés dans la surface totale de blé tendre cultivé en France entre 1995 et 2020. Source : [FranceAgriMer, 2020](#)

Différentes études scientifiques, portant sur l'utilisation des mélanges variétaux de blé tendre, ont mis en avant plusieurs avantages de cette pratique. En 2018, une méta-analyse par [Borg *et al.*](#) a ainsi rapporté des éléments suggérant un intérêt des mélanges pour le rendement et sa stabilité, la qualité des grains, la résistance aux maladies, le contrôle des insectes ravageurs, la compétition contre les adventices, la réduction de la verse, ou encore l'exploitation de l'eau et des nutriments du sol. L'intérêt des mélanges, en particulier sur le rendement, semble plus important en conditions stressantes. La culture des mélanges serait donc une pratique privilégiée pour les systèmes à bas niveau d'intrants qui sont développés dans le cadre de la transition agroécologique. Historiquement, c'est d'abord pour leur intérêt quant à la résistance aux maladies que les mélanges de blé ont été étudiés. En effet, de nombreuses études expérimentales ont montré que le mélange de variétés présentant des niveaux contrastés de résistance permettait de limiter l'impact de maladies telles que les rouilles, la septoriose ou encore l'oïdium, et des règles d'assemblage permettant de maximiser ces effets positifs ont été établies ([de Vallavieille-Pope, 2004](#)). Bien que moins étudiés, d'autres services écosystémiques peuvent être fournis par les mélanges de variétés ([Barot *et al.*, 2017](#) ; [Borg *et al.*, 2018](#)). Plusieurs études ont porté sur la capacité des mélanges de blé à produire de

l'overyielding (OY), c'est-à-dire à avoir un rendement supérieur à la moyenne des rendements des variétés qui les composent, lorsque ces dernières sont cultivées en pur. Si l'OY est en moyenne positif pour les associations variétales de blé (Borg *et al.*, 2018), les OY négatifs ne sont pas rares. La conception des mélanges ne peut donc pas se faire au hasard et le développement de règles d'assemblage paraît essentiel. Cependant, les mécanismes permettant d'expliquer les performances très variables des mélanges sont encore mal compris. De nouvelles études sont donc nécessaires pour apporter des éléments de compréhension sur les interactions qui prennent place entre plantes dans les peuplements hétérogènes de blé. Il serait notamment intéressant d'étudier l'impact, sur ces interactions, de facteurs tels que la densité de semis, la proportion des différents géotypes et leur agencement, ou encore les architectures aérienne et racinaire des plantes.

L'architecture aérienne dans les peuplements hétérogènes de blé

Malgré la réduction de diversité fonctionnelle entraînée par la domestication et la sélection du blé tendre, une variabilité conséquente existe encore aujourd'hui entre les variétés pour certains traits de l'architecture aérienne (Cantarel *et al.*, 2020). Les traits de l'architecture aérienne, en modifiant les caractéristiques biologiques de la plante, mais aussi son environnement, jouent un rôle clef dans la compétition pour la lumière entre les individus (Niinemets, 2010). De ce fait, ils constituent un élément critique de la performance et de l'évolution des peuplements hétérogènes et des individus qui les composent. Dès lors, il apparaît important de comprendre l'impact de ces traits sur les interactions entre plantes afin de le prendre en compte dans la conception et la gestion des peuplements hétérogènes. Pourtant, l'impact de l'architecture aérienne sur la performance et l'évolution des peuplements hétérogènes de blé a été peu étudié (Borg *et al.*, 2018 ; Vidal *et al.*, 2020).

Impact de l'architecture aérienne sur la performance des peuplements hétérogènes de blé

Les quelques études proposant des règles sur la conception de peuplements hétérogènes de blé conseillent de mélanger des variétés présentant une certaine homogénéité, notamment en ce qui concerne la hauteur des plantes (Bowden *et al.*, 2001 ; de Vallavieille-Pope, 2004 ; Dai *et al.*, 2012). Ces recommandations sont faites pour éviter que la compétition pour la lumière entre les cultivars de la parcelle ne conduise à un effet de sélection important, qui entraînerait un déséquilibre dans les proportions des différents géotypes en mélange, et pourrait s'avérer problématique pour la performance du peuplement si le gain des plantes les plus hautes était moins important que la perte de rendement pour les plantes les plus courtes. Ces recommandations sont en accord avec celles formulées pour les associations variétales d'autres espèces, comme le montre l'exemple des mélanges variétaux d'orge (*Hordeum vulgare* L.), qui ne sont autorisés au Danemark que si les différences de hauteur entre les cultivars sont inférieures à 15 cm (Labarthe *et al.*, 2018). L'impact d'une diversité sur d'autres traits d'architecture aérienne que la hauteur est rarement évoqué dans les publications sur les peuplements hétérogènes de blé et peu d'études ont démontré expérimentalement un effet négatif de différences d'architecture. On note tout de même que l'étude de Montazeaud *et al.* (2020) a mis en évidence une corrélation négative entre la diversité sur la capacité de tallage et le rendement, dans un essai expérimental impliquant 197 mélanges binaires de

variétés de blé dur (*Triticum turgidum* ssp. *durum*). Aucun effet négatif de la diversité de hauteur dans les mélanges n'a cependant été trouvé dans cette même étude. Enfin, il a été proposé qu'une certaine homogénéité soit préférable dans les cultures, indépendamment de l'impact de la compétition pour la lumière, car des mécanismes de reconnaissances entre apparentés pourraient impacter négativement la performance de peuplements trop diversifiés, bien que de tels mécanismes n'aient pas été mis en évidence chez le blé (Fréville *et al.*, 2019).

D'autre part, d'après les théories issues de l'écologie, sur lesquelles l'agroécologie se base, plusieurs mécanismes pourraient entraîner une bonne performance des peuplements hétérogènes dans lesquels les plantes présentent des architectures aériennes contrastées. En effet, des effets de complémentarité d'interception de la lumière ou de facilitation pourraient prendre place, ainsi que des effets de sélection (Barot *et al.*, 2017). Pour que l'effet de sélection ait un impact positif sur la performance du peuplement, il faut cependant que le gain de rendement du génotype dominant soit plus important que la perte de rendement de l'autre génotype. Barot *et al.* (2017) soulignent également l'importance de la plasticité phénotypique pour accentuer le potentiel impact positif d'un effet de sélection. La plasticité phénotypique est le mécanisme par lequel un individu modifie son phénotype en réponse à des signaux environnementaux. Chez le blé, le trait d'architecture aérienne le plus plastique, en réponse à la compétition pour la lumière, est le tallage (propriété de ramification) (Lecarpentier, 2017). Le nombre de tiges d'une plante de blé dépend ainsi de son environnement lumineux. L'importante plasticité de tallage chez le blé pourrait donc favoriser un effet positif de la diversité (Barot *et al.*, 2017). De plus, une diversité sur des traits de l'architecture aérienne pourrait être intéressante pour améliorer la compétitivité envers les adventices, en particulier dans les systèmes à bas niveau d'intrants dans lesquels la pression adventices est plus importante (Borg *et al.*, 2018). D'autres mécanismes, indépendants de la compétition pour la lumière, pourraient également entraîner une bonne performance des peuplements hétérogènes avec une diversité d'architecture aérienne. Des différences de hauteur entre les plantes d'un peuplement pourraient ainsi limiter la verse, par un effet tuteur des plantes courtes pour les plantes hautes, ou encore structurer le couvert de façon à limiter les pertes d'eau par évaporation (Adu-Gyamfi *et al.*, 2015 ; Faraji, 2011). Enfin, le contrôle de certaines maladies pourrait être favorisé par le mélange de plantes présentant des hauteurs contrastées, comme cela a été montré dans les mélanges de riz (*Oryza sativa* L.) pour la pyriculariose (Zhu *et al.*, 2005), ainsi que dans des mélanges de blé pour la septoriose (Vidal *et al.*, 2017). On note cependant que la diversité de hauteur n'avait pas d'impact (positif ou négatif) sur la sévérité de la septoriose et de la rouille jaune dans une étude de Vidal *et al.* (2020) portant sur des mélanges de blé. L'effet positif de différences de hauteur sur la performance de mélanges de blé semble confirmé par la méta-analyse de Borg *et al.* (2018), qui a montré une meilleure performance des mélanges de variétés de blé hétérogènes pour la hauteur et la précocité de floraison que des mélanges homogènes.

Bien que les traits de l'architecture aérienne puissent donc avoir un fort impact sur la performance des peuplements hétérogènes de blé via de multiples mécanismes, trop peu d'études ont porté sur ces questions à ce jour. Les données sur le sujet sont aujourd'hui trop peu nombreuses, présentent parfois des résultats contradictoires et permettent généralement uniquement de décrire des effets, sans pour autant établir de lien de causalité avec les mécanismes sous-jacents. Une meilleure compréhension des interactions prenant place dans les peuplements hétérogènes de blé à architectures contrastées est donc nécessaire.

Impact de l'architecture aérienne sur l'évolution des peuplements hétérogènes de blé

Au-delà de son effet sur la performance des peuplements hétérogènes, l'architecture aérienne des plantes impacte également l'évolution de ces peuplements, ce qui soulève des questions de biologie évolutive importantes. En effet, le resemis des peuplements hétérogènes de blé est une pratique courante qui présente plusieurs avantages. On note tout d'abord que les semences de ferme, qui sont resemées d'une année sur l'autre, représentent une proportion importante des semences de blé tendre cultivées en France (proportion comprise environ entre 40% et 50% depuis 1980 ; Perronne *et al.*, 2016). De plus, la culture de variétés population, comme les populations composites, implique généralement leur resemis sur plusieurs générations. Les variétés population peuvent ainsi être intégrées à des programmes de gestion dynamique (Enjalbert *et al.*, 2011) ou de sélection ('*evolutionary plant breeding*' ; Döring *et al.*, 2011) basés sur le resemis, qui peuvent être mis en place à la ferme par les agriculteurs, ou dans des stations expérimentales. Tous les peuplements resemés d'une année sur l'autre sont soumis aux forces évolutives, telles que la dérive génétique, les mutations et la sélection naturelle. Cette méthode de gestion présente ainsi l'avantage de permettre aux peuplements de s'adapter aux conditions environnementales et aux pratiques de culture auxquelles ils sont soumis (Döring *et al.*, 2011 ; Enjalbert *et al.*, 2011). Il a également été montré que cette pratique permet d'augmenter la diversité génétique des populations (Khan *et al.*, 2020). En plus d'augmenter la diversité cultivée, cette pratique favorise également d'autres principes de l'agroécologie, notamment l'autonomie des paysans. Cependant, les caractéristiques évolutives des populations de blé resemées sont encore trop peu comprises et gagneraient à être étudiées davantage (Khan *et al.*, 2020). L'architecture aérienne des plantes devrait notamment être prise en compte, puisqu'elle est un élément critique de la performance individuelle, principalement via son impact sur la compétition pour la lumière. Les traits d'architecture aérienne devraient ainsi fortement impacter l'évolution des peuplements hétérogènes resemés.

Si une diversité architecturale pourrait favoriser la performance du peuplement à l'échelle d'une saison de culture, elle pourrait également conduire à une évolution défavorable au cours des générations de resemis. En effet, dans le cas d'un effet de sélection lié à des différences importantes de compétitivité entre les génotypes de la parcelle, on s'attend à ce que le génotype dominant participe plus que les autres à la production de grains utilisés pour le semis de la génération suivante. Il en découle une augmentation en fréquence des génotypes les plus performants au cours des générations qui peut être intéressante pour l'adaptation de la population aux conditions environnementales et de culture (Döring *et al.*, 2011). Cependant, lorsque la population subit une pression de sélection directionnelle importante, cela entraîne également une perte de diversité au cours des générations. Ainsi, une sélection directionnelle pour la hauteur est généralement observée dans les populations de blé resemées (Goldringer *et al.*, 2001a ; Raquin *et al.*, 2008 ; Döring *et al.*, 2015 ; Knapp *et al.*, 2020). Dans ces populations, les plantes les plus hautes, plus compétitives, interceptent davantage de lumière que les plantes plus courtes. Ces plantes envahissent ainsi la population au cours des générations de resemis, conduisant à une baisse de la diversité fonctionnelle, bien que cette dernière puisse être favorable à la performance du peuplement. La baisse de diversité se produit également au niveau génétique, et peut concerner d'autres loci que ceux impliqués dans le déterminisme du trait sous sélection (hauteur ou autres), comme en témoignent les faibles valeurs d'effectifs efficaces estimées dans les populations de blé resemés (Raquin *et al.*, 2008 ; Goldringer *et al.*, 2001a ; Brumlop *et al.*, 2019). Pour ce qui est des mélanges

de variétés, l'effet de sélection, associé au faible taux d'allogamie du blé, peut entraîner une perte de diversité plus importante encore, et aboutir, en quelques générations, à un peuplement homogène, constitué uniquement de la variété la plus compétitive. Parallèlement à cette perte de diversité, il est intéressant de souligner que la sélection naturelle conduit à l'augmentation en fréquence des génotypes présentant la meilleure fitness individuelle dans la parcelle (plantes les plus hautes par exemple). Cependant, l'agroécologie évolutive, aussi appelée agriculture Darwinienne, repose sur la théorie qu'il existe un compromis (*trade-off*) entre la performance individuelle et la performance à l'échelle du peuplement (Denison *et al.*, 2003 ; Weiner *et al.*, 2010). Un tel compromis a d'ailleurs été mis en évidence expérimentalement chez le blé (Weiner *et al.*, 2017). Les génotypes qui augmentent en fréquence au cours des générations ne seraient donc pas ceux qui favorisent la performance du peuplement, alors que c'est le rendement à l'échelle de la parcelle qui est visé dans le cas de l'agriculture. Il en résulte une perte de performance du peuplement au cours des générations, ainsi que cela est énoncé par la théorie de la tragédie des biens communs (Anten et Vermeulen, 2016). Ce phénomène a également été mis en évidence expérimentalement dans des populations de blé ressemées dans lesquelles la sélection directionnelle pour la hauteur était associée à une baisse de rendement au cours des générations (Goldringer *et al.*, 2001a). Une diversité sur d'autres traits de l'architecture aérienne, comme la surface foliaire ou la capacité de tallage, pourrait conduire à des phénomènes similaires (Anten et Vermeulen, 2016).

Bien qu'une évolution défavorable des peuplements de blé présentant une diversité architecturale soit possible, plusieurs leviers pourraient permettre d'éviter ou de limiter ce cas de figure. Il a ainsi été suggéré que la sélection humaine vienne contrebalancer la sélection naturelle afin de maintenir une diversité dans les populations (sélection de plantes courtes par le paysan par exemple), ou que la densité de semis soit optimisée afin de limiter la compétition pour la lumière entre les plantes de la parcelle (Goldringer *et al.*, 2001b). On peut également penser que certaines combinaisons de traits architecturaux permettraient d'intégrer une diversité d'architecture dans le peuplement, tout en conférant des niveaux de compétitivité comparables entre individus, évitant ainsi un effet de sélection trop important. Des études sont cependant nécessaires pour déterminer le potentiel de ces différentes pratiques.

Une diversité d'architecture aérienne dans les peuplements hétérogènes de blé pourrait ainsi avoir des impacts favorables ou défavorables sur leur performance et leur évolution. La prise en compte de l'architecture aérienne des plantes est donc centrale pour la conception et la gestion de peuplements hétérogènes de blé. Il apparaît ainsi essentiel d'étudier le rôle de l'architecture aérienne dans les interactions prenant place dans ces peuplements, afin de mieux comprendre l'impact des différents traits sur la performance et l'évolution des parcelles hétérogènes.

Une approche par modélisation

Afin de mieux comprendre l'impact des traits d'architecture aérienne sur la performance et l'évolution des peuplements hétérogènes de blé, l'approche par modélisation semble tout indiquée. Les expérimentations, comme celle de Montazeaud *et al.* (2020) portant sur l'impact de 19 traits fonctionnels, dont 5 traits de l'architecture aérienne, sur la performance de mélanges de blé dur,

apportent des éléments très intéressants, mais sont très lourdes à mettre en place. En ce qui concerne l'évolution des peuplements hétérogènes, les programmes en gestion dynamique ou en "*evolutionary plant breeding*" sont également très lourds à mettre en place et difficiles à maintenir sur de longues périodes de temps sans un support politique conséquent (Enjalbert *et al.*, 2011). La relative rapidité de l'approche par modélisation, comparativement aux expérimentations, permet de simuler l'évolution de populations de blé sur de nombreuses générations. De plus, cela permet de répondre en partie au casse-tête combinatoire associé à l'étude des mélanges, le nombre de combinaisons de variétés étant bien trop important pour pouvoir toutes les tester, même en ne considérant que les centaines de variétés inscrites au catalogue officiel français ou en se concentrant seulement sur des mélanges binaires. À l'inverse de l'approche expérimentale, la modélisation permet également d'isoler l'impact des différents traits étudiés en s'affranchissant d'éventuels trade-offs entre traits et en contrôlant complètement les conditions testées.

Dans une revue récente, Gaudio *et al.* (2019) distinguent deux grandes catégories de modèles permettant de simuler des cultures hétérogènes : les modèles de culture (*process-based crop models*) et les modèles individus-centrés (IBM, pour *Individual Based Models*). Les modèles de culture considèrent les interactions entre les plantes, leur environnement (sol, atmosphère) et les pratiques agricoles (date de semis ou fertilisation azotée par exemple) à l'échelle de la parcelle (Boote *et al.*, 2013). Les plantes constituant la parcelle ne sont pas simulées individuellement, la canopée étant représentée comme une plante moyenne en considérant des grandeurs telles que la surface de feuilles par m², ou la biomasse de feuilles ou de fruits. D'abord développés pour représenter des cultures pures, ces modèles ne représentaient alors aucune hétérogénéité horizontale dans le couvert, bien qu'une hétérogénéité verticale (plusieurs couches de feuilles) soit parfois considérée. Plus récemment, des modèles de culture ont été développés spécifiquement pour représenter des rangs alternés de différentes cultures. L'hétérogénéité horizontale de la canopée est alors simulée par une alternance de rangs homogènes. Les caractéristiques associées à chaque rang homogène, comme la densité de feuilles par exemple, peuvent être ajustées pour rendre compte des différences de traits entre les rangs. Les modèles de culture ne peuvent cependant pas représenter certaines des interactions spatio-temporelles complexes qui prennent place entre plantes dans les peuplements hétérogènes. Bien que plus complexes, les modèles individus-centrés sont plus adaptés pour simuler ce type d'interactions complexes (Zakharova *et al.*, 2019). Contrairement aux modèles de culture, les modèles individus-centrés (Zhang et De Angelis, 2020) représentent explicitement chaque individu de la parcelle. Chaque individu interagit avec son environnement et avec les individus qui l'entourent, et présente des caractéristiques propres, qui peuvent évoluer dans le temps. Ces modèles ont un fonctionnement *bottom-up* : les caractéristiques du peuplement émergent des interactions des individus entre eux et avec leur environnement. Les IBM sont particulièrement adaptés pour l'approche basée sur les traits (Zakharova *et al.*, 2019) et sont donc des modèles de choix pour étudier l'impact des traits d'architecture aérienne dans les peuplements hétérogènes. Plusieurs types d'IBM peuvent être utilisés pour simuler des peuplements hétérogènes de blé. Pour simuler la compétition entre les individus en considérant leur répartition spatiale, on peut notamment citer l'approche FON (*Field Of Neighborhood*) qui considère un rayon d'influence autour de chaque plante, dont l'importance dépend des caractéristiques de l'individu (Zhang et De Angelis, 2020). Les interactions entre individus se font ainsi lorsque leurs rayons d'influence se recoupent et en fonction de la distance entre les individus et des caractéristiques des plantes

concernées. Cette approche assez complète présente tout de même des limites et ne considère généralement ni l'impact de l'individu sur la disponibilité locale en ressources, ni la plasticité phénotypique.

Les modèles de plante structure-fonction (FSPM pour *Functional-Structural Plant Models*) simulent quant à eux explicitement l'architecture des individus en 3 dimensions et son développement dans le temps (Godin et Sinoquet, 2005 ; Vos et al., 2010). Dans ces IBM, le développement des plantes et de leur architecture dépend de processus écophysologiques qui sont impactés par les facteurs environnementaux simulés (lumière, disponibilité en nutriments...). Le développement de la structure de chaque plante impacte également leur environnement et celui des individus qui l'entourent, par des effets d'ombrage dans le cas de la lumière par exemple. Ces modèles permettent ainsi de simuler les interactions complexes entre les facteurs environnementaux, la structure des plantes, et leur développement. De ce fait, ils sont particulièrement adaptés pour simuler des cultures hétérogènes, en particulier pour mieux comprendre les interactions entre plantes qui y prennent place, pour identifier les traits impliqués dans leur performance, déterminer les combinaisons de traits optimales et étudier l'impact de l'agencement spatial des plantes dans la parcelle (Evers et al., 2019 ; Gaudio et al., 2019). Bien que cela soit peu développé, il est possible d'intégrer un lien entre le phénotype des plantes simulées avec un FSPM et leur génotype (Louarn et Song, 2020). Le couplage entre un FSPM et un modèle de génétique des populations serait ainsi un moyen intéressant d'étudier l'évolution de peuplements hétérogènes de blé.

Cependant, comme souligné par Gaudio et al. (2019), la modélisation des cultures hétérogènes est encore peu développée et les modèles ne sont pas encore complètement opérationnels. Un certain nombre de défis, associés à la complexité des FSPM, ont ainsi été listés par Louarn et Song (2020). Les auteurs indiquent ainsi que le développement et la maintenance prennent généralement plus de temps pour les FSPM que pour des modèles plus simples et que la communication autour du modèle et sa diffusion sont compliquées. Ils soulignent également le coût computationnel élevé des FSPM. Cet élément particulièrement important limite les méthodes qui peuvent être employées avec les FSPM, en particulier pour l'analyse de sensibilité du modèle et sa calibration. De plus, la calibration des FSPM peut être limitée par le peu de données disponibles. Tous ces éléments doivent donc être pris en compte dans les travaux utilisant un FSPM pour étudier les peuplements hétérogènes de blé. Il apparaît important d'intégrer le développement des FSPM et de méthodes adaptées à l'étude par modélisation des peuplements hétérogènes.

Présentation du plan de la thèse

Cette thèse a ainsi pour objectif de développer et d'utiliser des outils et méthodes de modélisation structure-fonction pour étudier l'impact de l'architecture aérienne sur la compétition pour la lumière dans des peuplements hétérogènes de blé et les conséquences de cette compétition sur l'évolution des peuplements. Ce travail est basé sur l'utilisation d'un modèle structure-fonction de blé pré-existant qui est présenté dans le premier chapitre. Ce chapitre présente également les adaptations réalisées sur le modèle pour qu'il puisse répondre aux exigences de nos études. Le FSPM utilisé n'étant pas entièrement déterministe, le deuxième chapitre de la thèse pose la question de l'aléatoire

du modèle, dont l'impact doit être évalué pour utiliser le modèle et analyser ses résultats correctement. Ce chapitre permet ainsi de fixer les conditions (nombre de plantes, utilisation de réplicats) des simulations de la thèse. Le troisième chapitre présente l'analyse de sensibilité du modèle, qui est une étape essentielle en modélisation et qui permet d'éclairer les résultats des simulations à venir. En évaluant la réponse des sorties du modèle face à des variations de ses entrées, l'analyse de sensibilité identifie ainsi les paramètres du modèle ayant le plus d'influence sur les résultats des simulations et dont les valeurs doivent être sélectionnées avec soin pour la suite des analyses. Pour les paramètres les plus influents, identifiés grâce à l'analyse de sensibilité, l'application d'une méthode de calibration automatique est présentée dans le chapitre 4. Sachant des données observées, la méthode proposée permet d'estimer une distribution de probabilités a posteriori pour ces paramètres, avec un coût computationnel limité. Le modèle est ensuite utilisé pour simuler des peuplements hétérogènes, et le cinquième chapitre présente un travail exploratoire pour identifier les traits de l'architecture aérienne qui ont le plus d'impact sur la performance de mélanges binaires équilibrés. Une phase d'optimisation permet également d'identifier les combinaisons de traits qui, dans les simulations, maximisent la performance à l'échelle du couvert. Par la suite, le sixième chapitre présente un modèle de génétique des populations et son couplage au FSPM utilisé pour simuler l'évolution de peuplements hétérogènes. Une expérience *in silico* permet de valider le fonctionnement de ce couplage en simulant l'évolution d'une population composite de blé, en se concentrant sur l'impact de différences de hauteur dans le couvert. Enfin, la dernière partie de la thèse est consacrée à la discussion générale des résultats obtenus dans les différents chapitres et présente les perspectives d'utilisation des outils et méthodes développés.

RÉFÉRENCES

- Adu-Gyamfi, P., Mahmood, T., & Trethowan, R. (2015).** Can wheat varietal mixtures buffer the impacts of water deficit?. *Crop and Pasture Science*, 66(8), 757-769.
- Altieri, M. A. (1999).** The ecological role of biodiversity in agroecosystems. In *Invertebrate Biodiversity as Bioindicators of Sustainable Landscapes* (pp. 19-31). Elsevier.
- Altieri, M. A. (2018).** *Agroecology: the science of sustainable agriculture*. CRC Press.
- Altieri, M. A., & Nicholls, C. I. (2020).** Agroecology and the reconstruction of a post-COVID-19 agriculture. *The Journal of Peasant Studies*, 47(5), 881-898.
- Anten, N. P., & Vermeulen, P. J. (2016).** Tragedies and crops: understanding natural selection to improve cropping systems. *Trends in Ecology & Evolution*, 31(6), 429-439.
- Barot, S., Allard, V., Cantarel, A., Enjalbert, J., Gauffreteau, A., Goldringer, I., ... & Porcher, E. (2017).** Designing mixtures of varieties for multifunctional agriculture with the help of ecology. A review. *Agronomy for sustainable development*, 37(2), 13.
- Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., ... & Vasseur, D. A. (2011).** Why intraspecific trait variation matters in community ecology. *Trends in ecology & evolution*, 26(4), 183-192.
- Bonnin, I., Bonneuil, C., Goffaux, R., Montalent, P., & Goldringer, I. (2014).** Explaining the decrease in the genetic diversity of wheat in France over the 20th century. *Agriculture, ecosystems & environment*, 195, 183-192.
- Boote, K. J., Jones, J. W., White, J. W., Asseng, S., & Lizaso, J. I. (2013).** Putting mechanisms into crop production models. *Plant, cell & environment*, 36(9), 1658-1672.
- Borg, J., Kiær, L. P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., ... & Enjalbert, J. (2018).** Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*, 221, 298-313.
- Bowden, R.L., Shroyer, J.P., Roozeboom, K., Claassen, M., Evans, P., Gordon, B., ... & Schlegel, A. J. (2001).** Performance of wheat variety blends in Kansas. *Kans. State Univ. Agric. Ext. Bull.* 128.
- Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F. X., & Huard, F. (2010).** Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Research*, 119(1), 201-212.
- Cantarel, A. A. M., Allard, V., Andrieu, B., Barot, S., Enjalbert, J., Gervaix, J., ... & Le Roux, X. (2020).** Plant functional trait variability and trait syndromes among wheat (*Triticum aestivum*) varieties: the footprint of artificial selection. *Journal of Experimental Botany*.
- Cianciaruso, M. V., Batalha, M. A., Gaston, K. J., & Petchey, O. L. (2009).** Including intraspecific variability in functional diversity. *Ecology*, 90(1), 81-89.

- Coumou, D., & Rahmstorf, S. (2012).** A decade of weather extremes. *Nature climate change*, 2(7), 491-496.
- Dai, J., Wiersma, J. J., & Holen Jr, D. L. (2012).** Performance of hard red spring wheat cultivar mixtures. *Agronomy Journal*, 104(1), 17-21.
- de Vallavieille-Pope, C. (2004).** Management of disease resistance diversity of cultivars of a species in single fields: controlling epidemics. *Comptes rendus biologies*, 327(7), 611-620.
- Denison, R. F., Kiers, E. T., & West, S. A. (2003).** Darwinian agriculture: when can humans find solutions beyond the reach of natural selection?. *The quarterly review of biology*, 78(2), 145-168.
- Díaz, S., & Cabido, M. (2001).** Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in ecology & evolution*, 16(11), 646-655.
- Döring, T. F., Annicchiarico, P., Clarke, S., Haigh, Z., Jones, H. E., Pearce, H., ... & Wolfe, M. S. (2015).** Comparative analysis of performance and stability among composite cross populations, variety mixtures and pure lines of winter wheat in organic and conventional cropping systems. *Field Crops Research*, 183, 235-245.
- Döring, T. F., Knapp, S., Kovacs, G., Murphy, K., & Wolfe, M. S. (2011).** Evolutionary plant breeding in cereals—into a new era. *Sustainability*, 3(10), 1944-1971.
- Duru, M., Therond, O., Martin, G., Martin-Clouaire, R., Magne, M. A., Justes, E., ... & Sarthou, J. P. (2015).** How to implement biodiversity-based agriculture to enhance ecosystem services: a review. *Agronomy for sustainable development*, 35(4), 1259-1281.
- Enjalbert, J., Dawson, J. C., Paillard, S., Rhoné, B., Rousselle, Y., Thomas, M., & Goldringer, I. (2011).** Dynamic management of crop diversity: From an experimental approach to on-farm conservation. *Comptes rendus biologies*, 334(5-6), 458-468.
- Evenson, R. E., & Gollin, D. (2003).** Assessing the impact of the Green Revolution, 1960 to 2000. *science*, 300(5620), 758-762.
- Evers, J. B., Van Der Werf, W., Stomph, T. J., Bastiaans, L., & Anten, N. P. (2019).** Understanding and optimizing species mixtures using functional–structural plant modelling. *Journal of experimental botany*, 70(9), 2381-2388.
- Faraji, J. (2011).** Wheat cultivar blends: A step forward to sustainable agriculture. *African Journal of Agricultural Research*, 6(33), 6780-6789.
- Food and Agriculture Organization of the United Nations. (2018).** FAOSTAT statistical database. Available at <http://www.fao.org/faostat/en/#data/QC> (accessed 25 August 2020).
- FranceAgriMer (2020).** Enquête répartition variétale des céréales.
- Fréville, H., Roumet, P., Rode, N. O., Rocher, A., Latreille, M., Muller, M. H., & David, J. (2019).** Preferential helping to relatives: A potential mechanism responsible for lower yield of crop variety mixtures?. *Evolutionary Applications*, 12(9), 1837-1849.
- Gale, M. D., Youssefian, S., & Russell, G. E. (1985).** Dwarfing genes in wheat. *Progress in plant breeding*, 1, 1-35.

Gaudio, N., Escobar-Gutiérrez, A. J., Casadebaig, P., Evers, J. B., Gerard, F., Louarn, G., ... & Barillot, R. (2019). Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agronomy for Sustainable Development*, 39(2), 20.

Gliessman, S. R. (2014). *Agroecology: the ecology of sustainable food systems*. CRC press.

Godin, C., Sinoquet, H. (2005). Functional–structural plant modelling. *New phytologist*, 166(3), 705-708.

Goldringer, I., Bonnin, I., & David, J. (2013). Rôle de la sélection dans la dynamique de la diversité des Blés cultivés: de la domestication à la sélection moderne au 20ème siècle. *Innovations Agronomiques*, 29, 1-15

Goldringer, I., Enjalbert, J., David, J., Paillard, S., Pham, J. L., & Brabant, P. (2001b). Dynamic management of genetics resources: a 13-year experiment on wheat. *Broadening the genetic base of crop production*. *IPGRI/FAO*, 245-260.

Goldringer, I., Enjalbert, J., Raquin, A. L., & Brabant, P. (2001a). Strong selection in wheat populations during ten generations of dynamic management. *Genetics Selection Evolution*, 33(1), S441-S463.

Gomiero, T. (2016). Soil degradation, land scarcity and food security: Reviewing a complex challenge. *Sustainability*, 8(3), 281.

Hatt, S., Artru, S., Brédart, D., Lassois, L., Francis, F., Haubruge, E., ... & Boeraeve, F. (2016). Towards sustainable food systems: the concept of agroecology and how it questions current research practices. A review. *Biotechnologie, Agronomie, Société et Environnement*, 20(Special issue 1), 215-224.

Hooper, D. U., Chapin Iii, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., ... & Schmid, B. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*, 75(1), 3-35.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology letters*, 11(6), 609-623.

Jackson, L. E., Pascual, U., & Hodgkin, T. (2007). Utilizing and conserving agrobiodiversity in agricultural landscapes. *Agriculture, ecosystems & environment*, 121(3), 196-210.

Kahiluoto, H., Kaseva, J., Balek, J., Olesen, J. E., Ruiz-Ramos, M., Gobin, A., ... & Bezak, P. (2019). Decline in climate resilience of European wheat. *Proceedings of the National Academy of Sciences*, 116(1), 123-128.

Khan, A. R., Goldringer, I., & Thomas, M. (2020). Management Practices and Breeding History of Varieties Strongly Determine the Fine Genetic Structure of Crop Populations: A Case Study Based on European Wheat Populations. *Sustainability*, 12(2), 613.

Knapp, S., Döring, T. F., Jones, H. E., Snape, J., Wingen, L. U., Wolfe, M. S., ... & Griffiths, S. (2020). Natural selection towards wild-type in composite cross populations of winter wheat. *Frontiers in Plant Science*, 10, 1757.

- Labarthe, P., Coleno, F., Fugerey-Scarbel, A., Hannachi, M., & Lemarié, S. (2018).** Freins et leviers socio-économiques à la diffusion des mélanges variétaux pour la production de blé: une comparaison entre France et Danemark.
- Lecarpentier, C. (2017).** Plasticité de l'architecture aérienne du blé en réponse à la compétition pour la lumière au sein de cultures pure ou d'associations variétales : caractérisation expérimentale et développement d'un modèle 3D. Thèse de doctorat, Université Paris-Saclay.
- Loreau, M. (2000).** Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, 91(1), 3-17.
- Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., ... & Tilman, D. (2001).** Biodiversity and ecosystem functioning: current knowledge and future challenges. *science*, 294(5543), 804-808.
- Louarn, G., & Song, Y. (2020).** Two decades of functional–structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany*, 126(4), 501-509.
- Malézieux, E. (2012).** Designing cropping systems from nature. *Agronomy for sustainable development*, 32(1), 15-29.
- Migliorini, P., & Wezel, A. (2017).** Converging and diverging principles and practices of organic agriculture regulations and agroecology. A review. *Agronomy for sustainable development*, 37(6), 63.
- Montazeaud, G., Violle, C., Roumet, P., Rocher, A., Ecarnot, M., Compan, F., ... & Fréville, H. (2020).** Multifaceted functional diversity for multifaceted crop yield: Towards ecological assembly rules for varietal mixtures. *Journal of Applied Ecology*, 57(11), 2285-2295.
- Niinemets, Ü. (2010).** A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecological Research*, 25(4), 693-714.
- Østergård, H., Finckh, M. R., Fontaine, L., Goldringer, I., Hoad, S. P., Kristensen, K., ... & Wolfe, M. S. (2009).** Time for a shift in crop production: embracing complexity through diversity at all levels. *Journal of the Science of Food and Agriculture*, 89(9), 1439-1445.
- Perronne, R., Hannachi, M., Lemarié, S., Fugerey-Scarbel, A., & Goldringer, I. (2016).** L'évolution de la filière blé tendre en France entre 1980 et 2006: quelle influence sur la diversité cultivée. *Notes et Etudes Socio-Economiques*, (41), 83-113.
- Raquin, A. L., Brabant, P., Rhoné, B., Balfourier, F., Leroy, P., & Goldringer, I. (2008).** Soft selective sweep near a gene that increases plant height in wheat. *Molecular Ecology*, 17(3), 741-756.
- Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., ... & Scholes, B. (2008).** Greenhouse gas mitigation in agriculture. *Philosophical transactions of the royal Society B: Biological Sciences*, 363(1492), 789-813.

- Tilman, D., Balzer, C., Hill, J., & Belfort, B. L. (2011).** Global food demand and the sustainable intensification of agriculture. *Proceedings of the national academy of sciences*, 108(50), 20260-20264.
- van der Plas, F. (2019).** Biodiversity and ecosystem functioning in naturally assembled communities. *Biological Reviews*, 94(4), 1220-1245.
- Van Frank, G. (2018).** *Gestion participative de la diversité cultivée et création de mélanges diversifiés de blé tendre à la ferme* (Doctoral dissertation, Université Paris-Saclay (ComUE)).
- Vidal, T., Boixel, A. L., Durand, B., de Vallavieille-Pope, C., Huber, L., & Saint-Jean, S. (2017).** Reduction of fungal disease spread in cultivar mixtures: impact of canopy architecture on rain-splash dispersal and on crop microclimate. *Agricultural and Forest Meteorology*, 246, 154-161.
- Vidal, T., Saint-Jean, S., Lusley, P., Leconte, M., Ben Krima, S., Boixel, A. L., ... & de Vallavieille-Pope, C. (2020).** Cultivar mixture effects on disease and yield remain despite diversity in wheat height and earliness. *Plant Pathology*.
- Violle, C., Enquist, B. J., McGill, B. J., Jiang, L. I. N., Albert, C. H., Hulshof, C., ... & Messier, J. (2012).** The return of the variance: intraspecific variability in community ecology. *Trends in ecology & evolution*, 27(4), 244-252.
- Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007).** Let the concept of trait be functional!. *Oikos*, 116(5), 882-892.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- Weiner, J., Andersen, S. B., Wille, W. K. M., Griepentrog, H. W., & Olsen, J. M. (2010).** Evolutionary Agroecology: the potential for cooperative, high density, weed-suppressing cereals. *Evolutionary Applications*, 3(5-6), 473-479.
- Weiner, J., Du, Y. L., Zhang, C., Qin, X. L., & Li, F. M. (2017).** Evolutionary agroecology: individual fitness and population yield in wheat (*Triticum aestivum*). *Ecology*, 98(9), 2261-2266.
- Wezel, A., Bellon, S., Doré, T., Francis, C., Vallod, D., & David, C. (2009).** Agroecology as a science, a movement and a practice. A review. *Agronomy for sustainable development*, 29(4), 503-515.
- Wezel, A., Herren, B. G., Kerr, R. B., Barrios, E., Gonçalves, A. L. R., & Sinclair, F. (2020).** Agroecological principles and elements and their implications for transitioning to sustainable food systems. A review. *Agronomy for Sustainable Development*, 40(6), 1-13.
- Zakharova, L., Meyer, K. M., & Seifan, M. (2019).** Trait-based modelling in ecology: A review of two decades of research. *Ecological Modelling*, 407, 108703.
- Zhang, B., & DeAngelis, D. L. (2020).** An Overview of Agent-Based Models in Plant Biology and Ecology. *Annals of Botany*, 126(4), 539–557.
- Zhu, Y. Y., Fang, H., Wang, Y. Y., Fan, J. X., Yang, S. S., Mew, T. W., & Mundt, C. C. (2005).** Panicle blast and canopy moisture in rice cultivar mixtures. *Phytopathology*, 95(4), 433-438.

CHAPITRE 1

Adaptations du FSPM WALTer pour l'étude de peuplements hétérogènes de blé

INTRODUCTION

L'augmentation de la biodiversité cultivée est un levier majeur de l'agroécologie (Malézieux, 2012 ; Isbell, 2015). Parmi les approches possibles, l'augmentation de la diversité à l'échelle de la parcelle, avec la culture de mélanges d'espèces ou de variétés par exemple, pourrait participer à la transition vers des agrosystèmes plus durables. Les contraintes importantes associées à l'étude expérimentale de ces peuplements hétérogènes (coût, nombre de combinaisons très élevé, etc...), font de la modélisation en général, et de la modélisation structure fonction (Vos *et al.*, 2010 ; Godin & Sinoquet, 2005) en particulier, un outil privilégié pour étudier ces cultures (Evers *et al.*, 2019 ; Gaudio *et al.*, 2019 ; Muller et Martre, 2019). Ces dernières années, des modèles structure-fonction (FSPM) ont ainsi été développés ou adaptés pour simuler des parcelles diversifiées. Barillot *et al.* (2014) ont, par exemple, utilisé la plateforme OpenAlea (Pradal *et al.*, 2008) pour coupler un modèle de blé (*Triticum aestivum* L.) à un modèle de pois (*Pisum sativum* L.) et étudier la répartition de la lumière en fonction de variations d'architecture aérienne dans des mélanges blé-pois. Le système racinaire a également fait l'objet d'études avec, par exemple, les travaux de Postma et Lynch (2012), qui ont utilisé le FSPM SimRoot (Lynch *et al.*, 1997; Postma et Lynch, 2011) pour simuler des mélanges de maïs (*Zea mays* L.), haricots (*Phaseolus vulgaris* L.) et courges (*Cucurbita pepo* L.), et ainsi étudier l'impact de l'architecture racinaire sur l'assimilation de nutriments (azote, phosphore et potassium) et la production de biomasse. Verdenal *et al.* (2008) ont, pour leur part, développé un modèle visant à simuler des mélanges intra-spécifiques de différents génotypes de ray-grass (*Lolium perenne* L.) afin d'étudier l'impact de la compétition pour la lumière et de la coupe (pâturage ou tonte) sur le développement des prairies.

Pour ce qui est de la modélisation structure-fonction du blé, il est possible d'utiliser des modèles génériques, développés pour simuler, de manière simple, un grand nombre d'espèces. Avec des données de calibration adéquates, il est ainsi possible de modéliser le système racinaire du blé en utilisant ArchiSimple (Pagès *et al.*, 2014) ou d'utiliser le modèle développé par Evers et Bastiaans (2016) sur la plateforme GroIMP (Hemmerling *et al.*, 2008) pour simuler la photosynthèse dans des couverts de graminées. Ces modèles génériques ont l'avantage de simplifier la simulation de peuplements pluri-spécifiques, mais ils représentent généralement moins bien les spécificités du blé qu'un modèle spécifiquement développé pour cette espèce et sont donc moins adaptés à la simulation de mélanges intra-spécifiques. Parmi les FSPM adaptés spécifiquement au blé, ADEL-blé a été développé par Fournier *et al.* (2003) pour modéliser le développement de l'architecture aérienne du blé d'hiver et peut être utilisé pour simuler des mélanges de variétés, comme cela a été fait, par exemple, pour étudier l'impact des différences de hauteur entre plantes sur la dispersion des spores de *Zymoseptoria tritici* (champignon responsable de la septoriose) dans des mélanges binaires (Vidal *et al.*, 2018). Une limite de ce modèle est cependant sa modélisation empirique du tallage. Le tallage est la propriété de ramification présente chez certaines graminées, dont le blé, qui leur permet de former des tiges secondaires, appelées talles, par le débourrement de certains bourgeons axillaires de la plante. Ce processus, qui comporte trois étapes (émission de talles, arrêt de l'émission et régression de certaines talles ; Xie *et al.*, 2016), est capital car il influence directement le nombre d'épis produits par la plante, et donc le rendement. En outre, une modélisation fiable de ce processus est essentielle pour l'étude de l'impact de différences architecturales sur le développement des mélanges variétaux car le tallage est le trait le plus

plastique en réponse à la compétition pour la lumière chez le blé (Lecarpentier, 2017 ; Lecarpentier *et al.*, in-prep). Le modèle ADEL-blé a ainsi été adapté successivement en 2007 et en 2010 par Evers *et al.* pour modéliser les deux premières étapes du tallage par une approche plus mécaniste, en rendant compte de l'impact de la compétition pour la lumière sur la plasticité du tallage (Assuero et Tognetti, 2010). En 2019, Lecarpentier *et al.* ont développé un nouveau FSPM, appelé WALTer, qui s'appuie sur ADEL-blé pour le développement foliaire des plantes, mais propose un nouveau formalisme pour la régulation des trois étapes du tallage en réponse à la compétition pour la lumière. Dans sa version initiale (Lecarpentier *et al.*, 2019), WALTer permet ainsi de reproduire, de façon réaliste, les différences de tallage dans des parcelles monovariétales dans des conditions contrastées de compétition pour la lumière (simulées par des densités de semis variées) et permet également de simuler des mélanges binaires simples. WALTer a été développé spécifiquement dans l'optique d'étudier les questions associées à l'impact de l'architecture aérienne sur la performance et l'évolution de peuplements hétérogènes de blé et apparaît ainsi tout indiqué pour aborder les questions soulevées dans cette thèse.

Si les formalismes implémentés dans WALTer sont globalement bien adaptés aux simulations nécessaires à ce projet de thèse, certaines contraintes du modèle doivent cependant être surmontées. Comme beaucoup de FSPM, le temps de simulation de WALTer est important et, bien que le formalisme simple de ce modèle permette de limiter en partie la durée des simulations, le coût computationnel pourrait encore être optimisé, ce qui permettrait d'augmenter le nombre de simulations réalisables. D'autre part, le réalisme et la précision du modèle peuvent être améliorés. En effet, dans l'optique de simuler des mélanges de variétés, il est important que la compétition pour la lumière soit simulée avec précision. Enfin, quelques adaptations sont nécessaires pour rendre possibles certaines des simulations prévues (mélanges complexes, couplage avec un modèle de génétique des populations, etc..).

Ce chapitre présente les principales adaptations réalisées à partir du modèle WALTer (Lecarpentier *et al.*, 2019) pour aboutir à la version utilisée dans l'ensemble de cette thèse.

MATÉRIEL ET MÉTHODES

WALTer (Lecarpentier *et al.*, 2019)

Le modèle structure fonction (FSPM) WALTer (Lecarpentier *et al.*, 2019) permet de simuler le développement de l'architecture aérienne du blé d'hiver (*Triticum aestivum* L.) en 3D du semis à la maturité avec un pas de temps journalier. Dans le modèle, des éléments descriptifs basés sur le modèle ADEL-blé (Fournier *et al.*, 2003) permettent de représenter le développement foliaire et l'allongement des gaines et des entre-nœuds, de manière déterministe, en fonction du temps thermique. Le temps thermique (Monteith, 1984) est calculé comme une somme des différences entre la température moyenne journalière et la température de base de l'organisme considéré (température en dessous de laquelle l'organisme ne se développe pas) et s'exprime en degrés-jours (°Cd). Dans WALTer, comme dans de nombreux modèles de blé (Ritchie et Otter, 1985 ; Jamieson *et al.*, 1998 ; Fournier *et al.*, 2003), la température de base est fixée à 0°C pour l'ensemble du cycle de développement de la plante. WALTer modélise également la régulation du tallage en réponse à la

compétition pour la lumière grâce à un formalisme basé sur des courbes dose-réponse (Evers and Vos, 2013) : plus une plante est en compétition lumineuse avec ses voisines, moins elle aura de talles à maturité. Pour cela, WALTer s'appuie sur un modèle de rayonnement, CARIBU (Chelle and Andrieu, 1998), qui permet de connaître la quantité de lumière interceptée par chaque organe quotidiennement. Pour ce faire, le modèle de rayonnement utilise des données météo de rayonnement journalier ainsi qu'un ciel virtuel divisé en secteur. Chaque secteur du ciel représente une source lumineuse caractérisée par une direction et une intensité. Le processus de tallage peut être simplifié en trois étapes (Xie *et al.*, 2016), qui sont toutes modélisées par WALTer :

1) La phase d'émission des talles, pendant laquelle la plante génère des tiges secondaires, est modélisée dans WALTer par un processus partiellement stochastique. Le rythme d'émission des talles est synchronisé avec les dynamiques d'émergence des feuilles sur le brin maître (Kirby *et al.*, 1985), et un bourgeon a une probabilité P_T d'émerger, P_T étant un paramètre d'entrée de WALTer. Ce paramètre de probabilité est proche de 1, rendant l'émergence quasi systématique, et est commun à toutes les talles, à l'exception de la talle du coléoptile qui a une probabilité moindre d'émerger (P_{CT}).

2) L'arrêt du tallage, stade où le débourrement de tous les bourgeons de la plante est inhibé, est contrôlé dans WALTer par une perception précoce du voisinage. En effet, l'arrêt du tallage se fait à l'échelle de la plante, chaque plante arrêtant d'émettre de nouvelles talles lorsque le Green Area Index (GAI^a) des plantes composant son voisinage dépasse une valeur seuil : le GAI_c . Le voisinage d'une plante est défini comme l'ensemble des plantes situées à moins d'une distance fixe d_{GAIp} de cette plante. Dans ce formalisme, le GAI_c sert de proxy pour le ratio Lumière rouge claire/ Lumière rouge sombre (Smith, 1982), un indicateur de l'absorption de lumière par les plantes voisines. En effet, les plantes sont capables de détecter des variations de ce ratio et d'adapter leur développement en fonction de ce signal de qualité de la lumière (Casal, 2013) qui est très corrélé au GAI (Sattin *et al.*, 1994). Cette régulation détermine généralement le nombre maximal d'axes^b produits, mais la transition florale (lorsque toutes les feuilles du brin maître ont été initiées) peut également arrêter l'émission des talles avant que le GAI_c ne soit atteint.

3) La phase de régression des talles est régulée dans WALTer par la quantité de lumière interceptée par chaque talle. La régression des talles d'une plante est autorisée lorsque cette plante atteint un âge physiologique fixé en entrée du modèle : t_{beg}^{reg} . La décision de régression est ensuite prise à l'échelle de la talle : une talle meurt lorsque la quantité de rayonnement photosynthétiquement actif (PAR) qu'elle intercepte par unité de surface est inférieure à une valeur seuil : PAR_c . Cependant, après la mort d'une talle, un délai de protection Δ_{prot} doit s'écouler avant qu'une autre talle ne puisse régresser sur la même plante.

Le formalisme de régulation du tallage implémenté dans WALTer permet de n'utiliser le modèle de rayonnement (CARIBU) que pendant la phase de régression des talles, ce qui limite le coût computationnel des simulations. Il est cependant possible d'étendre l'utilisation de CARIBU à l'ensemble de la simulation pour étudier l'interception lumineuse de la parcelle pendant l'ensemble de son cycle de développement.

a Le GAI d'une plante est le rapport entre la surface photosynthétique de cette plante et la surface de sol qu'elle couvre.

b Axe : le brin maître et les talles

Il est important de noter que WALTer n'est pas un modèle entièrement déterministe. En effet, afin de représenter l'hétérogénéité observée dans l'environnement, plusieurs sources de stochasticité sont intégrées. En plus de son rôle dans la phase d'émission des talles, l'aléatoire intervient ainsi dans la détermination de (1) la position des plantes dans la parcelle et l'orientation des organes (feuilles et talles) dans l'espace, (2) la date d'émergence des plantes et (3) le nombre final de feuilles sur le brin maître de chaque plante. Il en résulte que, pour un même jeu de paramètres, les sorties du modèle peuvent différer d'une simulation à l'autre (cf. Chapitre 2).

Le choix des valeurs pour les différentes entrées de WALTer permet de représenter des situations variées. Il est ainsi possible de faire varier trois types d'entrées dans le modèle :

- les données météorologiques (température et rayonnement), pour représenter des conditions météorologiques contrastées ;
- les paramètres descriptifs de la parcelle à simuler (nombre de plantes, densité de semis...), permettant de simuler des parcelles allant du peuplement très dense à la plante isolée ;
- les paramètres génotypiques (hauteur finale du brin maître, probabilité d'émergence des talles...), permettant de représenter des plantes aux caractéristiques diverses.

Une liste des différentes entrées du modèle est présentée en Annexe 1.

Évolutions du modèle

Si WALTer est globalement adapté aux besoins de cette thèse, certaines modifications ont toutefois dû être apportées au modèle afin de mener à bien les analyses. De ce fait, la version utilisée dans cette thèse (3.0) diffère partiellement de celle publiée en 2019 par [Lecarpentier et al.](#)

Grâce à la généreuse collaboration de Christian Fournier (UMR LEPSE, France) et Christophe Pradal (UMR AGAP, France), WALTer a été intégré à l'incubateur OpenAlea. OpenAlea ([Pradal et al., 2008](#)) est une plateforme logicielle libre pour la modélisation structure-fonction des plantes qui intègre des outils de simulation et d'analyse. L'intégration de WALTer dans le projet OpenAlea permet de bénéficier d'un environnement de développement privilégié et d'accéder à l'expertise des chercheurs en modélisation des plantes en favorisant la collaboration avec des équipes de recherche à l'international. La participation à ce projet permet également d'augmenter la visibilité de WALTer et de favoriser sa diffusion. WALTer a ainsi dû être modifié pour répondre aux critères de qualité d'OpenAlea, notamment avec des adaptations de son architecture et la mise en place de tests qui améliorent la fiabilité du modèle et simplifient les processus de développement.

L'aide de C.Fournier et de C.Pradal a également été précieuse pour réaliser les autres modifications du modèle nécessaires au bon déroulement de cette thèse.

Tout d'abord, comme le projet de thèse implique un nombre important de simulations, un travail important a été réalisé pour limiter la durée de chaque simulation. L'infinetisation des parcelles, intégrée dans le modèle de rayonnement ([Chelle and Andrieu, 1998](#)) grâce à des répliques virtuelles qui permettent de simuler un couvert périodique composé d'un ensemble fini de plantes, a ainsi été

mise en place. Cette amélioration permet de réduire le nombre de plantes à simuler, en comparaison de l'ancienne version de WALTer, puisqu'il n'est plus nécessaire de simuler des plantes supplémentaires pour s'affranchir des effets de bordure. La représentation des plantes a également été simplifiée par rapport à l'ancienne version de WALTer. Seules les surfaces visibles sont désormais représentées dans le modèle et l'exclusion des organes non visibles (gaines, entre-nœuds ou limbes emboîtés les uns dans les autres) permet un gain de temps de simulation. Enfin, la parallélisation a été implémentée lorsque plusieurs simulations sont lancées.

D'autre part, le réalisme et la précision du modèle ont été améliorés. Ainsi, si les feuilles étaient auparavant droites et représentées par la juxtaposition d'un carré et d'un triangle isocèle, celles de la nouvelle version de WALTer sont courbes et représentées sur la base du modèle de [Christian Fournier et Christophe Pradal \(2012\)](#). La forme des limbes est désormais basée sur les équations de [Dornbusch *et al.* \(2009\)](#) et leur courbure dépend du formalisme décrit par [Perez *et al.* \(2016\)](#). De nouveaux paramètres permettent de faire varier la forme, l'inclinaison et la courbure des feuilles et ainsi de reproduire des génotypes ayant un port érectophile à planophile.

De plus, la discrétisation du ciel a été optimisée par l'utilisation d'un ciel TURTLE ([den Dulk, 1989](#)) tel qu'implémenté dans le package python [Alinea.astk](#) (version 2.1.0, 2019). Il est possible de modifier le nombre de secteurs considérés avec ce nouveau ciel (1, 6, 16 ou 46 secteurs) afin d'ajuster la précision du rayonnement simulé et le temps de simulation. Il est également possible de changer le type de rayonnement simulé pour représenter, au choix, un ciel couvert ou un ciel clair.

Enfin, de nouvelles fonctionnalités ont été ajoutées au modèle dans sa nouvelle version. Le modèle a en effet été adapté afin d'être capable de simuler des mélanges de génotypes avec des valeurs de paramètres génotypiques pouvant différer pour chaque variété. La nouvelle version de WALTer permet ainsi de représenter un nombre illimité de génotypes (il est possible de simuler autant de génotypes différents que de plantes dans la parcelle). Les proportions des différents génotypes dans la parcelle sont contrôlées en entrée du modèle et peuvent soit : (i) dépendre d'un tirage aléatoire dans lequel la probabilité de tirage de chaque génotype est identique ; (ii) être toutes similaires, avec un nombre de plantes identique pour chaque génotype (à 1 plante près en fonction du nombre total de plantes dans la parcelle) ; (iii) être personnalisées par l'utilisateur en spécifiant en entrée le nombre de plantes associé à chaque génotype. La nouvelle version de WALTer permet donc de simuler des mélanges de variétés plus complexes qu'auparavant. En effet, avant ces améliorations, le modèle permettait uniquement la simulation de mélanges binaires (mélanges de 2 variétés) dans lesquelles seules les hauteurs pouvaient différer entre génotypes et les proportions ne pouvaient dépendre que d'un tirage aléatoire.

Module de fitness

Un module de fitness^c a également été ajouté au modèle afin de générer un proxy du nombre de grains produits. L'objectif de ce module est de produire une sortie à l'échelle de la plante qui soit un indicateur du succès reproducteur de l'individu. Une fitness est ainsi calculée pour chaque plante de la parcelle. La fitness d'une plante dans WALTer dépend de son interception lumineuse et

^c La fitness, ou valeur sélective, est un indicateur utilisé principalement en biologie évolutive pour mesurer la contribution d'un individu au pool de gènes de la génération suivante.

correspond à la somme des fitness de ses axes. La fitness de chaque axe est calculée selon la formule suivante :

$$F_{\text{axe}} = \frac{PAR_{t45}}{T_{m45}}$$

dans laquelle F_{axe} est la fitness de l'axe (exprimée en $\mu\text{molPAR} \cdot ^\circ\text{Cd}^{-1}$), PAR_{t45} est la quantité journalière moyenne de PAR intercepté par l'axe pendant une période de 45 jours avant la floraison (en $\mu\text{molPAR} \cdot \text{d}^{-1}$) et T_{m45} est la température moyenne (en $^\circ\text{C}$) pendant cette même période. Cette formule est une adaptation du quotient photothermique (PTQ ; Nix, 1976), qui correspond au ratio entre la rayonnement incident quotidien et la température journalière moyenne. Il a en effet été démontré que la valeur de PTQ, calculée pour des périodes proches de la date de floraison est linéairement corrélée au nombre de grains produits par unité de surface chez le blé (Midmore *et al.*, 1984 ; Fischer, 1985 ; Abbate *et al.*, 1995 ; Demotes-Mainard and Jeuffroy, 2001).

PTQ est ici calculé sur une période de 45 jours avant la floraison de l'axe, mais la période (F_{period}) utilisée pour le calcul du quotient photothermique fait l'objet de débats (Midmore *et al.*, 1984 ; Fischer, 1985 ; Abbate *et al.*, 1995 ; Ortiz-Monasterio *et al.*, 1994 ; Demotes-Mainard and Jeuffroy, 2001 ; Nalley *et al.*, 2009). Afin de tester l'impact de F_{period} sur les fitness calculées par le module, une analyse de sensibilité (AS) (Iooss et Saltelli, 2017) a été réalisée. Pour ce faire, les simulations de parcelles monovariétales utilisées pour l'AS globale décrite dans le Chapitre 3 ont été utilisées. Pour chacune de ces simulations, la fitness a été calculée en se basant sur 3 périodes différentes qui ont chacune été largement utilisées pour le calcul de PTQ : une période de 45 jours avant la floraison (PTQ_{45}) (Demotes-Mainard and Jeuffroy, 2001 ; Ratjen *et al.*, 2012), une période de 30 jours avant la floraison (PTQ_{30}) (Fischer, 1985 ; Midmore *et al.*, 1984) et une période de 20 jours avant et 10 jours après la floraison (PTQ_{20+10}) (Abbate *et al.*, 1995 ; Lázaro et Abbate 2012). Les valeurs de fitness moyenne par plante et d'écart-type de fitness entre plantes ainsi obtenues ont été comparées en fonction de F_{period} (PTQ_{45} , PTQ_{30} ou PTQ_{20+10}) et un test de Kruskal-Wallis a été réalisé. L'hypothèse nulle (H_0) de ce test est qu'il n'existe pas de différence significative entre les fitness (moyenne et écart-type) calculées avec différents F_{period} .

Des indices de sensibilité ont également été calculés via un modèle d'analyse de la variance permettant d'intégrer les interactions d'ordre 1 et 2 ainsi qu'une résiduelle, tel que décrit dans le Chapitre 3. Ces indices permettent d'estimer l'impact du choix de F_{period} sur la fitness moyenne et l'écart-type de la fitness. Les indices étant calculés pour toutes les entrées du plan factoriel fractionnaire, ils permettent également de connaître l'impact relatif du choix de F_{period} par rapport à des changements de valeurs pour d'autres entrées du modèle. Enfin, les indices de sensibilité permettent d'estimer les interactions entre le choix de F_{period} et l'impact des autres entrées sur le calcul de la fitness. Les résultats des indices de sensibilité sont présentés dans l'Annexe 1.

Simulation de stress

Une nouvelle fonctionnalité, permettant de simuler des stress (représentés par la mort d'une partie des plantes de la parcelle), a également été ajoutée au modèle. En effet, dans la version initiale de

WALTer (Lecarpentier *et al.*, 2019) toutes les plantes semées dans la parcelle virtuelle atteignent la maturité, car aucune mortalité d'individus n'est considérée. La nouvelle version du modèle permet de fixer en entrée le nombre de plantes de la parcelle qui sont sensibles au stress ainsi que le moment où le stress se manifeste. Ainsi, à la date à laquelle le stress doit se manifester, le modèle simule la mort de toutes les plantes identifiées comme sensibles. Cette nouvelle fonctionnalité permet d'étudier la capacité de compensation offerte par le tallage en fonction du moment du stress et de la proportion des plantes impactées. La simulation de stress n'est pas étudiée dans cette thèse, mais fait l'objet de travaux en collaboration avec d'autres collègues (Lecarpentier *et al.*, in-prep).

RÉSULTATS

WALTer et adaptation du modèle

WALTer permet désormais de simuler des mélanges complexes, comme illustré en Figure 1 avec une parcelle de 33 plantes composée de 8 géotypes différant par leur hauteur finale et leur port foliaire. Les adaptations réalisées sur le modèle ont permis de réduire significativement le coût computationnel des simulations. Le temps d'une simulation a ainsi été réduit d'un facteur 9 grâce, notamment, à la parallélisation des simulations et à l'infinetisation du couvert, qui est utilisée pour toutes les simulations de cette thèse. Le réalisme du modèle a également été amélioré par la modification de la représentation des feuilles. Pour les nouveaux paramètres de forme de limbes, les valeurs utilisées correspondent à celles identifiées par Dornbusch *et al.* (2011b) pour le cultivar Soissons.



Figure 1. Parcelle virtuelle d'un mélange variétal de blé simulé avec la nouvelle version de WALTer à 3 stades de développement différents (de gauche à droite : 80, 200 et 300 jours après le semis). La parcelle représentée est composée de 33 plantes semées à une densité de 170 plantes/m². La parcelle est composée de 8 géotypes différents, avec des hauteurs finales et des ports foliaires contrastés. La sénescence, simulée par WALTer, n'est pas représentée dans cette image.

De plus, l'utilisation du ciel TURTLE permet d'estimer avec plus de précision la quantité de lumière interceptée par chaque organe de la parcelle virtuelle. Pour toutes les simulations de cette thèse, une discrétisation du ciel à 16 secteurs est utilisée. En effet, étant donné la taille limitée des parcelles simulées, les sorties de WALTer obtenues avec une discrétisation à 16 secteurs sont très peu

différentes de celles obtenues avec une discrétisation à 46 secteurs (Annexe 1 - Figures 1 et 2). De plus, l'utilisation d'un nombre moindre de secteurs permet de limiter le temps de simulation.

Enfin, pour toutes les simulations de cette thèse, seul du rayonnement diffus a été considéré, ce qui correspond à un ciel couvert ("Standard OverCast", Moon et Spencer, 1942).

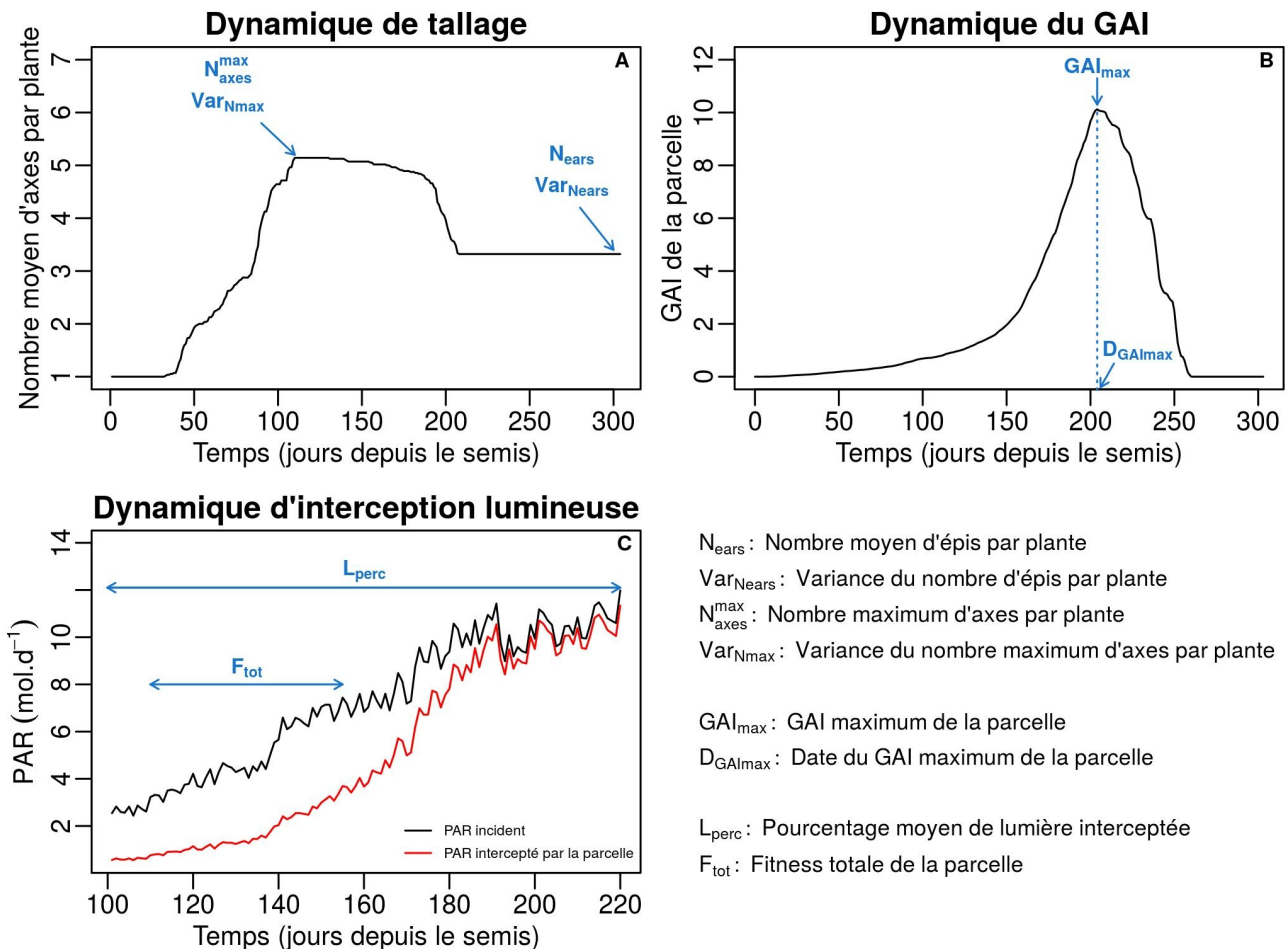


Figure 2. Sorties brutes et résumées extraites de WALTer pour la dynamique de tallage (A), la dynamique de GAI (B) et la dynamique d'interception lumineuse (C). L'utilisation du modèle de rayonnement n'étant pas nécessaire pour toutes les étapes de la simulation, la dynamique d'interception lumineuse n'est connue que pour une partie du cycle de développement. Ces figures ont été réalisées à partir des sorties d'une simulation de 56 plantes semées à une densité de 200 plantes/m². Les valeurs par défaut des entrées ont été utilisées (Annexe 1).

Les simulations réalisées avec WALTer génèrent de nombreuses sorties, pour chaque pas de temps, à l'échelle de la plante, de l'axe et de l'organe. Un pré-traitement des données est donc nécessaire afin de résumer les informations extraites des simulations, et les valeurs moyennes des sorties, à l'échelle de la parcelle, sont souvent utilisées dans nos analyses. Trois sorties dynamiques moyennes (Figure 2) sont ainsi utilisées dans les analyses de cette thèse : (i) la dynamique de tallage (nombre moyen d'axes par plante en fonction du temps), (ii) la dynamique du GAI de la parcelle et (iii) la dynamique de la quantité de lumière interceptée par la parcelle. Étant donné les difficultés liées à l'utilisation de sorties dynamiques dans les analyses, des sorties résumées ont été définies. Ces sorties résumées permettent d'extraire des informations importantes à partir des sorties dynamiques et ont l'avantage d'être des scalaires. De plus, elles correspondent à des éléments biologiques facilement interprétables. La dynamique de tallage moyenne peut ainsi être décrite par le nombre maximum d'axes produit par plante (N_{axes}^{max}) et sa variance (entre plantes de la parcelle) ($Var_{N_{max}}$)

ainsi que par le nombre final d'épis produits par plante (N_{ears}) et sa variance ($\text{Var}_{N_{\text{ears}}}$). Le GAI maximum atteint par la parcelle (GAI_{max}) et la date à laquelle ce GAI est atteint ($D_{\text{GAI}_{\text{max}}}$) permettent de résumer une partie de l'information portée par la dynamique de GAI moyenne. Enfin, deux sorties résumées sont utilisées pour décrire la dynamique d'interception lumineuse : le pourcentage moyen de lumière interceptée par la parcelle (L_{perc}), qui est calculé sur toute la période d'utilisation du modèle de rayonnement, et la fitness totale de la parcelle (F_{tot}), qui correspond à une somme des fitness de toutes les plantes de la parcelle. Le module de fitness permet ainsi, en plus de comparer les plantes ou les génotypes au sein d'une parcelle, d'estimer la performance à l'échelle du peuplement. La fitness moyenne des plantes de la parcelle est parfois utilisée (F_{mean}) et permet de comparer des parcelles avec des nombres de plantes différents. L'écart-type de la fitness (entre plantes de la parcelle) (F_{sd}) peut également être utilisé.

Module de fitness

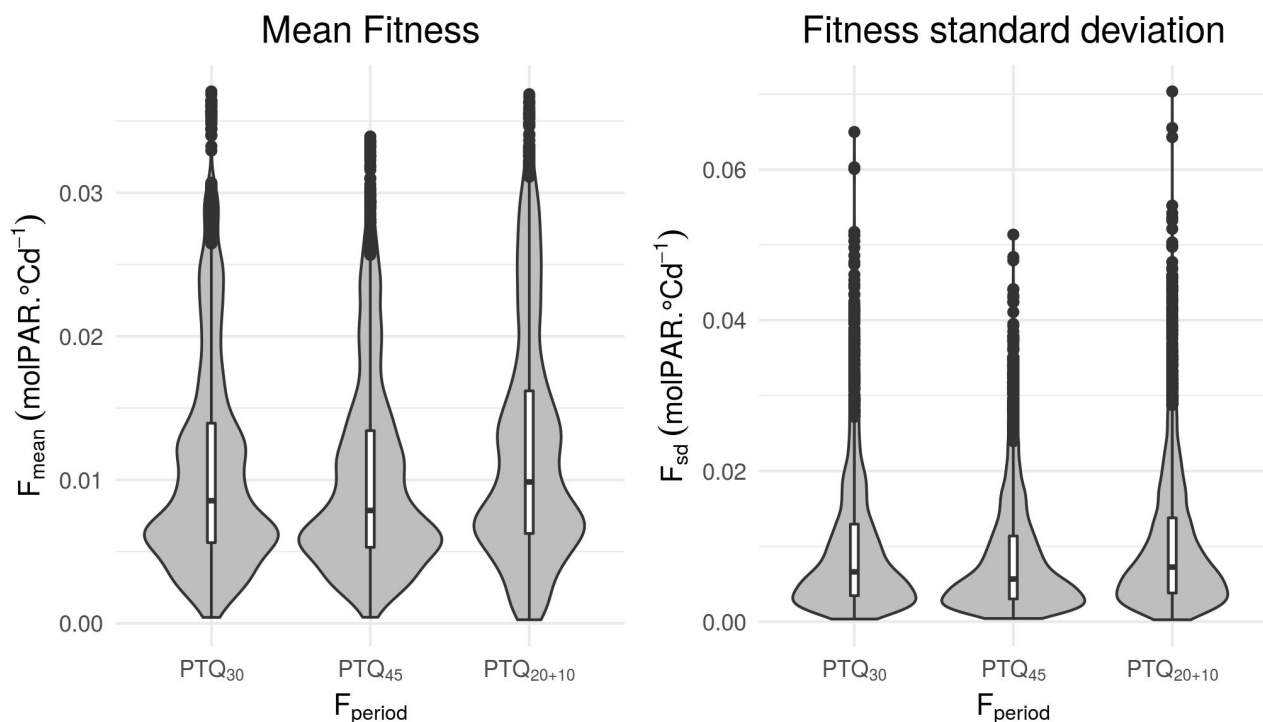


Figure 3. Distributions de la fitness moyenne (gauche) et de l'écart-type de la fitness (droite) en fonction de la période utilisée pour le calcul de la fitness.

Les tests de Kruskal-Wallis indiquent que les fitness (moyenne et écart-type) calculées à partir de PTQ₃₀, de PTQ₂₀₊₁₀ et de PTQ₄₅ sont significativement différentes les unes des autres. En effet, les F_{mean} et F_{sd} calculés avec PTQ₄₅ sont inférieurs, en moyenne, à ceux calculés à partir des autres F_{period} testés. Bien que les différences entre les F_{mean} et F_{sd} calculés avec différents F_{period} soient statistiquement significatives, la figure 3 laisse apparaître des écarts relativement faibles et des distributions qui se recourent largement. De plus, le calcul des indices de sensibilité (Annexe 1) met en avant l'impact relativement faible des variations de F_{period} en comparaison à des variations d'autres entrées du modèle. Ainsi, dans le plan de simulation décrit au Chapitre 3, des variations de F_{period} expliquent uniquement 2,1% de la variance totale de F_{mean} et 1,4% de la variance de F_{sd} . Bien

que le choix de F_{period} ait un impact notable sur l'influence des autres entrées sur F_{sd} , le calcul de la fitness moyenne de la parcelle (F_{mean}) semble peu impacté par les interactions entre F_{period} et les autres entrées du modèle. Enfin, étant donné que la fitness est calculée à l'échelle de la plante, il est intéressant de noter que le choix de F_{period} a peu d'impact sur le classement des individus d'une parcelle selon leur fitness : au sein d'une parcelle, les plantes présentant les fitness les plus élevées sont les mêmes quelle que soit la F_{period} considérée (Annexe 1 – Figure 7).

DISCUSSION

La modélisation est un outil majeur de l'étude des associations variétales et, les FSPM, en particulier, permettent d'examiner des questions essentielles sur ces peuplements hétérogènes. Le modèle WALTer (Lecarpentier *et al.*, 2019) a ainsi été développé dans l'optique d'étudier les mélanges de blé tendre. Le formalisme et les fonctionnalités de WALTer sont donc particulièrement adaptés à l'étude de ces couverts complexes. Cependant, comme tout modèle, WALTer repose sur un certain nombre d'hypothèses sous-jacentes et de simplifications qui doivent être prises en compte lors de son utilisation et de l'analyse de ses résultats.

WALTer permet de modéliser la compétition pour la lumière entre plantes d'une parcelle grâce à l'utilisation d'un modèle (CARIBU: Chelle and Andrieu, 1998) qui simule la distribution du rayonnement solaire incident dans le couvert. Si, en conditions réelles, ce rayonnement peut généralement être divisé en une part de rayonnement diffus et une part de rayonnement direct, WALTer permet simplement de simuler un rayonnement uniquement diffus ou uniquement direct. Les simulations réalisées avec le modèle impliquent donc nécessairement une hypothèse sous-jacente sur la météo (ciel couvert ou ciel clair), qui s'applique à tous les jours de la simulation et qui impacte les sorties générées. Il est cependant intéressant de noter que les sorties du modèle pour des simulations réalisées avec un rayonnement diffus sont peu différentes de celles des simulations qui utilisent du rayonnement direct (Annexe 1 - Figures 3 et 4). Ainsi, l'utilisation de rayonnement diffus seul, pour l'ensemble des simulations de cette thèse, apparaît comme une simplification acceptable.

D'autre part, la lumière est le seul facteur environnemental considéré dans WALTer : l'eau, les nutriments (azote, phosphore, potassium, soufre...) ou les maladies ne sont pas explicitement simulés par le modèle. Ces différents facteurs sont cependant connus pour impacter le développement du blé. En particulier, de nombreuses études ont mis en avant le rôle important de l'azote dans la plasticité du tallage (Alzueta *et al.*, 2012 ; Assuero et Tognetti, 2010 ; Dornbusch *et al.*, 2011a ; Sparkes *et al.*, 2006). Malgré cette simplification, WALTer permet de simuler, de façon réaliste, la dynamique du tallage pour une large gamme de densités de semis dans des conditions non limitantes en azote (Lecarpentier *et al.*, 2019). Néanmoins, la disponibilité en azote influence probablement la paramétrisation du modèle, et cet impact ne doit pas être négligé. Cette limite est en lien avec une simplification importante de WALTer : le modèle simule uniquement la partie aérienne des plantes et aucun système racinaire n'est donc considéré. Par ailleurs, le rythme de développement des plantes est basé sur le temps thermique dans WALTer, mais aucun autre impact de la température n'est considéré dans le modèle. La nouvelle fonctionnalité intégrée à WALTer, permettant de simuler la mort d'une partie des plantes de la parcelle, pourrait être utilisée pour

étudier la compensation par le tallage après un stress thermique, comme le gel. Une prise en compte plus fine des dégâts liés au gel ou à un stress dû à de hautes températures, tel que l'échaudage, demanderait cependant des développements conséquents.

Si les hypothèses intégrées au modèle permettent de simuler de façon réaliste la plasticité du tallage en réponse à des variations dans l'intensité de la compétition pour la lumière, une des limites de WALTer est que la plasticité n'est intégrée pour aucun autre trait que le tallage. En effet, la hauteur finale du brin maître et les dimensions des feuilles à maturité, par exemple, sont fixées en entrée du modèle. Bien que ces traits soient plastiques, une étude de [Lecarpentier \(2017\)](#), portant sur 21 traits d'architecture aérienne mesurés sur 20 géotypes de blé tendre, a montré que le tallage est, de loin, le trait d'architecture aérienne le plus plastique en réponse à des variations de densité de semis. Le tallage apparaît donc central pour l'étude de la compétition pour la lumière. En conséquence, le formalisme de WALTer suppose que la plasticité des autres traits d'architecture est négligeable comparativement à la plasticité du tallage.

Le nombre final de feuilles sur le brin maître fait également partie des traits fixés en entrée du modèle et pour lesquels aucun contrôle environnemental n'est considéré. Le nombre final de feuilles du brin maître étant étroitement lié à la précocité de floraison chez le blé, l'absence de contrôle environnemental sur ce trait est une limite notable pour la modélisation de la phénologie par WALTer. De plus, d'autres simplifications ont également été faites dans la modélisation de la phénologie dans WALTer. Ainsi, dans le modèle, le phyllochrone (*i.e.* l'intervalle de temps entre l'émergence de 2 feuilles successives) est un paramètre génotypique exprimé en temps thermique qui est identique pour tous les axes et fixe au cours du cycle. Ce formalisme est fréquemment utilisé pour la modélisation du blé mais ne rend pas compte des différences de phyllochrone entre le brin maître et les talles et de son évolution au cours du cycle ([Abichou et al., 2018](#) ; [Baccar, 2011](#) ; [Kirby et al., 1985](#)). De plus, le calcul du temps thermique dans WALTer peut être discuté. En particulier, une température de base de 0°C est utilisée tout au long du cycle de développement, mais des variations de la température de base en fonction du stade phénologique et du géotype considéré ont été proposées par plusieurs auteurs ([Salazar-Gutierrez et al., 2013](#) ; [Slafer et Rawson, 1995](#)). Pour toutes ces raisons, dans le cadre de l'étude du développement de mélanges variétaux, WALTer est davantage indiqué pour analyser l'impact de différences d'architecture entre géotypes que pour s'intéresser à l'influence de différences de phénologie entre variétés. De futurs développements de WALTer, notamment l'intégration d'un module de précocité de floraison permettant de déterminer le nombre final de feuilles en fonction de la séquence climatique ([He et al., 2012](#)), pourraient permettre d'examiner des questions en lien avec la phénologie.

Il existe des FSPM de blé permettant de dépasser certaines des limites listées ci-dessus. [Barillot et al. \(2019\)](#) ont ainsi couplé plusieurs modèles pour simuler des mélanges binaires avec des ports foliaires contrastées. Le couplage a permis aux auteurs d'étudier la période post-floraison en prenant en compte les interactions entre le métabolisme du carbone et de l'azote, l'architecture des plantes et de nombreux facteurs environnementaux (lumière, température, humidité, vent et concentration en azote dans le sol). [Gauthier et al. \(2020\)](#) ont étendu ce couplage afin de permettre de simuler également des stades de développement plus précoces : du semis à la montaison. Ces modèles assez complets ont cependant un coût computationnel plus élevé que WALTer, et ils ne permettent pas de simuler les différentes étapes de la dynamique de tallage, qui sont centrales dans la compétition

pour la lumière. Les simplifications intégrées dans WALTer permettent de faire des simulations relativement rapides avec un nombre relativement restreint de paramètres et en font donc un modèle de choix pour des études exploratoires sur les différences d'architecture aérienne dans des mélanges variétaux de blé.

Module de fitness

Une amélioration importante de WALTer depuis sa publication (Lecarpentier *et al.*, 2019) est l'ajout d'un module de fitness permettant de générer un proxy du nombre de grains produits par chaque plante. Le nombre de grains produits à l'échelle de la parcelle étant une des composantes majeures du rendement, cette nouvelle sortie est un indicateur de performance intéressant à l'échelle de la parcelle que ce soit pour des couverts en pur ou en mélange. De plus, l'objectif étant d'utiliser WALTer pour simuler des mélanges de variétés, ce nouveau module est essentiel pour connaître la proportion de chaque variété du mélange dans la parcelle. En effet, la proportion des différentes variétés dans le mélange détermine en partie la performance de la parcelle. La proportion initiale est connue car c'est un paramètre d'entrée de WALTer, mais cette proportion est amenée à évoluer entre le semis et la récolte, en réponse à la compétition pour la lumière entre les différentes variétés. La proportion des différents génotypes dans le mélange peut être estimée via le nombre d'axes de chaque variété dans la parcelle et, en fin de cycle, via le nombre d'épis de chaque génotype. À la récolte, le nombre de grains produits par chaque variété, qui dépend en partie du nombre d'épis de chaque variété, joue également un rôle crucial dans l'évolution des proportions au cours des générations lorsqu'un mélange est resemé d'une année sur l'autre. Le nombre d'épis produits par chaque plante est connu dans WALTer grâce à la modélisation réaliste des différentes phases du tallage. La connaissance de la fitness de chaque plante (proxy du nombre de grains par plante) permet de préciser les proportions des différentes variétés à la récolte. De plus, la fitness calculée par ce nouveau module est un indicateur du succès reproducteur de chaque individu, ce qui est un élément essentiel dans l'optique de coupler WALTer avec un modèle de génétique.

Ce nouveau module est basé sur le quotient photothermique (PTQ), qui est un indice introduit par Nix en 1976 comme une mesure simple de l'énergie lumineuse disponible pour la photosynthèse par unité de temps de développement de la plante. L'auteur suggère alors une relation linéaire positive entre le rendement en grains et les valeurs de PTQ pendant la période de remplissage des grains, bien que cette relation n'ait pas pu être testée correctement dans ces travaux, par manque de données adéquates. Plus tard, Midmore *et al.* (1984) et Fischer (1985) mettent en évidence une corrélation linéaire entre PTQ, calculé sur une période de 30 jours précédant la floraison (PTQ₃₀), et le nombre de grains produits par m² (GM²) dans des parcelles de blé de printemps. Cependant, la période précise pendant laquelle PTQ doit être calculé pour être corrélé au mieux à GM² a fait l'objet de débats. En se basant sur les travaux de Wardlaw (1970) et de Jenner (1979), qui ont montré un impact de changements dans le rayonnement et la température dans la période de 10 jours suivant la floraison sur GM², Abbate *et al.* (1995) ont ainsi décidé de calculer PTQ pour une période allant de 20 jours avant la floraison à 10 jours après (PTQ₂₀₊₁₀). Ces auteurs ont mis en évidence une forte corrélation entre PTQ₂₀₊₁₀ et GM² pour des parcelles de blé de printemps cultivées en Argentine. Ortiz-Monasterio *et al.* (1994) ont également décidé d'utiliser PTQ₂₀₊₁₀, après avoir comparé les

corrélations entre GM^2 et PTQ calculé pour différentes périodes. Plus récemment, [Làzaro et Abbate \(2012\)](#) ont trouvé de bonnes corrélations entre PTQ_{20+10} et GM^2 chez 5 variétés sud-américaines de blé de printemps. Cependant, les travaux de [Armstrong et al. \(1987\)](#) ont montré que le nombre de grains produits était fixé très tôt après la floraison, ce qui remet en question l'utilisation de PTQ_{20+10} . Après avoir comparé les corrélations de GM^2 et PTQ pour 106 intervalles différents, [Nalley et al. \(2009\)](#) ont d'ailleurs conclu que la corrélation était bien moins bonne lorsque la période considérée incluait plus de 4 jours après la floraison. [Demotes-Mainard et Jeuffroy \(2001\)](#) ont également trouvé une moins bonne corrélation entre PTQ_{20+10} et GM^2 qu'entre PTQ_{30} et GM^2 . Dans leur étude sur des parcelles de blé d'hiver, c'est avec un PTQ calculé pour une période de 45 jours avant la floraison (PTQ_{45}) que la corrélation avec GM^2 était la meilleure. C'est d'ailleurs l'utilisation de PTQ_{45} , associé à un indice de nutrition azoté, qui a donné les meilleurs résultats dans une étude de [Ratjen et al. \(2012\)](#) comparant différents modèles de prédiction du nombre de grains chez le blé d'hiver.

WALTer ayant été développé pour simuler des parcelles de blé d'hiver, le module de fitness est basé sur l'utilisation de PTQ_{45} , comme proposé par [Demotes-Mainard et Jeuffroy \(2001\)](#) et [Ratjen et al. \(2012\)](#). L'AS a permis d'estimer l'impact de l'incertitude de F_{period} sur les résultats du module de fitness. Ainsi, il est important de garder à l'esprit, lors de l'analyse des résultats, que les fitness calculées avec d'autres F_{period} sont significativement différentes de celles calculées avec PTQ_{45} . Cependant, l'impact de F_{period} sur les sorties de fitness est relativement faible par rapport à l'impact de variations d'autres entrées du modèle. De plus, malgré des interactions notables entre F_{period} et certaines autres entrées du modèle pour le calcul de F_{sd} , l'influence des différentes entrées de WALTer sur la fitness moyenne (F_{mean}) n'est que faiblement impactée par le choix de F_{period} .

Outre le choix de F_{period} , d'autres éléments du calcul de la fitness peuvent être discutés. En effet, si PTQ a été défini par [Nix \(1976\)](#) comme le ratio entre le rayonnement global incident et la température moyenne, c'est le PAR intercepté qui a été utilisé dans le module de fitness de WALTer. Plusieurs auteurs ont en effet mis en avant les relations plus fortes entre GM^2 et PTQ lorsque ce dernier est calculé à partir du rayonnement intercepté plutôt qu'à partir du rayonnement incident ([Midmore et al., 1984](#) ; [Fischer, 1985](#)). De la même manière, l'utilisation du PAR, plutôt que du rayonnement global, a été privilégiée dans plusieurs études. Ainsi, c'est le PAR incident qui a été utilisé pour les études de [Demotes-Mainard et Jeuffroy \(2001\)](#) et de [Ratjen et al. \(2012\)](#), et le PAR intercepté a été utilisé par [Abbate et al. \(1995\)](#) et [Làzaro et Abbate \(2012\)](#).

D'autre part, il est important de noter que, pour le module de fitness de WALTer, PTQ est utilisé comme proxy du nombre de grains par épi. Cependant, c'est habituellement la relation linéaire entre PTQ et le nombre de grains/m², et non entre PTQ et le nombre de grains par épi, qui est mise en avant. Néanmoins, GM^2 est déterminé en partie par le nombre d'épis/m² et par le nombre de grains par épi ([Slafer et al., 2014](#)) et l'on peut penser que les mécanismes biologiques sur lesquels repose la corrélation entre PTQ GM^2 (disponibilité en ressource pendant une période clef) doivent être applicables à l'échelle de l'axe. Les données disponibles dans l'étude de [Demotes-Mainard et Jeuffroy \(2001\)](#) permettent de tester la corrélation entre PTQ et GM^2 , mais également entre PTQ et le nombre d'épis/m² et entre PTQ et le nombre de grains par épi. En conditions non limitantes en azote, la corrélation entre PTQ_{45} et le nombre de grains par épi ($r = 0,93$; p -value = 0,0024) est presque aussi importante que celle entre PTQ_{45} et GM^2 ($r = 0,94$; p -value = 0,0019) et bien plus

importante que celle entre PTQ_{45} et le nombre d'épis/m² ($r = 0,78$; $p\text{-value} = 0,038$). L'utilisation de PTQ pour prédire le nombre de grains/épi apparaît donc cohérente. De plus, le calcul de PTQ à l'échelle de l'axe permet de se baser sur la date de floraison de chaque épi plutôt que sur la date de floraison moyenne de la parcelle. Il est ainsi possible de prendre en compte la variabilité entre plantes sur la date de floraison qui peut impacter les résultats (Nalley *et al.*, 2009).

Enfin, comme pour toutes les sorties de WALTer, seul le rayonnement et la température ont une influence sur les sorties du module de fitness. Les autres facteurs environnementaux pouvant impacter le nombre de grains produits ne sont pas considérés. En particulier, les disponibilités en eau et en azote ne sont pas prises en compte, malgré leur rôle clef dans le déterminisme du nombre de grains (Jamieson *et al.*, 1995 ; Jeuffroy et Bouchard, 1999 ; Sinclair et Jamieson, 2006). Or, la corrélation entre PTQ et GM² a été établie dans des conditions non limitantes en eau et en azote. Différentes études ont ainsi montré que la relation linéaire positive entre PTQ et GM² n'était plus valable avec un déficit en azote, même transitoire (Abbate *et al.*, 1995 ; Demotes-Mainard et Jeuffroy, 2001). Les valeurs de fitness générées par WALTer sont donc calculées sous l'hypothèse que les autres ressources que la lumière, en particulier l'eau et l'azote, ne sont pas limitantes.

Les FSPM sont des outils privilégiés pour étudier les associations variétales de blé et les peuplements hétérogènes en général. Les améliorations apportées à WALTer (Lecarpentier *et al.*, 2019) en font un modèle de choix pour examiner l'impact de différences d'architecture aérienne sur le développement, la performance et l'évolution des mélanges de blé tendre. Comme pour tout travail de modélisation, il est cependant important de prendre en considération les hypothèses sous-jacentes et les limites du modèle pour être en mesure de tirer des conclusions fiables à partir des simulations de cette thèse.

RÉFÉRENCES

- Abbate, P. E., Andrade, F. H., & Culot, J. P. (1995).** The effects of radiation and nitrogen on number of grains in wheat. *The Journal of Agricultural Science*, 124(3), 351-360.
- Abichou, M., Fournier, C., Dornbusch, T., Chambon, C., De Solan, B., Gouache, D., & Andrieu, B. (2018).** Parameterising wheat leaf and tiller dynamics for faithful reconstruction of wheat plants by structural plant models. *Field Crops Research*, 218, 213-230.
- Alinea.astk [package python]. (2019).** URL : <https://github.com/openalea-incubator/astk.git>
- Alzueta, I., Abeledo, L. G., Mignone, C. M., & Miralles, D. J. (2012).** Differences between wheat and barley in leaf and tillering coordination under contrasting nitrogen and sulfur conditions. *European Journal of Agronomy*, 41, 92-102.
- Armstrong, T. A., Soong, T. S., & Hinchee, M. A. (1987).** Culture of detached spikes and the early development of the fourth floret caryopsis in wheat. *Journal of plant physiology*, 131(3-4), 305-314.
- Assuero, S. G., & Tognetti, J. A. (2010).** Tillering regulation by endogenous and environmental factors and its agricultural management. *Am. J. Plant Sci. Biotechnol*, 4(1), 35-48.
- Baccar, R. (2011).** *Plasticité de l'architecture du blé d'hiver modulée par la densité et la date de semis et son effet sur les épidémies de Septoria tritici* (Doctoral dissertation, Paris, AgroParisTech).
- Barillot, R., Chambon, C., Fournier, C., Combes, D., Pradal, C., & Andrieu, B. (2019).** Investigation of complex canopies with a functional–structural plant model as exemplified by leaf inclination effect on the functioning of pure and mixed stands of wheat during grain filling. *Annals of botany*, 123(4), 727-742.
- Barillot, R., Escobar-Gutiérrez, A. J., Fournier, C., Huynh, P., & Combes, D. (2014).** Assessing the effects of architectural variations on light partitioning within virtual wheat–pea mixtures. *Annals of botany*, 114(4), 725-737.
- Casal, J. J. (2013).** Photoreceptor signaling networks in plant responses to shade. *Annual review of plant biology*, 64, 403-427.
- Chelle, M., & Andrieu, B. (1998).** The nested radiosity model for the distribution of light within plant canopies. *Ecological Modelling*, 111(1), 75-91.
- Demotes-Mainard, S., & Jeuffroy, M. H. (2001).** Incorporating radiation and nitrogen nutrition into a model of kernel number in wheat. *Crop Science*, 41(2), 415-423.
- Dornbusch, T., Baccar, R., Watt, J., Hillier, J., Bertheloot, J., Fournier, C., & Andrieu, B. (2011a).** Plasticity of winter wheat modulated by sowing date, plant population density and nitrogen fertilisation: dimensions and size of leaf blades, sheaths and internodes in relation to their position on a stem. *Field Crops Research*, 121(1), 116-124.
- Dornbusch, T., Watt, J., Baccar, R., Fournier, C., & Andrieu, B. (2009).** Towards a quantitative evaluation of cereal lamina shape using an empirical shape model. In *2009 Third International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications* (pp. 229-236). IEEE.

- Dornbusch, T., Watt, J., Baccar, R., Fournier, C., & Andrieu, B. (2011b).** A comparative analysis of leaf shape of wheat, barley and maize using an empirical shape model. *Annals of botany*, 107(5), 865-873.
- den Dulk, J. A. (1989).** *The interpretation of remote sensing: a feasibility study* (Doctoral dissertation, Den Dulk).
- Evers, J. B., & Bastiaans, L. (2016).** Quantifying the effect of crop spatial arrangement on weed suppression using functional-structural plant modelling. *Journal of plant research*, 129(3), 339-351.
- Evers, J. B., Van Der Werf, W., Stomph, T. J., Bastiaans, L., & Anten, N. P. (2019).** Understanding and optimizing species mixtures using functional-structural plant modelling. *Journal of experimental botany*, 70(9), 2381-2388.
- Evers, J. B., & Vos, J. (2013).** Modeling branching in cereals. *Frontiers in plant science*, 4, 399.
- Evers, J. B., Vos, J., Chelle, M., Andrieu, B., Fournier, C., and Struik, P. C. (2007).** Simulating the effects of localized red:far-red ratio on tillering in spring wheat (*Triticum aestivum*) using a three-dimensional virtual plant model. *New Phytologist*, 176(2):325–336.
- Evers, J. B., Vos, J., Yin, X., Romero, P., van der Putten, P. E. L., and Struik, P. C. (2010).** Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *Journal of Experimental Botany*, 61(8):2203–2216.
- Fischer, R. A. (1985).** Number of kernels in wheat crops and the influence of solar radiation and temperature. *The Journal of Agricultural Science*, 105(2), 447-461.
- Fournier, C., Andrieu, B., Ljutovac, S., & Saint-Jean, S. (2003).** ADEL-wheat: a 3D architectural model of wheat development. *Plant Growth Modeling and Applications* (eds B.-G. Hu & M. Jaeger), 54-66.
- Fournier, C., & Pradal, C. (2012).** A plastic, dynamic and reducible 3D geometric model for simulating gramineous leaves. In *2012 IEEE 4th International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications* (pp. 125-132). IEEE.
- Gaudio, N., Escobar-Gutiérrez, A. J., Casadebaig, P., Evers, J. B., Gerard, F., Louarn, G., ... & Barillot, R. (2019).** Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agronomy for sustainable development*, 39(2), 20.
- Gauthier, M., Barillot, R., Schneider, A., Chambon, C., Fournier, C., Pradal, C., ... & Andrieu, B. (2020).** A functional structural model of grass development based on metabolic regulations and coordination rules. *Journal of Experimental Botany*.
- Godin, C., Sinoquet, H. (2005).** Functional-structural plant modelling. *New phytologist*, 166(3), 705-708.
- He, J., Le Gouis, J., Stratonovitch, P., Allard, V., Gaju, O., Heumez, E., ... & Semenov, M. A. (2012).** Simulation of environmental and genotypic variations of final leaf number and anthesis date for wheat. *European Journal of Agronomy*, 42, 22-33.

- Hemmerling, R., Kniemeyer, O., Lanwert, D., Kurth, W., & Buck-Sorlin, G. (2008).** The rule-based language XL and the modelling environment GroIMP illustrated with simulated tree competition. *Functional plant biology*, 35(10), 739-750.
- Iooss, B., & Saltelli, A. (2017).** Introduction to Sensitivity Analysis. In *Handbook of Uncertainty Quantification* (pp. 1103–1122).
- Isbell, F. (2015).** Agroecology: agroecosystem diversification. *Nature plants*, 1(4), 1-2.
- Jamieson, P. D., Martin, R. J., & Francis, G. S. (1995).** Drought influences on grain yield of barley, wheat, and maize. *New Zealand journal of crop and horticultural science*, 23(1), 55-66.
- Jamieson, P. D., Semenov, M. A., Brooking, I. R., & Francis, G. S. (1998).** Sirius: a mechanistic model of wheat response to environmental variation. *European Journal of Agronomy*, 8(3-4), 161-179.
- Jenner, C. F. (1979).** Grain-filling in wheat plants shaded for brief periods after anthesis. *Functional Plant Biology*, 6(6), 629-641.
- Jeuffroy, M. H., & Bouchard, C. (1999).** Intensity and duration of nitrogen deficiency on wheat grain number. *Crop science*, 39(5), 1385-1393.
- Kirby, E. J. M., Appleyard, M., & Fellowes, G. (1985).** Leaf emergence and tillering in barley and wheat. *Agronomie*, 5(3), 193-200.
- Lázaro, L., & Abbate, P. E. (2012).** Cultivar effects on relationship between grain number and photothermal quotient or spike dry weight in wheat. *The Journal of Agricultural Science*, 150(4), 442-459.
- Lecarpentier, C. (2017).** Plasticité de l'architecture aérienne du blé en réponse à la compétition pour la lumière au sein de cultures pure ou d'associations variétales : caractérisation expérimentale et développement d'un modèle 3D. Thèse de doctorat, Université Paris-Saclay.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTER: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Lecarpentier, C., Flutre, T., Blanc, E., Gawinowski, M., Enjalbert, J. (in-prep.).** Tillering as a major driver of plant compensation.
- Lynch, J. P., Nielsen, K. L., Davis, R. D., & JablOKow, A. G. (1997).** SimRoot: modelling and visualization of root systems. *Plant and Soil*, 188(1), 139-151.
- Malézieux, E. (2012).** Designing cropping systems from nature. *Agronomy for sustainable development*, 32(1), 15-29.
- Midmore, D. J., Cartwright, P. M., & Fischer, R. A. (1984).** Wheat in tropical environments. II. Crop growth and grain yield. *Field Crops Research*, 8, 207-227.
- Monod, H., Naud, C., Makowski, D. (2006).** Uncertainty and sensitivity analysis for crop models. Working with dynamic crop models: Evaluation, analysis, parameterization, and applications, 4, 55-100.

- Monteith, J. L. (1984).** Consistency and convenience in the choice of units for agricultural science. *Experimental Agriculture*, 20(2), 105-117.
- Moon, P., Spencer, D. (1942).** Illumination from a non-uniform sky, Transactions of the Illumination Engineering Society 37 (1942) 707712.
- Muller, B., & Martre, P. (2019).** Plant and crop simulation models: powerful tools to link physiology, genetics, and phenomics. *Journal of Experimental Botany* 70(9), 2339–2344.
- Nalley, L. L., Barkley, A. P., & Sayre, K. (2009).** Photothermal quotient specifications to improve wheat cultivar yield component models. *Agronomy Journal*, 101(3), 556-563.
- Nix, H. A. (1976).** Climate and crop productivity in Australia. *Climate and rice. IRRI, Los Baños, The Philippines*, 495-507.
- Ortiz-monasterio, J. I., Dhillon, S., & Fischer, R. (1994).** Date of sowing effects on grain yield and yield components of irrigated spring wheat cultivars and relationships with radiation and temperature in Ludhiana, India. *Field crops research*, 37(3), 169-184.
- Pagès, L., Bécel, C., Boukcim, H., Moreau, D., Nguyen, C., & Voisin, A. S. (2014).** Calibration and evaluation of ArchiSimple, a simple model of root system architecture. *Ecological Modelling*, 290, 76-84.
- Perez, R. P., Pallas, B., Le Moguédec, G., Rey, H., Griffon, S., Caliman, J. P., ... & Dautat, J. (2016).** Integrating mixed-effect models into an architectural plant model to simulate inter-and intra-progeny variability: a case study on oil palm (*Elaeis guineensis* Jacq.). *Journal of experimental botany*, 67(15), 4507-4521.
- Postma, J. A., & Lynch, J. P. (2011).** Theoretical evidence for the functional benefit of root cortical aerenchyma in soils with low phosphorus availability. *Annals of botany*, 107(5), 829-841.
- Postma, J. A., & Lynch, J. P. (2012).** Complementarity in root architecture for nutrient uptake in ancient maize/bean and maize/bean/squash polycultures. *Annals of botany*, 110(2), 521-534.
- Pradal, C., Dufour-Kowalski, S., Boudon, F., Fournier, C., & Godin, C. (2008).** OpenAlea: a visual programming and component-based software platform for plant modelling. *Functional plant biology*, 35(10), 751-760.
- Ratjen, A. M., Böttcher, U., & Kage, H. (2012).** Improved modeling of grain number in winter wheat. *Field Crops Research*, 133, 167-175.
- Ritchie, J. R., & Otter, S. (1985).** Description and performance of CERES-Wheat: a user-oriented wheat yield model. *ARS-United States Department of Agriculture, Agricultural Research Service (USA)*.
- Salazar-Gutierrez, M. R., Johnson, J., Chaves-Cordoba, B., & Hoogenboom, G. (2013).** Relationship of base temperature to development of winter wheat. *International Journal of Plant Production*, 7(4), 741-762.
- Sattin, M., Zuin, M. C., & Sartorato, I. (1994).** Light quality beneath field-grown maize, soybean and wheat canopies—red: far red variations. *Physiologia Plantarum*, 91(2), 322-328.

- Sinclair, T. R., & Jamieson, P. D. (2006).** Grain number, wheat yield, and bottling beer: an analysis. *Field Crops Research*, 98(1), 60-67.
- Slafer, G. A. (2003).** Genetic basis of yield as viewed from a crop physiologist's perspective. *Annals of Applied Biology*, 142(2), 117-128.
- Slafer, G. A., & Rawson, H. M. (1995).** Base and optimum temperatures vary with genotype and stage of development in wheat. *Plant, Cell & Environment*, 18(6), 671-679.
- Slafer, G. A., Savin, R., & Sadras, V. O. (2014).** Coarse and fine regulation of wheat yield components in response to genotype and environment. *Field Crops Research*, 157, 71-83.
- Smith, H. (1982).** Light quality, photoperception, and plant strategy. *Annual review of plant physiology*, 33(1), 481-518.
- Sparkes, D. L., Holme, S. J., & Gaju, O. (2006).** Does light quality initiate tiller death in wheat?. *European Journal of Agronomy*, 24(3), 212-217.
- Verdenal, A., Combes, D., & Escobar-Gutiérrez, A. J. (2008).** A study of ryegrass architecture as a self-regulated system, using functional–structural plant modelling. *Functional Plant Biology*, 35(10), 911-924.
- Vidal, T., Gigot, C., de Vallavieille-Pope, C., Huber, L., & Saint-Jean, S. (2018).** Contrasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. *Annals of botany*, 121(7), 1299-1308.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- de Vries, J. (2019).** *How plants balance competitive growth and defence: an analysis of virtual plants in dynamic interactions* (Doctoral dissertation, Wageningen University).
- Wardlaw, I. F. (1970).** The early stages of grain development in wheat: response to light and temperature in a single variety. *Australian Journal of Biological Sciences*, 23(4), 765-774.
- Woods, D. C., & Lewis, S. M. (2015).** Design of experiments for screening. arXiv preprint arXiv:1510.05248.
- Xie, Q., Mayes, S., & Sparkes, D. L. (2016).** Optimizing tiller production and survival for grain yield improvement in a bread wheat× spelt mapping population. *Annals of botany*, 117(1), 51-66.

CHAPITRE 2

Impact de l'aléatoire dans le modèle WALTer

INTRODUCTION

L'architecture et le développement des plantes dépendent de leur génome, de l'environnement dans lequel elles se développent et des interactions entre ces deux éléments. Cependant, des variations inter-individuelles existent, même entre plantes du même génotype soumises aux mêmes conditions environnementales. Ainsi, au sein d'une parcelle agricole, des différences sont habituellement observées entre plantes génétiquement semblables. Ces différences sont le résultat de micro-variations environnementales locales, mais leur déterminisme est difficile à établir, et elles sont, de ce fait, souvent considérées comme des variations aléatoires.

Les modèles de culture font abstraction de ces variations individuelles et se concentrent sur l'étude d'un ensemble de plantes moyennes. A l'inverse, les modèles individus-centrés, aussi appelés modèles à base d'agents (IBM pour "individual-based models" ou ABM pour "agent-based models"), modélisent un ensemble d'entités autonomes en interaction les unes avec les autres et permettent ainsi de représenter explicitement les variations inter-individuelles (voir [Zhang et De Angelis, 2020](#) pour une revue récente des applications de ces modèles en biologie des plantes et écologie). Le fonctionnement "bottom-up" de ces modèles permet de faire émerger des propriétés à l'échelle de la population à partir des interactions des individus entre eux et avec leur environnement. Les modèles de plantes structure-fonction (FSPM) ([Vos et al., 2010](#)), comme WALTer ([Lecarpentier et al., 2019](#)) sont individus-centrés et permettent donc de représenter les différences entre plantes qui peuvent être le résultats de paramétrisations différentes (pour représenter des différences génétiques) ou la conséquence des processus biologiques modélisés. Dans le cas des modèles stochastiques, les différences entre plantes peuvent également être en partie aléatoires. Ainsi, un certain nombre de modèles simulent des variations aléatoires entre plantes, ce qui leur permet d'augmenter leur réalisme par rapport à des versions uniquement déterministes. C'est le cas notamment du modèle GreenLab qui, dans sa version GL2 ([Kang et al., 2008](#)), associe des probabilités au fonctionnement des bourgeons, ce qui permet de représenter des topologies partiellement stochastiques, comme cela a été fait chez le pin (*Pinus sylvestris* L.) ([Wang et al., 2009](#)) et le coton (*Gossypium hirsutum* L.) ([Li et al., 2010](#)) par exemple. C'est également le cas d'autres modèles comme ADELwheat ([Fournier et al., 2003](#)), qui intègre une part d'aléatoire sur la configuration du tallage et la courbure des limbes de plantes de blé (*Triticum aestivum* L.) ; L-grass ([Verdenal et al., 2008](#)), qui simule l'émission des talles chez le ray-grass (*Lolium perenne* L.) comme un processus partiellement aléatoire ; ou encore le modèle de canne à sucre (*Saccharum officinarum* L.) développé par [Vianna \(2018\)](#), qui intègre un aléa dans la position et l'orientation des plantes.

Dans le cas du modèle WALTer, de la stochasticité est intégrée sur quatre éléments ([Lecarpentier, 2017](#)). Tout d'abord, la position des plantes et l'orientation des organes dans l'espace présentent une part d'aléatoire. Des tirages dans des lois uniformes sont utilisés pour représenter l'imprécision du semis via des déviations de quelques centimètres par rapport au plan de semis "parfait" et pour représenter des variations de quelques degrés dans les angles d'insertion foliaire et dans la disposition des feuilles et des talles autour de la tige. D'autre part, la date de levée des plantes présente une part de stochasticité. Le délai entre le semis et la date d'émergence des plantes suit en effet une loi Normale dont la moyenne et l'écart-type sont des paramètres du modèle choisis pour représenter l'hétérogénéité observée expérimentalement. Afin de reproduire l'hétérogénéité observée

dans les parcelles de blé, le nombre final de feuilles sur le brin maître est également en partie aléatoire : cette caractéristique des plantes est contrôlée dans le modèle par le paramètre N_{MS}^B qui fixe le nombre moyen de feuilles par brin-maître pour un génotype donné à l'échelle de la parcelle. Pour chaque plante de la parcelle, un tirage aléatoire dans une loi uniforme permet de déterminer si le nombre final de feuilles sur le brin maître sera égal à la partie entière de N_{MS}^B ou à sa partie entière + 1 feuille. Enfin, le processus de tallage du modèle est partiellement stochastique, puisque l'émission des talles dépend d'une probabilité fixe (P_T) : la décision de débourrer pour un bourgeon axillaire dépend ainsi d'un tirage aléatoire dans une loi uniforme.

L'intégration de ces variations aléatoires permet d'améliorer le réalisme du modèle et de limiter la sensibilité des sorties à certains paramètres incertains. Des simulations préliminaires réalisées avec WALTer (non présentées ici) ont ainsi montré que le nombre final de feuilles et la date de levée avaient un poids très important dans le modèle lorsque aucun aléatoire n'était intégré sur ces éléments. De petites variations dans les valeurs de ces paramètres pouvaient alors conduire à des changements drastiques dans les sorties du modèle par des effets de seuil. Cependant, si l'intégration de variations inter-individuelles aléatoires présente des avantages, elle soulève également des questions. Tout d'abord, la question du nombre de plantes à simuler se pose. En effet, pour chaque pas de temps, WALTer génère de nombreuses sorties à l'échelle de la plante, de l'axe^a et de l'organe. Afin de résumer les informations extraites des simulations, les valeurs moyennes des sorties à l'échelle de la parcelle sont souvent utilisées dans nos analyses (cf. Chapitre 1). Ainsi, le modèle permet de suivre la dynamique de tallage (nombre d'axes en fonction du temps) de chaque plante, mais c'est principalement la dynamique de tallage moyenne de la parcelle qui est analysée. Il en va de même pour la dynamique du GAI (Green Area Index)^b et pour la dynamique de la quantité de lumière interceptée. Pour que ces moyennes soient fiables, elles doivent être calculées sur un nombre suffisant de plantes. Cependant, plus le nombre de plantes simulées est important, plus les simulations sont coûteuses en temps computationnel. Il convient donc de trouver le bon compromis entre la précision des sorties analysées et le temps de simulation.

D'autre part, il est important de déterminer si la réalisation de plusieurs réplicats par simulation est nécessaire car la part de stochasticité du modèle implique qu'une simulation réalisée avec un même jeu de paramètres ne donne pas systématiquement les mêmes sorties moyennes. En effet, même si le nombre de plantes simulées est suffisant, les sorties moyennes à l'échelle de la parcelle peuvent différer d'un réplicat à l'autre car les plantes au sein d'une parcelle ne sont pas indépendantes du fait des interactions entre individus. Il est alors essentiel de savoir si des différences significatives existent entre des réplicats d'une simulation réalisée avec un même jeu de paramètres, auquel cas des réplicats seront nécessaires. Ainsi, il n'est pas rare que des réplicats soient réalisés dans des études utilisant des FSPM stochastiques (Mao *et al.*, 2016 ; Verdenal *et al.*, 2008 ; Vianna, 2018). Cependant, la question du coût computationnel de ces réplicats se pose également, limitant le nombre de répétitions possibles.

L'objectif de ce chapitre est d'explorer ces questions afin de déterminer les conditions de simulations (nombre de plantes, utilisation de réplicats) qui devront être utilisées par la suite. En étudiant l'impact du nombre de plantes et de la graine d'aléatoire sur les sorties du modèle, cette

a Axe : le brin maître et les talles

b GAI : rapport entre la surface photosynthétique et la surface correspondante au sol

analyse permettra d'enrichir nos connaissances sur l'impact de la stochasticité dans WALTer et sur la précision des sorties étudiées.

MATÉRIEL ET MÉTHODES

Nombre de plantes

Afin de déterminer le nombre de plantes à simuler, une même simulation a été réalisée pour des nombres de plantes différents (12 ; 49 ; 56 ; 64 ; 72 et 100 plantes). Pour chacune de ces valeurs de nombre de plantes, 20 réplicats de la même simulation ont été réalisés. Pour chaque simulation, les sorties résumées suivantes (telles que définies dans le Chapitre 1) ont été extraites : le nombre moyen d'épis par plante (N_{ears}) et sa variance ($\text{Var}_{N_{\text{ears}}}$) ; le nombre maximum d'axes produit par plante en moyenne ($N_{\text{axes}}^{\text{max}}$) et sa variance ($\text{Var}_{N_{\text{axes}}^{\text{max}}}$) ; le GAI maximum de la parcelle (GAI_{max}) et la date à laquelle il est atteint ($D_{\text{GAI}_{\text{max}}}$) ; le pourcentage moyen de lumière interceptée par la parcelle (L_{perc}) et la fitness moyenne des plantes de la parcelle (F_{mean}). Pour chaque nombre de plantes et pour chaque sortie, le coefficient de variation entre les différents réplicats a été calculé. Cette analyse a été réalisée pour 3 densités de semis contrastées (100, 200 et 400 plantes par m^2) afin de mesurer l'impact de ce paramètre sur la variance du modèle et sur le nombre de plantes à simuler. D'autres entrées que la densité sont également susceptibles d'impacter la variance du modèle. C'est le cas des probabilités d'émergence des talles et de la talle du coléoptile (P_T et P_{CT}), du nombre final de feuille (N_{MS}^B) et des paramètres GAI_c et PAR_t , qui contrôlent tous partiellement le nombre de talles qui peuvent être émises et conservées par les plantes de la parcelle.

Afin d'estimer l'impact de ces différents paramètres sur la variance de WALTer, 20 réplicats de la même simulation ont également été réalisés à une densité de 200 plantes/ m^2 avec $N_{MS}^B = 8,3^c$ et 15,3, avec $P_T = 0,5$ et 1, avec $P_{CT} = 0,5$, avec $\text{GAI}_c = 0,25$ et 1,25 et avec $\text{PAR}_t = 200000$ et 400000 $\mu\text{mol}.\text{cm}^{-2}.\text{degrés-jours}^{-1}$ (Tableau 1).

c La partie décimale du nombre final de feuilles correspond à la proportion de plantes présentant une feuille supplémentaire. Le reste des plantes présentant un nombre final de feuilles égal à la partie entière de N_{MS}^B .

Tableau 1. Récapitulatif des simulations réalisées pour estimer le nombre de plantes à simuler et la variance du modèle. La première ligne du tableau correspond à la simulation de référence.

Densité	N_{MS}^B	P_T	P_{CT}	GAI_c	PAR_t	Nombre de plantes	Nb simulations
200	11,3	0,88	0	0,75	300000	12, 49, 56, 64, 72, 100 (20 rep/ nb plantes)	120
100	11,3	0,88	0	0,75	300000		120
400	11,3	0,88	0	0,75	300000		120
200	8,3	0,88	0	0,75	300000		120
200	15,3	0,88	0	0,75	300000		120
200	11,3	0,5	0	0,75	300000		120
200	11,3	1	0	0,75	300000		120
200	11,3	0,88	0,5	0,75	300000		120
200	11,3	0,88	0	0,25	300000		120
200	11,3	0,88	0	1,25	300000		120
200	11,3	0,88	0	0,75	200000		120
200	11,3	0,88	0	0,75	400000		120
Total							1440

Réplicats

Une fois le nombre de plantes à simuler identifié, il est important de s'assurer que les sorties analysées ne sont pas significativement différentes d'un réplicat à l'autre. Si des différences significatives sont détectées d'une simulation à l'autre pour un même jeu de paramètres, les sorties devront être moyennées sur plusieurs réplicats pour être fiables.

Afin de tester si des différences significatives existent entre réplicats d'une même simulation, les simulations effectuées à l'étape précédente ont été utilisées (20 réplicats de la même simulation pour 3 densités : 100, 200 et 400 plantes par m²). Pour chaque réplicat, les sorties ont été calculées plante à plante. Un test de Kruskal-Wallis a ensuite été réalisé pour chaque sortie et pour chaque densité. L'hypothèse nulle (H0) de ce test est qu'il n'existe pas de différence significative entre les sorties calculées sur les différents réplicats. D'après H0, il n'y aurait donc pas plus de différences entre 2 plantes issues de réplicats différents qu'entre 2 plantes issues d'un même réplicat.

Les simulations ont été réalisées avec un semis en maillage (distance inter-rangs = distances entre plantes dans le rang) et en utilisant les données de rayonnement disponibles dans l'étude de [Darwinkel \(1978\)](#) réalisée à Lelystad et une séquence de température correspondant aux moyennes journalières calculées entre 2004 et 2014 pour la même localité. Les valeurs par défaut des entrées du modèle (Annexe 1) ont été utilisées (exception faite des entrées listées dans le Tableau 1). Les analyses ont été réalisées avec le logiciel R ([R Core Team, 2018](#)).

RÉSULTATS

Nombre de plantes

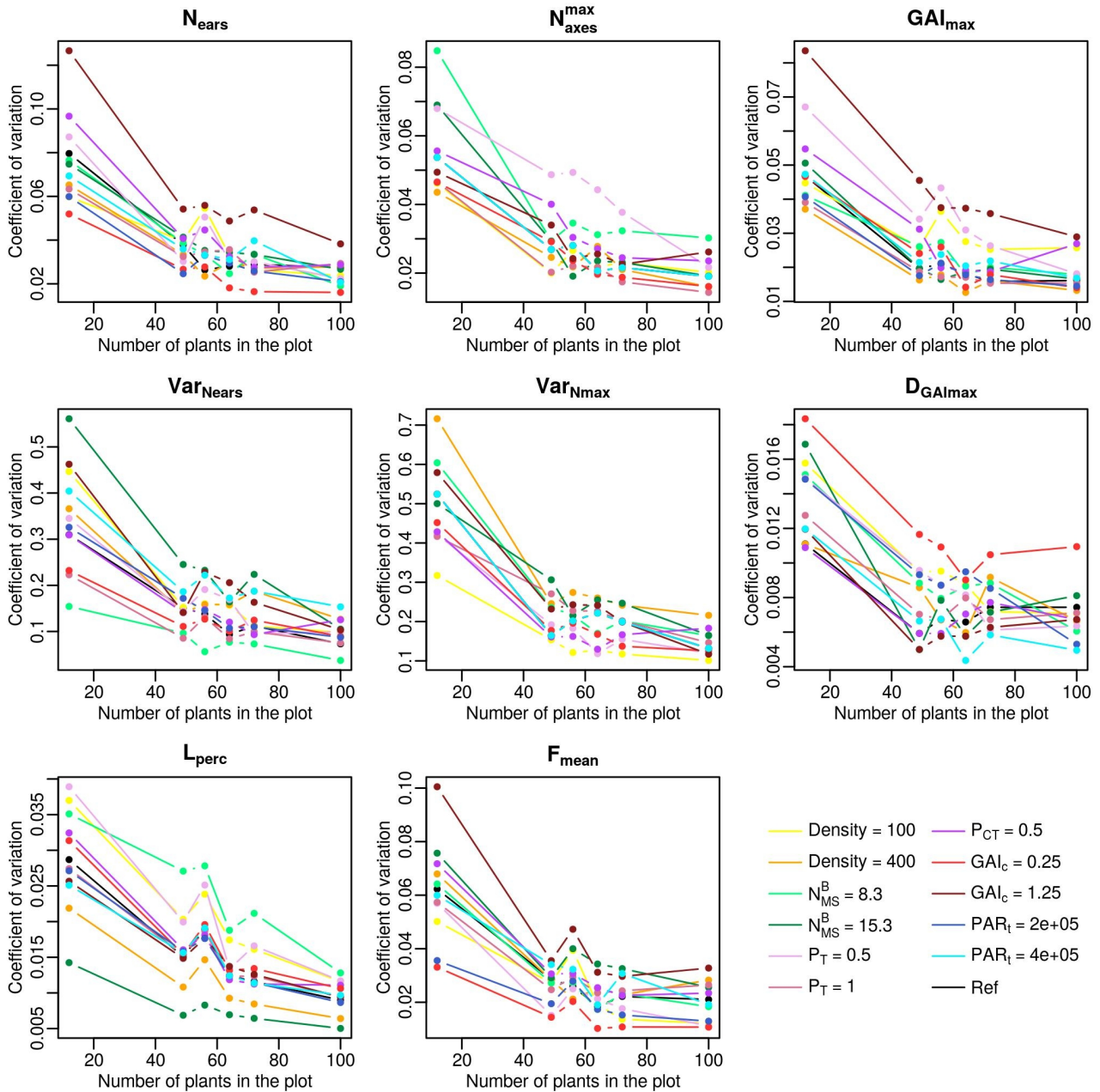


Figure 1. Coefficients de variation entre les sorties de 20 réplicats d'une même simulation en fonction du nombre de plantes simulées. La courbe noire correspond au paramétrage de référence (première ligne du Tableau 1) et les courbes de couleur correspondent à différentes valeurs d'entrées.

Conformément à ce qui était attendu, la figure 1 montre une diminution générale du coefficient de variation entre les sorties des différents réplicats (CV) avec l'augmentation du nombre de plantes simulées. La diminution de CV est de l'ordre de $\frac{1}{\sqrt{n}}$ (avec n le nombre de plantes), comme cela est attendu théoriquement. Pour des simulations de 56 plantes, les coefficients de variation entre les

sorties des différents réplicats sont relativement faibles. En effet, pour ce nombre de plantes, le CV moyen est compris entre 2% et 4% pour la plupart des sorties (N_{ears} , $N_{\text{axes}}^{\text{max}}$, GAI_{max} , L_{perc} et F_{mean}). Le CV de $D_{GAI_{\text{max}}}$ est même particulièrement faible, avec une moyenne de 0,8% pour des simulations de 56 plantes. En revanche, on remarque que les CV de $\text{Var}_{N_{\text{ears}}}$ et $\text{Var}_{N_{\text{max}}}$ sont bien supérieurs à ceux des autres sorties. Ainsi, pour des simulations de 56 plantes, le CV moyen de $\text{Var}_{N_{\text{ears}}}$ est de 16% et celui de $\text{Var}_{N_{\text{max}}}$ est de 20%.

On note, en outre, que les CV diffèrent en fonction du jeu de paramètres utilisé pour les simulations. Cependant, il ne semble pas y avoir de configurations de paramètres pour lesquelles les CV de toutes les sorties sont systématiquement supérieurs ou inférieurs aux autres. Pour les simulations de 56 plantes, les CV des différentes configurations de paramètres présentent toujours un écart inférieur à un facteur 3 pour toutes les sorties sauf $\text{Var}_{N_{\text{ears}}}$ et L_{perc} . Pour ces deux sorties, l'écart maximum entre les CV des différentes configurations de paramètres est un facteur 4,1 et 3,4 respectivement.

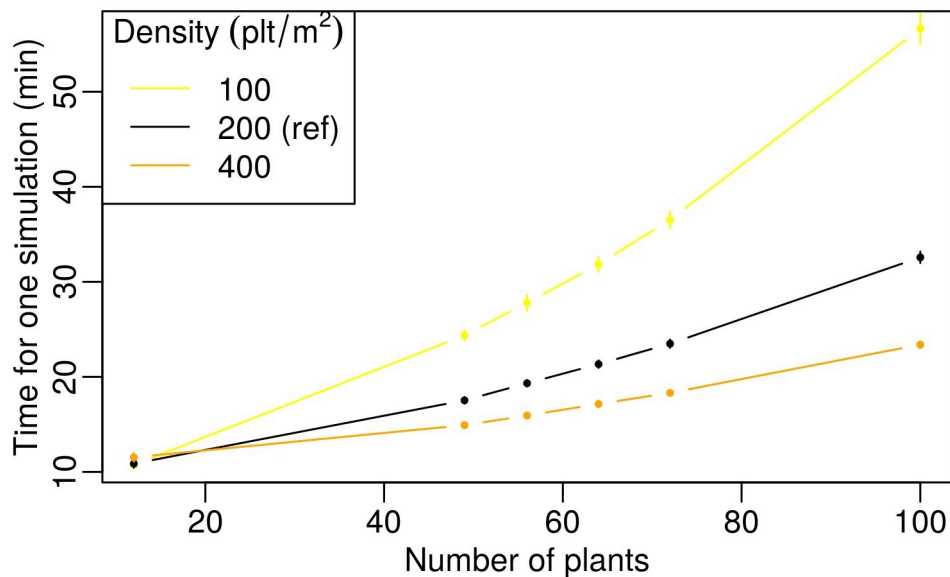


Figure 2. Temps moyen d'une simulation (en minutes) \pm son écart-type en fonction du nombre de plantes simulées et pour 3 densités de semis.

La figure 2 montre que, conformément à ce qui était attendu, le temps de simulation augmente avec le nombre de plantes simulées. Les simulations durent entre environ 10 minutes pour 12 plantes et environ 1h pour 100 plantes à une densité de 100 plantes par m^2 .

A l'exception des simulations de 12 plantes, pour lesquelles il y a peu de différences de temps de simulation entre les 3 densités, on note que, pour un nombre de plantes donné, le temps de simulation est d'autant plus important que la densité est faible. Par exemple, la simulation de 100 plantes dure en moyenne environ 57 minutes, 33 minutes et 23 minutes pour des densités de 100, 200 et 400 plantes par m^2 respectivement.

Réplicats

Tableau 2. P-values obtenues pour les tests de Kruskal-Wallis entre les 20 répliquats d'une même simulation de 56 plantes pour différentes valeurs d'entrée et pour chaque sortie. La présence d'une astérisque indique une p-value inférieure à 0.05.

Entrée \ Sortie	N_{ears}	$N_{\text{axes}}^{\text{max}}$	GAI_{max}	$D_{GAI_{\text{max}}}$	L_{perc}	F_{mean}
Ref	1,00	0,920	1,00	0,119	1,00	1,00
Densité = 100	0,890	0,964	0,962	$2,17 \times 10^{-07} *$	1,00	0,999
Densité = 400	1,00	0,808	0,981	0,0482 *	0,998	1,00
$N_{MS}^B = 8,3$	0,996	0,775	0,873	0,390	0,996	1,00
$N_{MS}^B = 15,3$	0,991	0,996	0,999	0,386	0,999	0,994
$P_T = 0,5$	0,982	0,623	0,945	0,0383 *	1,00	1,00
$P_T = 1,0$	1,00	0,797	1,00	0,618	1,00	1,00
$P_{CT} = 0,5$	0,999	0,976	1,00	0,359	1,00	1,00
$GAI_c = 0,25$	0,993	0,991	0,990	$1,48 \times 10^{-03} *$	1,00	1,00
$GAI_c = 1,25$	0,927	0,983	0,975	0,0150 *	0,853	0,841
$PAR_t = 200000$	1,00	0,920	0,998	$2,28 \times 10^{-04} *$	1,00	1,00
$PAR_t = 400000$	0,997	0,920	0,992	0,0314 *	0,990	0,996

Exception faite de $D_{GAI_{\text{max}}}$, les test de Kruskal-Wallis donnent systématiquement des p-value supérieure au seuil de 0,05 pour toutes les sorties (Tableau 2), ce qui indique qu'il n'y a pas de différences significatives entre les sorties des différents répliquats. En revanche, pour certaines configurations de paramètres, le test de Kruskal-Wallis donne des p-value inférieures au seuil de 0,05 pour la sortie $D_{GAI_{\text{max}}}$ (date à laquelle le GAI maximum est atteint). En effet, si la simulation de référence donne une p-value de 0,119, des variations de densité de semis, des changement de valeur pour les paramètres GAI_c , PAR_t ou encore une probabilité d'émergence des talles (P_T) trop faible conduisent tous à une p-value inférieure à 0,05. Pour ces configurations, il existe donc une différence significative entre les valeurs de $D_{GAI_{\text{max}}}$ pour différents répliquats d'une même simulation.

DISCUSSION

WALTer est un FSPM stochastique : le modèle intègre une part d'aléatoire dans son fonctionnement afin d'améliorer son réalisme et de limiter des effets de seuil indésirables. Dans ce chapitre, l'étude de l'impact de l'aléatoire du modèle sur ses sorties a permis de déterminer les conditions de simulation pour la suite de la thèse.

Bien que le temps de simulation de WALTer soit relativement restreint par rapport à celui d'autres FSPM, le coût computationnel de l'utilisation du modèle est non négligeable. Les conditions de simulation ont un impact important sur le temps des simulations. En particulier, les simulations sont d'autant plus longues que le nombre de plantes simulées est élevé. De plus, certaines configurations de paramètres, comme une faible densité de semis par exemple, conduisent à la représentation d'un couvert complexe dans lequel beaucoup de talles sont produites, ce qui augmente encore le temps de simulation. Ces éléments soulignent l'importance de choisir le nombre minimal de plantes pour lequel les sorties du modèle sont fiables, afin de limiter le coût computationnel, d'autant que

certaines méthodes employées par la suite (calage et optimisation notamment) impliqueront un nombre d'appels important à WALTer. D'après les résultats de ce chapitre, le nombre de plantes à simuler pour des parcelles en pur (1 seul génotype) est ainsi fixé à 56. Pour la simulation de mélanges binaires équilibrés (2 génotypes en proportions 50/50), des parcelles de 110 plantes seront utilisées. Les coefficients de variation entre les sorties des différents réplicats sont globalement faibles pour des parcelles de 56 plantes en pur. On note cependant que ces coefficients de variations sont bien plus élevés pour $\text{Var}_{\text{NearS}}$ et Var_{Nmax} que pour les autres sorties. Ce résultat s'explique, au moins en partie, par le fait que la variance d'une sortie est partiellement dépendante de la valeur moyenne de cette sortie. Ainsi, des différences de N_{ears} entre réplicats peuvent entraîner des différences de $\text{Var}_{\text{NearS}}$ entre réplicats, et il en va de même pour $N_{\text{axes}}^{\text{max}}$ et Var_{Nmax} . C'est pourquoi les coefficients de variations mesurés entre réplicats pour $\text{Var}_{\text{NearS}}$ reflètent en partie la variation inter-réplicats de N_{ears} . Pour s'affranchir de cette confusion d'effet, il est possible d'étudier le coefficient de variation intra-parcelle de N_{ears} à la place de $\text{Var}_{\text{NearS}}$ et le coefficient de variation intra-parcelle de $N_{\text{axes}}^{\text{max}}$ à la place de Var_{Nmax} . Les coefficients de variation entre réplicats pour ces deux nouvelles sorties sont bien moins importants que ceux pour $\text{Var}_{\text{NearS}}$ et Var_{Nmax} . La variabilité intra-parcelle semble donc moins impactée par la graine d'aléatoire que ce que laisse penser la figure 1.

Le test de Kruskal-Wallis présenté ci-dessus indique qu'il n'y a pas de différences significatives entre les sorties calculées pour des réplicats différents dans des parcelles de 56 plantes en pur. Il sera ainsi possible de s'abstenir de réaliser des réplicats pour la suite des analyses de cette thèse. On note cependant une exception pour la sortie $D_{\text{GAI}_{\text{max}}}$, pour laquelle des différences existent entre réplicats dans certaines configurations de paramètres. Ces différences pourraient venir du fait que, comme toutes les sorties qui concernent la séquence temporelle du développement, $D_{\text{GAI}_{\text{max}}}$ est très sensible à l'aléatoire dans la simulation (date de levée, nombre final de feuilles et émission des talles). Cette sortie n'est pas critique pour la suite des expériences, mais il conviendra tout de même d'être prudent dans son analyse.

L'estimation de la variance des sorties du modèle est intéressante pour améliorer notre connaissance de WALTer et de son fonctionnement. Cette variance peut varier en fonction des valeurs de certains paramètres d'entrée. C'est pourquoi, pour déterminer si la variance est homogène sur l'intégralité de l'espace des paramètres, il est intéressant de la mesurer pour différentes configurations de paramètres. Les simulations de ce chapitre ont mis en avant des différences de variance pour des simulations réalisées avec des valeurs différentes de N_{MS}^{B} , P_{T} , P_{CT} , GAI_{c} , PAR_{t} ou de densité de semis. Cependant, ces différences de variance sont relativement faibles, et aucune configuration de paramètre n'a été identifiée comme entraînant des variances systématiquement plus élevées ou plus faibles que les autres.

On notera néanmoins que l'analyse présentée ici correspond à une exploration assez restreinte de l'espace des entrées, puisque la variance des sorties n'a été mesurée que pour quelques valeurs de paramètres. Pour approfondir ces résultats, il serait intéressant de connaître la variance de chaque sortie dans l'ensemble de l'espace des paramètres d'entrée. Il est possible d'obtenir une approximation relativement rapide de cette information en faisant appel à un métamodèle (Baker *et al.*, 2020). Le métamodèle se construit à partir d'un ensemble de simulations de WALTer qui suivent un plan d'expérience présentant de bonnes propriétés d'exploration de l'espace des paramètres d'entrée, tel qu'un hypercube latin maximin par exemple (McKay *et al.*, 1979 ; Johnson *et al.*,

1990). Pour chaque point du plan d'expérience, plusieurs réplicats de WALTer doivent être simulés, ce qui permet de connaître la variance des différentes sorties pour chacun de ces points. À partir de ces simulations, la construction d'un métamodèle permettrait alors d'extrapoler la variance des sorties en n'importe quel point de l'hyper-espace des paramètres. Ces informations supplémentaires peuvent être utiles pour des étapes de calibration ou d'optimisation. Cependant, malgré l'utilisation d'un métamodèle, l'estimation des variances pour tout l'espace des paramètres impliquerait un nombre relativement élevé de simulations.

Enfin, un certain nombre d'entrées contrôlant le niveau d'aléatoire intra-parcelle (délai entre le semis et l'émergence ou aléatoire dans la position des plantes et l'orientation des organes par exemple) n'ont pas été étudiées dans ce chapitre. Malgré l'impact évident de ces entrées sur l'aléatoire des simulations, il n'a pas été jugé nécessaire de les faire varier car leur valeur ont été fixées par Lecarpentier (2017) pour reproduire des niveaux de variabilité intra-parcelle observés expérimentalement et aucune variation de ces entrées n'est prévue dans la suite des analyses.

L'utilisation de modèles stochastiques soulève de nombreuses questions et plusieurs analyses pourraient être réalisées pour étudier spécifiquement la stochasticité de WALTer. Les analyses réalisées dans ce chapitre permettent d'approfondir la compréhension de l'impact de l'aléatoire dans le modèle et de fixer les conditions de simulation pour le reste de la thèse. Les choix suivants sont ainsi faits : (i) les simulations sont faites sous l'hypothèse que la variance des sorties est homogène dans l'espace des paramètres, (ii) les parcelles simulées comportent 56 plantes en pur ou 110 plantes pour les mélanges binaires équilibrés, et (iii) un seul réplicat peut suffire pour chaque simulation.

RÉFÉRENCES

- Baker, E., Barbillon, P., Fadikar, A., Gramacy, R. B., Herbei, R., Higdon, D., ... & Sacks, J. (2020).** Stochastic Simulators: An Overview with Opportunities. *arXiv preprint arXiv:2002.01321*.
- Darwinkel, A. (1978).** Patterns of tillering and grains production of winter wheat at a wide range of plant densities. *Netherland Journal of Agricultural Science*, 26:383–398.
- Fournier, C., Andrieu, B., Ljutovac, S., & Saint-Jean, S. (2003).** ADEL-wheat: a 3D architectural model of wheat development. *Plant Growth Modeling and Applications* (eds B.-G. Hu & M. Jaeger), 54-66.
- Kang, M. Z., Cournède, P. H., De Reffye, P., Auclair, D., & Hu, B. G. (2008).** Analytical study of a stochastic plant growth model: Application to the GreenLab model. *Mathematics and Computers in Simulation*, 78(1), 57-75.
- Lecarpentier, C. (2017).** Plasticité de l'architecture aérienne du blé en réponse à la compétition pour la lumière au sein de cultures pure ou d'associations variétales : caractérisation expérimentale et développement d'un modèle 3D. Thèse de doctorat, Université Paris-Saclay.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTER: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Li, D., Letort, V., Kang, M. Z., Wang, F., Guo, Y., Li, B., ... & Zhan, Z. (2010).** A stochastic, functional-structural model for cotton plants. University of California.
- Mao, L., Zhang, L., Evers, J. B., Henke, M., van der Werf, W., Liu, S., ... & Li, Z. (2016).** Identification of plant configurations maximizing radiation capture in relay strip cotton using a functional–structural plant model. *Field crops research*, 187, 1-11.
- R Core Team (2018).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Verdenal, A., Combes, D., & Escobar-Gutiérrez, A. J. (2008).** A study of ryegrass architecture as a self-regulated system, using functional–structural plant modelling. *Functional Plant Biology*, 35(10), 911-924.
- Vianna, M. D. S. (2018).** *Functional, structural and agrohydrological sugarcane crop modelling: towards a simulation platform for Brazilian farming systems*. Doctoral dissertation, Universidade de São Paulo.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- Wang, F., Kang, M., Lu, Q., Han, H., Letort, V., Guo, Y., ... & Li, B. (2009, November).** Calibration of topological development in the procedure of parametric identification: application of the stochastic GreenLab model for *Pinus sylvestris* var. *mongolica*. In *2009 Third International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications* (pp. 26-33). IEEE.

Zhang, B., & DeAngelis, D. L. (2020). An Overview of Agent-Based Models in Plant Biology and Ecology. *Annals of Botany*.

CHAPITRE 3

Analyse de sensibilité de WALTer

INTRODUCTION

En agronomie, comme dans de nombreux domaines, la modélisation est largement utilisée lorsque les expérimentations ne sont pas possibles (coût trop élevé, nombre important de combinaisons etc...), pour mieux comprendre certains processus complexes, pour contrôler strictement les conditions testées, ou pour faire de la prédiction par exemple. Les modèles génèrent des sorties qui dépendent d'un nombre plus ou moins important de paramètres et de variables. Ces paramètres et variables d'entrées sont tous entachés d'une certaine variabilité ou incertitude (Monod *et al.*, 2006). Les paramètres sont des éléments du modèle qui ne peuvent pas être directement mesurés et sont donc fixés à partir de la bibliographie, ou estimés à partir de données mesurées. Ainsi, les valeurs des paramètres présentent une certaine incertitude, liée au processus d'estimation et aux erreurs d'observation des données utilisées. Les variables d'entrée d'un modèle sont des valeurs qui sont mesurées et portent également une certaine incertitude car elles présentent des imprécisions de mesure ou ne sont pas directement observables et sont donc estimées. Il est alors impératif d'étudier les répercussions éventuelles des incertitudes des entrées du modèle sur les sorties qu'il produit, afin de s'assurer de sa bonne utilisation.

L'analyse de sensibilité (AS) (Iooss et Saltelli, 2017), qui évalue la réponse des sorties du modèle suite à des variations effectuées sur les entrées (variables ou paramètres), est alors une étape essentielle du processus de modélisation. Elle peut répondre à des objectifs variés (Saltelli *et al.*, 2000), et peut notamment permettre :

- d'améliorer la compréhension du fonctionnement du modèle ;
- de vérifier que la réponse des sorties face aux variations des entrées est en accord avec la réalité du système modélisé ;
- de détecter et de quantifier des interactions entre entrées (variables ou paramètres) ;
- d'identifier les entrées ayant le plus et le moins d'impact sur les sorties ;
- de simplifier le modèle, notamment en fixant les entrées ayant un faible impact sur les sorties ;
- d'identifier les variables qui, parce-qu'elles ont une forte influence sur les sorties du modèle, devront être mesurées avec précision ;
- de faciliter l'étape d'estimation des paramètres en se concentrant sur ceux qui auront été identifiés comme très influents sur les sorties du modèle.

Afin de répondre à ces objectifs variés, plusieurs méthodes d'analyses de sensibilité existent et peuvent être classifiées de différentes manières (Cariboni *et al.*, 2007 ; Wu, 2012). En se basant sur les objectifs de l'analyse, il est par exemple possible de distinguer : d'une part, les analyses qualitatives, qui permettent de classer les entrées en différentes catégories en fonction de leurs effets sur les sorties (comme la méthode de Morris (Morris, 1991)); et d'autre part, les analyses quantitatives, qui permettent de mesurer la contribution de chaque entrée à la variance totale des sorties (comme la méthode de Sobol (Sobol, 1993)). Il est également possible de classer les méthodes d'analyses de sensibilité en séparant celles basées sur l'utilisation d'un échantillonnage ("sampling-based methods" ; Helton *et al.*, 2006) de celles qui utilisent un métamodèle (Le Gratiet *et al.*, 2017). Enfin, une des classifications souvent utilisée fait la distinction entre les méthodes

locales (Turányi, 1990), qui s'intéressent à de petites variations autour des valeurs nominales des entrées ; et les méthodes globales, qui étudient l'impact de variations des entrées dans l'ensemble de leur gamme de variation (Iooss et Lemaître, 2015). Le choix de la méthode à utiliser pour réaliser une analyse de sensibilité dépend de l'objectif visé, des hypothèses sur la complexité et la régularité du modèle, du nombre d'entrées considérées et du coût computationnel d'une simulation (voir Cariboni *et al.*, 2007 et Iooss et Saltelli, 2017 pour des schémas et arbres de décisions sur la sélection de méthodes d'AS).

Les modèles de plantes structure-fonction (FSPM) (Vos *et al.*, 2010 ; Godin & Sinoquet, 2005), dont WALTer (Lecarpentier *et al.*, 2019) fait partie, présentent un certain nombre de caractéristiques qui doivent être prises en compte lors de l'analyse de sensibilité de ce type de modèle (voir Wu, 2012 pour une étude approfondie de l'AS des FSPM). En effet, la plupart du temps, les FSPM décrivent de nombreux processus en interaction, génèrent des sorties dynamiques, comptent un grand nombre de paramètres et ont un coût de calcul élevé. Des méthodes d'AS locales ont été utilisées sur des FSPM (voir Evers *et al.*, 2010 pour un exemple sur le blé (*Triticum aestivum* L.)) ainsi que des méthodes plus globales, comme dans l'étude de Drouet et Pagès (2007) sur un modèle simulant l'architecture aérienne et racinaire de plantes de maïs (*Zea mays* L.), dans l'étude de Kang *et al.* (2014) dans laquelle des indices de Sobol (Sobol, 1993) ont été calculés sur le modèle SUNLAB de tournesol (*Helianthus annuus* L.) ou encore dans le cas du riz (*Oryza sativa* L.) avec la méthode FAST (Saltelli *et al.*, 1999) employée sur le modèle Ecomeristem par Luquet *et al.* en 2012.

Le modèle WALTer a également fait l'objet d'une AS (Lecarpentier *et al.*, 2019) par une méthode de criblage (Iooss et Lemaître, 2015). Cependant, étant donné les modifications importantes apportées au modèle (cf. Chapitre 1), il apparaît judicieux de réaliser une nouvelle analyse, afin d'améliorer la compréhension du fonctionnement de cette nouvelle version de WALTer. De plus, le temps de simulation ayant été réduit dans la nouvelle version du modèle, cette nouvelle AS pourra inclure un plus grand nombre de facteurs d'entrées. Enfin, l'AS nous servira de base pour identifier les paramètres sur lesquels se concentrer lors de l'étape de calibration.

MATÉRIEL ET MÉTHODES

Toutes les analyses ont été effectuées à l'aide du logiciel R (R Core Team, 2018).

Méthode d'analyse de sensibilité

Une analyse de sensibilité de type criblage a été effectuée sur le FSPM WALTer (Lecarpentier *et al.*, 2019). Cette méthode globale permet d'identifier qualitativement, parmi un nombre relativement important d'entrées, lesquelles ont le plus et le moins d'influence sur les sorties du modèle. La sensibilité de WALTer à 13 entrées (récapitulées dans le tableau 1) a été testée.

Tableau 1. Facteurs d'entrée de l'analyse de sensibilité, description, valeurs sélectionnées pour le plan factoriel fractionnaire et unités.

Facteur d'entrée	Description	Valeurs			Unité
		100	200	400	
Density	Densité de semis	100	200	400	Plante/m ²
GAI _c	Valeur de GAI ^a des plantes du voisinage d'un individu au-dessus de laquelle l'individu concerné arrête sa phase d'émission de talles	0,25	0,75	1,25	-
d _{GAIp}	Distance de prise en compte du GAI de proximité	0,20	0,5	1	m
PAR _t	Quantité seuil de PAR ^b intercepté par un axe, en dessous de laquelle l'axe meurt	2x10 ⁵	3x10 ⁵	4x10 ⁵	μmol.cm ⁻² . degrés-jours ⁻¹
Δ _{prot}	Délai minimum entre la mort de 2 talles successives sur une même plante	10	50	100	Degrés-jours
t _{beg} ^{reg}	Âge physiologique de la plante à partir duquel la régression des talles devient possible	3,2	4,2	5,2	- ^c
N _{MS} ^B	Nombre final de feuilles sur le brin maître	8,3 ^d	11,3	15,3	-
L _{max} ^B	Longueur finale du plus long limbe du brin maître	8	16,6	35	cm
Leaf type	Port des feuilles : combinaison du paramètre d'angle d'insertion foliaire et de la courbure du limbe	étalé	moyen	dressé	-
P _T	Probabilité d'émergence d'une talle (sauf talle du coléoptile)	0,50	0,75	0,95	-
P _{CT}	Probabilité d'émergence de la talle du coléoptile	0	0,35	0,7	-
Height	Hauteur finale du brin maître	40	100	160	cm
Phl	Phyllochrone : intervalle de temps thermique entre l'émergence de 2 feuilles successives	80	100	125	Degrés-jours

a Green Area Index (GAI) : rapport entre la surface photosynthétique et la surface correspondante au sol

b Photosynthetically Active Radiation (PAR) : part du rayonnement global pouvant être utilisé pour la photosynthèse

c Le paramètre t_{beg}^{reg} est exprimé en nombre de feuilles non ligulées, *i.e.* en nombre de feuilles dont l'élongation n'est pas terminée, sur le brin maître. Une valeur de 4.2 indique ainsi que 4 feuilles doivent encore finir leur croissance, ainsi qu'une 5^{ème}, qui doit encore effectuer 20% de son élongation.

d La partie décimale du nombre final de feuilles correspond à la proportion de plantes présentant une feuille supplémentaire. Le reste des plantes présentant un nombre final de feuilles égal à la partie entière de N_{MS}^B . Ce formalisme permet de reproduire l'hétérogénéité observée dans les cultures.

Un plan factoriel fractionnaire (Woods et Lewis, 2015 ; Monod *et al.*, 2006) a été sélectionné. Ce type de plan optimise la combinaison des valeurs explorées pour réduire le nombre de simulations nécessaires, tout en permettant d'accéder à une partie des interactions entre entrées. Dans le cas présent, 3 valeurs différentes ont été sélectionnées pour chaque facteur d'entrée, et un plan de résolution V comportant 729 simulations a été choisi. Le plan permet d'estimer les effets principaux et les interactions d'ordre 2 (interactions entre 2 facteurs) sans confusion d'effet, en supposant que les interactions d'ordre supérieur à 2 sont négligeables. Le plan de simulation a été généré à l'aide du package planor (Kobilinsky *et al.*, 2017). L'analyse a été réalisée avec des simulations de 56 plantes réparties dans un plan de semis en maillage (distance inter-rang = distance entre plantes dans le rang). La séquence climatique utilisée est une séquence moyenne obtenue à partir de données collectées sur 10 ans dans cinq localités de France, tel que décrit dans l'Annexe 2. À l'exception des facteurs d'entrée de l'analyse (Tableau 1), les simulations ont été réalisées avec les valeurs par défaut des entrées du modèle (Annexe 1). Afin de compléter les informations du Chapitre 2 sur l'impact de l'aléatoire dans le modèle, 5 réplicats ont été réalisés pour chaque simulation (3645 simulations au total). La graine d'aléatoire a ainsi pu être considérée comme un facteur d'entrée supplémentaire.

Estimation des gammes de variation pour les facteurs d'entrée de l'analyse

Les 3 valeurs prises par chaque facteur d'entrée dans le plan factoriel fractionnaire ont été sélectionnées en se basant sur des connaissances empiriques, des données expérimentales et des références bibliographiques. En particulier, les résultats des thèses de Mariem Abichou (Abichou, 2016), de Rim Baccar (Baccar, 2011) et de Christophe Lecarpentier (Lecarpentier, 2017) ont servi à définir ces valeurs, ainsi que des données non publiées provenant d'expérimentations réalisées à l'INRA Thiverval-Grignon.

Si, pour beaucoup, les gammes de variation des entrées de WALTer ont pu être estimées grâce à la bibliographie ou à des données expérimentales, certains paramètres d'entrée sont toutefois peu documentés et difficiles à estimer expérimentalement. C'est le cas des paramètres GAI_c , d_{GAIp} , PAR_t et Δ_{prot} . Or, ces 4 paramètres ont un rôle central dans la régulation du tallage dans le modèle. Il a donc été nécessaire de comparer des résultats de simulations à des données expérimentales pour estimer leur gamme de variation : cette comparaison permet en effet d'identifier quelles gammes de variation des paramètres conduisent à des résultats biologiquement réalistes.

Des simulations ont ainsi été réalisées avec WALTer pour reproduire au mieux une partie de l'expérimentation réalisée à Lelystad par Darwinkel (1978), dans laquelle la dynamique de tallage de plantes de blé a été mesurée pour différentes densités de semis (allant de 5 à 800 plantes/m²). 56 plantes ont donc été simulées dans un semis en maillage (*i.e* distance inter-rang = distance inter-plantes) à des densités de 100, 200 et 400 plantes par m². Les données de rayonnement utilisées sont celles disponibles dans l'étude de Darwinkel (1978) réalisée à Lelystad et la séquence de température correspond aux moyennes journalières calculées entre 2004 et 2014 pour la même localité.

À l'exception des paramètres GAI_c , d_{GAIp} , PAR_t et Δ_{prot} , qui sont étudiés dans cette partie, les paramètres génotypiques utilisés correspondent aux valeurs estimées pour le génotype Maxwell

dans le cadre de la thèse de Christophe [Lecarpentier \(2017\)](#).

Paramètres GAI_c et d_{GAIp}

Dans un premier temps, les gammes de variation des paramètres GAI_c et d_{GAIp} ont été estimées. Pour chaque densité (100, 200 et 400 plantes par m^2), une simulation a été réalisée pour 10 valeurs de GAI_c (0,05 ; 0,1 ; 0,25 ; 0,5 ; 0,75 ; 1,0 ; 1,25 ; 1,5 ; 1,75 ; 2,0) et pour 2 valeurs de d_{GAIp} (0,3 et 1 m). Pour chaque combinaison de GAI_c / d_{GAIp} , 5 réplicats ont été réalisés, ce qui représente un total de 500 simulations pour cette étape. Afin de limiter le temps de calcul, les simulations ont été réalisées sans la phase de régression des talles (simulations de 300 jours sans utiliser de module de rayonnement).

Pour chaque simulation, le nombre moyen d'axe par plante dans la parcelle a été calculé et son maximum (N_{axes}^{max}) a été comparé aux valeurs expérimentales de [Darwinkel \(1978\)](#) et à des données sur la dynamique de tallage de différentes variétés de blé tendre issues de la thèse de [Lecarpentier \(2017\)](#).

Paramètres PAR_t et Δ_{prot}

Une fois l'estimation des paramètres GAI_c et d_{GAIp} effectuée, un travail similaire a été réalisé pour les paramètres PAR_t et Δ_{prot} . Pour chaque densité (100, 200 et 400 plantes par m^2), la simulation a été réalisée pour 15 valeurs de PAR_t (5000, 50000, 100000, 150000, 200000, 250000, 300000, 350000, 400000 ; 500000 ; 600000 ; 700000 ; 800000 ; 900000 ; 1000000 $\mu mol.cm^{-2}.degrés-jours^{-1}$) et 2 valeurs de Δ_{prot} (25 et 75 degrés-jours). Pour chaque combinaison de PAR_t / Δ_{prot} , 3 réplicats ont été réalisés, portant le nombre total de simulations de cette étape à 270.

Les paramètres GAI_c et d_{GAIp} ont chacun été fixés à la valeur optimale identifiée à l'étape précédente.

Pour chaque simulation, le nombre moyen d'axes par plante dans la parcelle a été calculé et sa valeur finale (N_{ears} , *i.e* le nombre moyen d'épis produits par plante) a été comparée aux valeurs expérimentales de [Darwinkel \(1978\)](#) et à des données tirées de [Lecarpentier \(2017\)](#).

Calcul des indices de sensibilité

À partir des 3645 simulations du plan factoriel fractionnaire, l'analyse de sensibilité a tout d'abord été réalisée sur les sorties résumées de WALTer suivantes (telles que définies dans le Chapitre 1) : le nombre moyen d'épis par plante (N_{ears}) et sa variance ($Var_{N_{ears}}$) ; le nombre maximum d'axes produit par plante en moyenne (N_{axes}^{max}) et sa variance ($Var_{N_{max}}$) ; le GAI maximum de la parcelle (GAI_{max}) et la date à laquelle il est atteint ($D_{GAI_{max}}$) ; le pourcentage moyen de lumière interceptée par la parcelle (L_{perc}) et la fitness totale de la parcelle (F_{tot}). Pour chacune de ces sorties, un modèle d'analyse de la variance, intégrant les interactions d'ordre 1 et 2 ainsi qu'une résiduelle, a permis d'estimer l'impact de chaque facteur d'entrée par le calcul des indices de sensibilité principaux (Msi : effet du facteur d'entrée seul ; équation 1) et des indices d'interactions d'ordre 2 (Isi : effet du facteur d'entrée en interaction avec un autre facteur d'entrée ; équation 2) ([Monod et al., 2006](#)). Pour chaque sortie, les indices d'interactions totaux ($Tisi$; équation 3) ont également été calculés, pour

estimer l'impact de toutes les interactions d'ordre 2 sur la variance. Ces indices sont calculés pour chaque facteur d'entrée (Tableau 1) ainsi que pour la graine d'aléatoire, qui est considérée comme un facteur d'entrée supplémentaire. L'évaluation de la résiduelle permet d'estimer l'importance des effets d'interaction d'ordre supérieur à 2.

Pour une sortie donnée, soient a et b des facteurs d'entrée ; n le nombre total de facteurs d'entrée de l'analyse ; SS_a la somme des carrés des écarts (SCE) associée à l'effet principal de a ; $SS_{a,b}$ la SCE associée à l'effet d'interaction entre a et b ; et SS_T la SCE totale de la sortie ; les différents indices sont calculés grâce aux formules suivantes :

$$Msi_a = \frac{SS_a}{SS_T} \quad (\text{Équation 1})$$

$$Isi_a = \frac{1}{SS_T} \sum_{\substack{b=1,n \\ a \neq b}} SS_{a,b} \quad (\text{Équation 2})$$

$$TIsi_a = \frac{1}{SS_T} \sum_{\substack{b=1,n \\ b=1,n \\ a < b}} SS_{a,b} \quad (\text{Équation 3})$$

Analyses complémentaires

Afin de compléter ces résultats, une analyse de sensibilité globale séquentielle, telle qu'implémentée dans le package multisensi (Lamboni *et al.*, 2009), a été réalisée sur la dynamique de tallage et sur la dynamique du GAI. Dans cette analyse, les indices de sensibilité sont calculés pour chaque pas de temps de chaque dynamique, puis leur évolution est représentée graphiquement.

Une analyse de sensibilité globale multivariée basée sur une analyse en composantes principales (ACP) (Campbell *et al.*, 2006) a également été réalisée via le package multisensi (Lamboni *et al.*, 2009). Dans cette analyse, une ACP est tout d'abord réalisée pour chaque dynamique (tallage et GAI). Les composantes principales ainsi extraites servent de sortie sur lesquelles les indices de sensibilité sont calculés pour chaque facteur d'entrée.

Enfin, pour chaque sortie résumée (N_{ears} , $\text{Var}_{\text{Nears}}$, $N_{\text{axes}}^{\text{max}}$, Var_{Nmax} , GAI_{max} , $D_{\text{GAI}_{\text{max}}}$, L_{perc} et F_{tot}), les facteurs les plus influents (qui expliquent plus de 10% de la variance de la sortie) ont été identifiés. Les valeurs de chaque sortie ont ensuite été représentées graphiquement en fonction des valeurs de leurs facteurs influents.

Les résultats de ces analyses complémentaires sont présentés dans l'Annexe 2.

RÉSULTATS

Estimation des gammes de variation pour les paramètres d'entrée GAI_c , d_{GAIp} , PAR_t et Δ_{prot}

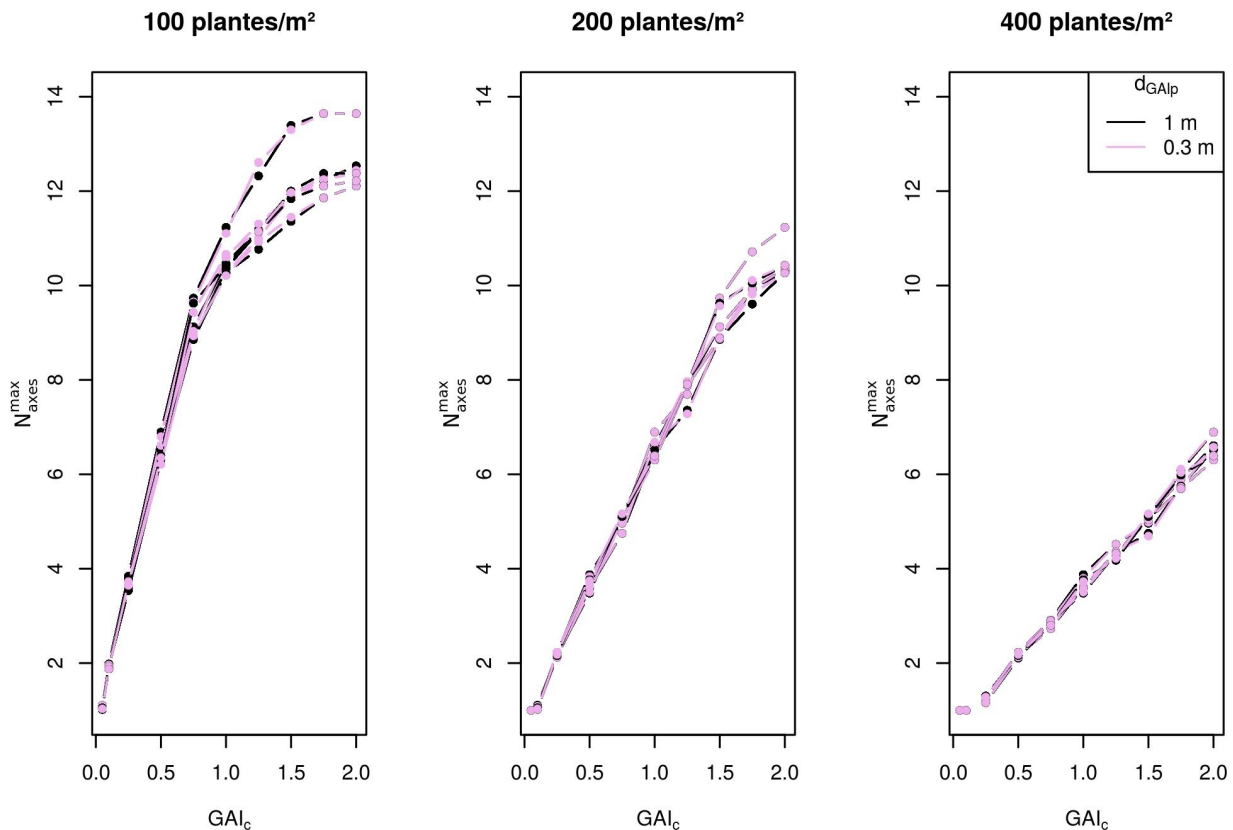


Figure 1. Moyenne parcellaire du nombre maximum d'axes par plante pour différentes valeurs de GAI_c . Les résultats sont présentés pour 3 densités de semis (de gauche à droite 100, 200 et 400 plantes par m²) et pour 2 valeurs de d_{GAIp} (1m ou 0,3m).

La figure 1 laisse apparaître que, pour une valeur de GAI_c donnée, le nombre maximum d'axes par plante (N_{axes}^{max}) est d'autant plus élevé que la densité est faible. La valeur du paramètre GAI_c influence fortement le nombre maximum d'axes produits : quelle que soit la densité, N_{axes}^{max} est d'autant plus important que la valeur de GAI_c est élevée. À densité 400 plantes/m², N_{axes}^{max} augmente linéairement avec l'élévation du GAI_c , en passant de 1 axe par plante pour une valeur de GAI_c de 0,05 à près de 7 axes par plante pour un GAI_c de 2,0. Cette augmentation linéaire se retrouve à la densité 200 plantes/m² pour des valeurs de GAI_c comprises entre 0,05 et 1,5 et à densité 100 plantes/m² pour des valeurs de GAI_c allant de 0,05 à 0,75. Pour les valeurs de GAI_c supérieures à 0,75 à densité 100 plantes/m² et supérieures à 1,5 à densité 200 plantes/m², N_{axes}^{max} semble tendre vers un plateau avec l'augmentation du GAI_c .

La figure 1 montre qu'il existe très peu de différences entre les résultats des simulations réalisées avec une valeur de d_{GAIp} de 0,3 et ceux des simulations faites avec une valeur de d_{GAIp} de 1, ce qui témoigne du faible impact de ce paramètre sur N_{axes}^{max} dans ces conditions.

La comparaison de ces résultats au nombre maximum d'axes produits par plante dans l'expérience de [Darwinkel \(1978\)](#) a permis d'identifier une valeur de GAI_c biologiquement réaliste de 0,75 qui a été utilisée pour les simulations d'exploration de la gamme de variation des paramètres PAR_t et Δ_{prot} . En comparant ces résultats à d'autres données de dynamique de tallage ([Lecarpentier, 2017](#)), la gamme de variation du paramètre GAI_c pour l'AS a été fixée afin de représenter des dynamiques biologiquement réalistes (Tableau 1).

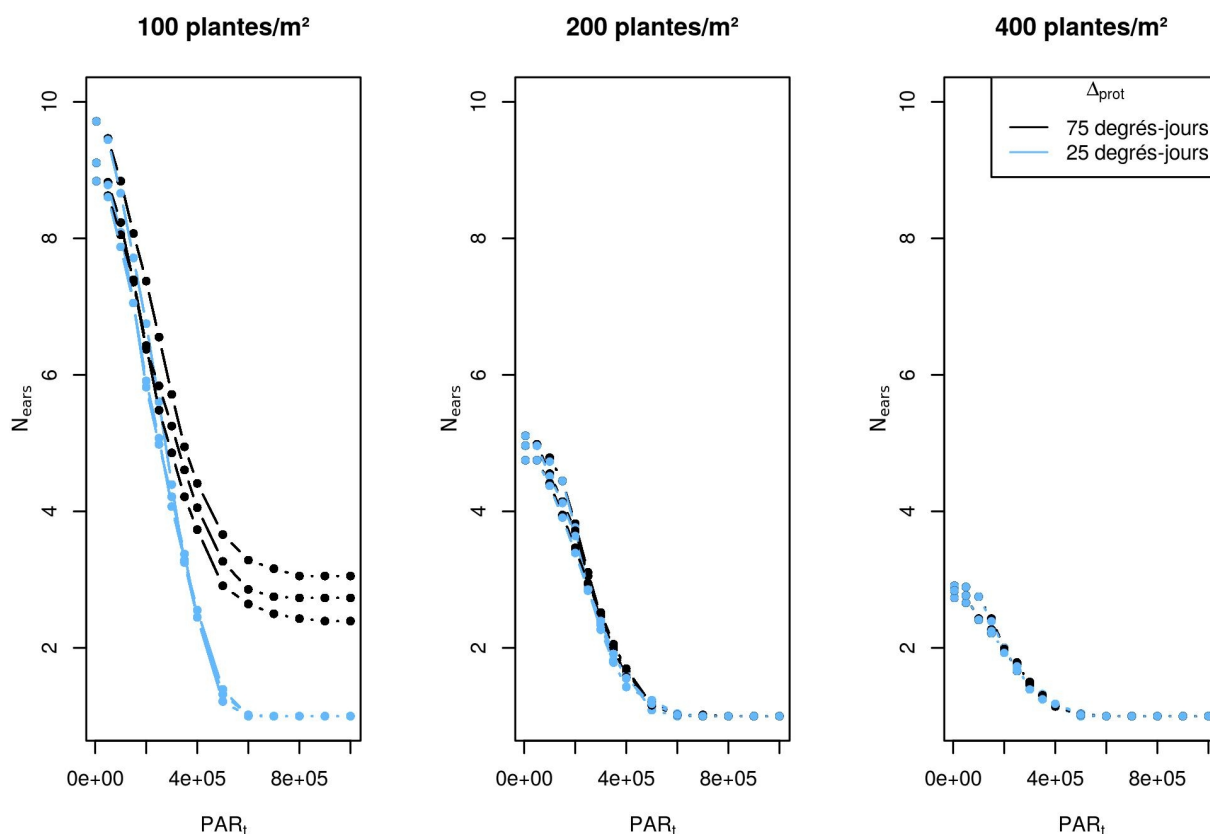


Figure 2. Moyenne parcellaire du nombre final d'épis par plante pour différentes valeurs de PAR_t . Les résultats sont présentés pour 3 densités de semis (de gauche à droite 100, 200 et 400 plantes par m²) et pour 2 valeurs de Δ_{prot} (25 degrés-jours ou 75 degrés-jours).

La figure 2 montre que la valeur du paramètre PAR_t influence fortement le nombre final d'épis produit par plante (N_{ears}), principalement à faible densité. En effet, à densité 100 plantes/m² par exemple, on observe en moyenne plus de 9 épis par plante pour un PAR_t de 5000 $\mu\text{mol}\cdot\text{cm}^{-2}\cdot\text{degrés-jours}^{-1}$, mais seulement 1 épis par plante à partir d'un PAR_t de 600000 $\mu\text{mol}\cdot\text{cm}^{-2}\cdot\text{degrés-jours}^{-1}$ (pour un Δ_{prot} de 25 degrés-jours). Pour une valeur de PAR_t donnée, N_{ears} est d'autant plus élevé que la densité est faible, excepté pour les valeurs élevées de PAR_t , pour lesquelles N_{ears} vaut 1 pour toutes les densités. Pour toutes les densités, on observe une diminution de N_{ears} avec l'augmentation de PAR_t qui tend vers un plateau pour les valeurs les plus élevées de PAR_t . Ce plateau de régression maximale est atteint pour un PAR_t de 600000 $\mu\text{mol}\cdot\text{cm}^{-2}\cdot\text{degrés-jours}^{-1}$ environ pour les densités 100 et 200 plantes/m² et pour un PAR_t de 500000 $\mu\text{mol}\cdot\text{cm}^{-2}\cdot\text{degrés-jours}^{-1}$ à densité 400 plantes/m².

Le paramètre Δ_{prot} semble avoir principalement de l'influence sur N_{ears} à faible densité. En effet, à 100 plantes/m² et pour des valeurs de PAR_t élevées, N_{ears} est significativement plus élevé pour un Δ_{prot} de 75 degrés-jours que pour un Δ_{prot} de 25 degrés-jours. Ainsi, si le plateau de régression

maximale correspond à un N_{ears} de 1 pour un Δ_{prot} de 25 degrés-jours (ce qui équivaut à la mort de toutes les talles émises), ce plateau est d'environ 3 épis par plante dans le cas des simulations avec $\Delta_{\text{prot}} = 75$ degrés-jours. Pour les densités plus élevées, l'impact de Δ_{prot} sur N_{ears} semble être réduit. La comparaison de ces résultats à des données expérimentales a permis de sélectionner une gamme de variation du paramètre PAR_i conduisant à des valeurs de N_{ears} biologiquement réalistes pour l'AS (Tableau 1).

Analyse de sensibilité

Les indices de sensibilités totaux, principaux et les indices des effets d'interactions d'ordre 2 ont été calculés sur les différentes sorties résumées pour chacun des 13 facteurs d'entrée ainsi que pour la graine d'aléatoire (rep). Ces indices, qui permettent d'estimer l'impact de chaque facteur d'entrée sur la variance de chaque sortie, sont résumés dans la figure 3.

La décomposition de la variance des sorties en effets principaux, d'interactions d'ordre 2 et en résiduelle permet tout d'abord d'améliorer notre compréhension du déterminisme de ces sorties. On note que la résiduelle, qui représente l'effet de certaines interactions d'ordre supérieur à 2, est faible (elle représente moins de 5% de la variance) pour la plupart des sorties. En revanche, elle s'élève à 7% pour N_{ears} et même à 12% et 15% pour $Var_{N_{\text{max}}}$ et $Var_{N_{\text{ears}}}$ respectivement, ce qui tend à montrer que ces sorties ont un déterminisme plus complexe, faisant intervenir des interactions d'ordre supérieur à 2. Pour certaines sorties, la variance est expliquée majoritairement par des effets principaux. C'est le cas des sorties descriptives de la dynamique de GAI (GAI_{max} et $D_{GAI_{\text{max}}}$) et des sorties qui décrivent la dynamique d'interception lumineuse (L_{perc} et F_{tot}), pour lesquelles les effets principaux représentent entre 73% et 83% de la variance totale selon la sortie considérée. En revanche, les sorties qui décrivent la dynamique de tallage (N_{ears} , $N_{\text{axes}}^{\text{max}}$, $Var_{N_{\text{ears}}}$ et $Var_{N_{\text{max}}}$) semblent avoir un déterminisme plus complexe, comme le laissait également penser l'analyse de la résiduelle. En effet, pour ces sorties, les effets des interactions d'ordre 2 sont plus importants, puisqu'ils représentent systématiquement plus de 30% de la variance totale de ces sorties, et peuvent aller jusqu'à expliquer 44% de la variance de $Var_{N_{\text{max}}}$ et 45% de la variance de $Var_{N_{\text{ears}}}$.

Il est ensuite possible de détailler le déterminisme de chaque sortie en étudiant l'impact de chaque facteur d'entrée. La figure 3 montre ainsi que, lorsque l'on considère l'ensemble des sorties, les facteurs d'entrée les plus influents sont le nombre final de feuilles sur le brin maître (N_{MS}^B), la densité de semis, la longueur finale des limbes (L_{max}^B), le paramètre GAI_c et la valeur du phyllochrone (Phl). Les effets cumulés de ces 5 facteurs d'entrée permettent d'expliquer au moins 80% de la variance de chaque sortie, à l'exception de la sortie $Var_{N_{\text{ears}}}$, pour laquelle ils permettent d'expliquer 76% de la variance totale. En particulier, l'importance du paramètre N_{MS}^B peut être soulignée, puisqu'il explique systématiquement plus de 10% de la variance de chaque sortie.

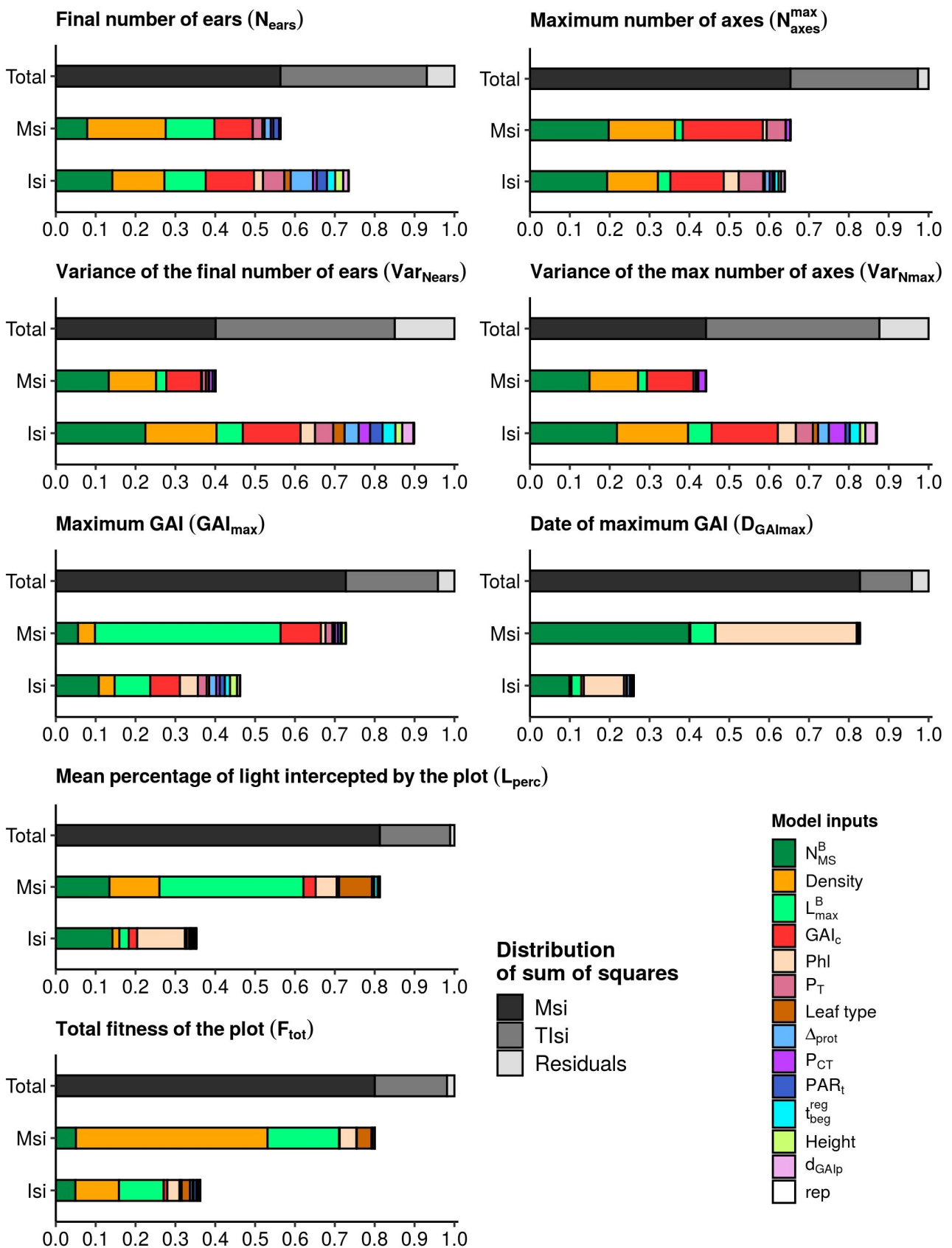


Figure 3. Indices de sensibilité pour le nombre final d'épis moyen par plante et sa variance ; le nombre maximum d'axes moyen par plante et sa variance ; le GAI maximum de la parcelle et la date à laquelle il est atteint ; le pourcentage moyen de lumière interceptée par la parcelle et la fitness totale de la parcelle. Les barres supérieures montrent la distribution des effets principaux, des effets d'interactions de second ordre et de la résiduelle. Les deux barres inférieures montrent l'impact des différents facteurs d'entrée et de la graine d'aléatoire (rep).

Si les 5 facteurs d'entrées cités ci-avant sont globalement les plus influents, on distingue cependant des différences en fonction du type de sortie concerné. Ainsi, si les sorties qui décrivent la dynamique de tallage sont toutes beaucoup impactées par N_{MS}^B , par la densité et par GAI_c , L_{max}^B a un fort impact sur N_{ears} (en expliquant 23% de sa variance) mais moins sur les 3 autres sorties. De la même manière, P_T a une influence importante sur N_{axes}^{max} , dont il explique 11% de la variance, mais un peu moins sur les autres sorties.

Pour ce qui est des sorties descriptives de la dynamique de GAI, plus de la moitié de la variance de GAI_{max} est expliquée par L_{max}^B . Les paramètres GAI_c et N_{MS}^B sont également importants pour cette sortie. De son côté, la variance de $D_{GAI_{max}}$ est largement expliquée par N_{MS}^B (50%) et Phl (46%).

Enfin, les sorties qui décrivent la dynamique d'interception lumineuse sont fortement impactées par la densité (cette entrée explique près de 60% de la variance totale de F_{tot}), mais également par L_{max}^B , par N_{MS}^B et par Phl .

Malgré les différences observées entre sorties, certains facteurs d'entrée n'ont que très peu d'impact sur l'ensemble des sorties. C'est le cas de la graine d'aléatoire (rep), de d_{GAIp} , de t_{beg}^{reg} et de PAR_t , qui expliquent tous moins de 4% de la variance de chaque sortie. L'impact de rep sur l'ensemble des sorties est même négligeable, puisque le pourcentage de variance expliquée par cette entrée est systématiquement inférieur à 0,25%. En outre, il est intéressant de noter que, pour ces quatre entrées (rep, d_{GAIp} , t_{beg}^{reg} et PAR_t), ainsi que pour le paramètre Δ_{prot} , les effets observés sont très majoritairement des effets d'interactions. Enfin, l'importance particulière de certaines interactions peut être soulignée. Ainsi, l'interaction entre N_{MS}^B et la densité, impacte fortement toutes les sorties de la dynamique de tallage et l'interaction entre N_{MS}^B et Phl a une forte influence sur $D_{GAI_{max}}$ et L_{perc} .

DISCUSSION

L'analyse de sensibilité (AS) est une étape essentielle du développement d'un modèle. L'AS de WALTer, par une méthode de type criblage, a permis d'améliorer notre compréhension du fonctionnement du modèle et d'identifier clairement les entrées (variables et paramètres) qui ont le plus d'impact sur les sorties du modèle.

En particulier, l'analyse a mis en avant l'impact majeur du paramètre N_{MS}^B sur l'ensemble des sorties étudiées. Ce résultat est attendu, car il est en accord avec les concepts écophysologiques intégrés au modèle. En effet, le nombre final de feuilles sur le brin maître a une influence forte sur le développement de la plante via 3 mécanismes. Premièrement, le nombre final de feuilles impacte directement la surface photosynthétique de la plante et donc l'interception lumineuse et la compétition pour la lumière. D'autre part, le nombre final de feuilles est en lien étroit avec la phénologie de la plante, comme en témoigne l'importance de ce paramètre sur $D_{GAI_{max}}$. En effet, plus une plante a un nombre de feuilles important, plus sa floraison est tardive. Enfin, les talles étant émises à l'aisselle des feuilles, au niveau des bourgeons axillaires, le nombre final de feuilles du

brin maître détermine le nombre de sites de tallage et donc le nombre potentiel de talles sur la plante. C'est pourquoi N_{MS}^B est l'entrée qui a le plus d'impact sur N_{axes}^{max} , avec un nombre maximum de talles d'autant plus important que le nombre final de feuilles sur le brin maître est élevé (Annexe 2 - Figure 5).

L'AS a également révélé que les variations de densité de semis, en changeant radicalement l'intensité de la compétition pour la lumière, impactent fortement les sorties du modèle. La forte influence de la densité sur les sorties descriptives de la dynamique de tallage, témoigne du bon fonctionnement du modèle. En effet, les résultats montrent que le nombre maximum de talles produites et le nombre final d'épis par plante sont d'autant plus faibles que la densité de semis est élevée (Figures 1 et 2 et Annexe 2 - Figure 5). Ces éléments sont en accord avec les observations expérimentales (Darwinkel, 1978 ; Sparkes *et al.*, 2006) qui mettent en avant la réduction importante de la ramification en lien avec l'augmentation de la densité, mécanisme caractéristique du syndrome d'évitement de l'ombre ("*shade avoidance syndrome*") (Franklin, 2008) observé chez le blé et de nombreuses autres graminées cultivées (Gommers *et al.*, 2013).

Si la densité impacte fortement les sorties de la dynamique de tallage et de la dynamique d'interception lumineuse, l'impact de cette variable est moindre sur les sorties de la dynamique de GAI. Ce résultat peut s'expliquer par l'effet de compensation lié au tallage : le tallage accru à faible densité permet de compenser partiellement la baisse de GAI associée à une plus faible densité de semis. Cette hypothèse est appuyée par des observations expérimentales qui ont montré que la plasticité du tallage permettait de diminuer fortement l'écart de GAI entre des parcelles de blé semées à 30 plantes par m² et des parcelles semées à 150 plantes par m² (Lecarpentier, 2017).

Comme cela pouvait être attendu, l'AS a également mis en avant la forte influence de la longueur des limbes sur le GAI de la parcelle et sur les sorties d'interception lumineuse. Les variations de cette entrée ont également un fort impact sur le nombre moyen d'épis par plante : N_{ears} est d'autant plus faible que les limbes sont longs (Annexe 2 - Figure 5). La réponse de N_{ears} aux variations de L_{max}^B s'explique probablement par des effets d'ombrages qui contrôlent la régression des talles. Ce résultat va dans le sens de nombreuses études, comme celles de Donald (1968) ou de Anten et Vermeulen (2016), qui avancent que la culture de variétés avec de petites feuilles, et donc un potentiel compétitif moindre, permettrait d'aboutir à de meilleures performances à l'échelle de la parcelle.

L'AS a également révélé le rôle central du paramètre GAI_c , dont la valeur impacte fortement les sorties de la dynamique de tallage et GAI_{max} . Ce résultat souligne l'importance de l'étape de détection précoce du voisinage dans la régulation du tallage (Evers *et al.*, 2007 ; Sparkes *et al.*, 2006) telle qu'implémentée dans WALTER.

En revanche, PAR_i , qui est un paramètre clef de l'étape de régression des talles (Xie *et al.*, 2016) dans le modèle, semble avoir un impact très limité sur les sorties étudiées. Il apparaît pourtant surprenant que l'analyse fasse ressortir une faible importance de PAR_i au vu de la construction du modèle et des résultats de l'AS réalisée sur la précédente version de WALTER (Lecarpentier *et al.*, 2019), laquelle avait montré un impact majeur de ce paramètre sur les sorties du modèle. Les simulations réalisées pour estimer la gamme de variation de PAR_i (Figure 2) ont d'ailleurs montré que des variations de ce paramètre peuvent conduire à des changements importants dans la valeur

de N_{ears} pour les conditions testées. Il est alors probable que le faible impact de PAR_t identifié dans l'AS soit lié à un choix peu adapté de sa gamme de variation. En effet, l'impact de PAR_t sur N_{ears} par exemple, est dépendant des étapes d'émission des talles et de cessation du tallage qui ont lieu plus tôt dans les simulations et sont contrôlées par les valeurs d'un certain nombre d'autres entrées (N_{MS}^B , GAI_c , densité et P_T notamment, Figure 3) ; ainsi que de la valeur de Δ_{prot} , comme en témoignent les interactions $\Delta_{\text{prot}}/\text{PAR}_t$ mises en avant dans la Figure 2. La gamme de variation de PAR_t sélectionnée pour l'analyse permet donc bien d'explorer des fluctuations de sorties dans certaines conditions (conditions dans lesquelles cette gamme a été sélectionnée, à savoir : GAI_c peu restrictif et nombre final de feuilles suffisant pour permettre l'émission d'un nombre suffisant de talles, densité de semis restreinte, Δ_{prot} limité...). Cependant, pour les configurations de paramètres pour lesquelles ces conditions ne sont pas réunies, une gamme de variation élargie serait probablement nécessaire pour révéler l'influence de PAR_t sur le déterminisme des sorties de WALTER.

D'autre part, il est intéressant de noter que le phyllochrone (Phl) fait partie des entrées les plus influentes identifiées par l'AS (avec un fort impact sur les sorties d'interception lumineuse et sur $D_{\text{GAI}_{\text{max}}}$). Le fort impact de ce paramètre, associé à l'influence majeure de N_{MS}^B , souligne la capacité du modèle à représenter l'importance bien établie des paramètres phénologiques sur le développement des parcelles (Jamieson *et al.*, 2007).

Il est également intéressant de souligner le très faible impact de certaines entrées, dont les valeurs peuvent être fixées sans risquer de modifier fortement les résultats des simulations (d_{GAI_p} et la hauteur finale du brin maître notamment). Ces résultats ne semblent pas surprenants étant donné que l'AS a été réalisée en simulant des cultures pures. Cependant, il est intéressant de garder à l'esprit que des simulations de mélanges de variétés pourraient donner des résultats différents : le rôle important de la hauteur sur la compétition pour la lumière est bien documenté (Falster et Westoby, 2003 ; Mason *et al.*, 2008), ce qui en fait un trait d'intérêt pour l'étude des mélanges variétaux de blé (Borg *et al.*, 2018).

L'effet négligeable de la graine d'aléatoire (rep) sur la variance de l'ensemble des sorties peut également être souligné. Ce résultat appuie les conclusions du Chapitre 2 sur la non-nécessité des répliqués.

Enfin, cette analyse permet d'orienter les efforts de calibration du modèle en se concentrant sur l'estimation des entrées les plus influentes. En effet, fixer la valeur de ces entrées sans connaissances a priori reviendrait à faire des hypothèses fortes pouvant avoir des conséquences importantes sur les sorties des simulations. Ainsi, le nombre final de feuilles sur le brin maître et la longueur finale des limbes devront être estimés avec précision lorsqu'ils n'ont pas été mesurés. De la même manière, le paramètre GAI_c , qui est très peu documenté, devra être estimé par calibration. De plus, bien que l'AS ait montré un faible impact des paramètres Δ_{prot} et PAR_t sur les sorties du modèle, l'estimation de ces paramètres par calibration semble être indiquée étant donné que (i) l'impact de ces paramètres, et principalement de PAR_t , est probablement sous-estimé par l'AS du fait d'une gamme de variation peu adaptée et (ii) que ces paramètres ne sont pas mesurables expérimentalement et qu'aucune information n'est disponible sur leur valeur. En revanche, il ne devrait pas être nécessaire d'inclure la densité dans la calibration du modèle, malgré son influence majeure sur les sorties, car cette variable est normalement toujours connue dans les

expérimentations.

Cette analyse de sensibilité approfondie a ainsi permis d'améliorer notre compréhension de la nouvelle version de WALTer, de confirmer le bon fonctionnement du modèle et d'orienter les efforts de calibration.

RÉFÉRENCES

Abichou, M. (2016). Modélisation de l'architecture 4D du blé : identification des patterns dans la morphologie, la sénescence et le positionnement spatial des organes pour une large gamme de situations de croissance. Thèse de doctorat, Université Paris-Saclay.

Anten, N. P., & Vermeulen, P. J. (2016). Tragedies and crops: understanding natural selection to improve cropping systems. *Trends in Ecology & Evolution*, 31(6), 429-439.

Baccar, R. (2011). Plasticité de l'architecture du blé d'hiver modulée par la densité, la date de semis et son effet sur les épidémies de *Septoria tritici*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement.

Borg, J., Kiær, L. P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., ... & Enjalbert, J. (2018). Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*, 221, 298-313.

Campbell, K., McKay, M. D., & Williams, B. J. (2006). Sensitivity analysis when model outputs are functions. *Reliability Engineering & System Safety*, 91(10-11), 1468-1472.

Cariboni, J., Gatelli, D., Liska, R., & Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. *Ecological modelling*, 203(1-2), 167-182.

Darwinkel, A. (1978). Patterns of tillering and grains production of winter wheat at a wide range of plant densities. *Netherland Journal of Agricultural Science*, 26:383-398.

Donald, C. T. (1968). The breeding of crop ideotypes. *euphytica*, 17(3), 385-403.

Drouet, J. L., & Pagès, L. (2007). GRAAL-CN: A model of GRowth, Architecture and ALlocation for Carbon and Nitrogen dynamics within whole plants formalised at the organ level. *Ecological Modelling*, 206(3-4), 231-249.

Evers, J. B., Vos, J., Chelle, M., Andrieu, B., Fournier, C., & Struik, P. C. (2007). Simulating the effects of localized red: far-red ratio on tillering in spring wheat (*Triticum aestivum*) using a three-dimensional virtual plant model. *New Phytologist*, 176(2), 325-336.

Evers, J. B., Vos, J., Yin, X., Romero, P., Van Der Putten, P. E. L., & Struik, P. C. (2010). Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *Journal of Experimental Botany*, 61(8), 2203-2216.

Falster, D. S., & Westoby, M. (2003). Plant height and evolutionary games. *Trends in Ecology & Evolution*, 18(7), 337-343.

Franklin, K. A. (2008). Shade avoidance. *New Phytologist*, 179(4), 930-944.

Godin, C., Sinoquet, H. (2005). Functional-structural plant modelling. *New phytologist*, 166(3), 705-708.

Gommers, C. M., Visser, E. J., St Onge, K. R., Voeselek, L. A., & Pierik, R. (2013). Shade tolerance: when growing tall is not an option. *Trends in plant science*, 18(2), 65-71.

- Le Gratiet, L., Marelli, S., & Sudret, B. (2017).** Metamodel-based sensitivity analysis: polynomial chaos expansions and Gaussian processes. In *Handbook of Uncertainty Quantification*. Springer International Publishing.
- Helton, J. C., Johnson, J. D., Sallaberry, C. J., & Storlie, C. B. (2006).** Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10-11), 1175-1209.
- Iooss, B., & Lemaître, P. (2015).** A review on global sensitivity analysis methods. In *Uncertainty management in simulation-optimization of complex systems* (pp. 101-122). Springer, Boston, MA.
- Iooss, B., & Saltelli, A. (2017).** Introduction to Sensitivity Analysis. In *Handbook of Uncertainty Quantification* (pp. 1103–1122).
- Jamieson, P. D., Brooking, I. R., Semenov, M. A., McMaster, G. S., White, J. W., & Porter, J. R. (2007).** Reconciling alternative models of phenological development in winter wheat. *Field Crops Research*, 103(1), 36-41.
- Kang, F., Cournede, P. H., Lecoeur, J., & Letort, V. (2014).** SUNLAB: a functional–structural model for genotypic and phenotypic characterization of the sunflower crop. *Ecological modelling*, 290, 21-33.
- Kobilinsky, A., Bouvier, A., Monod, H. (2017).** PLANOR: an R package for the automatic generation of regular fractional factorial designs. R package version 1.3-6
- Lamboni, M., Makowski, D., Lehuger, S., Gabrielle, B., & Monod, H. (2009).** Multivariate global sensitivity analysis for dynamic crop models. *Field Crops Research*, 113(3), 312-320.
- Lecarpentier, C. (2017).** Plasticité de l'architecture aérienne du blé en réponse à la compétition pour la lumière au sein de cultures pure ou d'associations variétales : caractérisation expérimentale et développement d'un modèle 3D. Thèse de doctorat, Université Paris-Saclay.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Luquet, D., Soulié, J. C., Rebolledo, M. C., Rouan, L., Clément-Vidal, A., & Dingkuhn, M. (2012).** Developmental dynamics and early growth vigour in rice 2. Modelling genetic diversity using Ecomeristem. *Journal of Agronomy and Crop Science*, 198(5), 385-398.
- Mason, H., Goonewardene, L., & Spaner, D. (2008).** Competitive traits and the stability of wheat cultivars in differing natural weed environments on the northern Canadian Prairies. *The Journal of Agricultural Science*, 146(1), 21-33.
- Monod, H., Naud, C., Makowski, D. (2006).** Uncertainty and sensitivity analysis for crop models. Working with dynamic crop models: Evaluation, analysis, parameterization, and applications, 4, 55-100.
- Morris, M.D., 1991.** Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, 161–174.

- R Core Team (2018).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Saltelli, A., Tarantola, S., & Chan, K. S. (1999).** A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1), 39-56.
- Saltelli, A., Tarantola, S., & Campolongo, F. (2000).** Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15(4), 377-395.
- Sobol, I. M. (1993).** Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4), 407-414.
- Sparkes, D. L., Holme, S. J., & Gaju, O. (2006).** Does light quality initiate tiller death in wheat?. *European Journal of Agronomy*, 24(3), 212-217.
- Turányi, T. (1990).** Sensitivity analysis of complex kinetic systems. Tools and applications. *Journal of mathematical chemistry*, 5(3), 203-248.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- Woods, D. C., & Lewis, S. M. (2015).** Design of experiments for screening. *arXiv preprint arXiv:1510.05248*.
- Wu, Q. (2012).** Sensitivity Analysis for Functional Structural Plant Modelling. Doctoral dissertation, École centrale de Paris.
- Xie, Q., Mayes, S., & Sparkes, D. L. (2016).** Optimizing tiller production and survival for grain yield improvement in a bread wheat× spelt mapping population. *Annals of Botany*, 117(1), 51-66.

CHAPITRE 4

Calibration of WALTer using an adaptive design and a metamodelling approach

INTRODUCTION

Modelling is a powerful tool to study complex systems. Numerical experiments are often used when real experiments are too difficult or too expensive to carry out. Furthermore, computer models allow to strictly control the conditions of the experiments and to explore conditions beyond what can be covered experimentally. The advantages of computer models have made them essential tools in many disciplines. For example, modelling is widely used in plant sciences and agronomy and has been identified as a promising tool to tackle some of the challenges associated with major issues such as food security (Christensen *et al.*, 2018). In particular, functional-structural plant models (FSPM ; Godin & Sinoquet, 2005 ; Vos *et al.*, 2010) have been more and more used since the 1990s, and they have become a major research subject (Guo *et al.*, 2011 ; Sievänen *et al.*, 2014 ; Evers *et al.*, 2018). These individual-based models explicitly represent the plant architecture in 3D. Thus, FSPM take into account the complex interactions between environmental factors, plant development, plant architecture and the underlying physiological processes. These models allow the integration of different scales (from gene to community) to study complex systems in plant science. Thereby, FSPM have been used to study a wide variety of species (both wild and cultivated) and a broad range of questions. Examples of application include prediction of fire behaviour in tree crowns (Parsons *et al.*, 2011) ; study of the evolutionary emergence of life history strategies along gradients of stress intensity and disturbance frequency (Bornhofen *et al.*, 2011) ; and study of the impact of plant height on the control of rain-borne diseases in wheat cultivar mixtures (Vidal *et al.*, 2018).

In order to obtain reliable results from FSPM, or from any model, their input parameters must be accurately set. Indeed, the quality of the parameter estimation is critical for the quality of the model predictions. Some parameters can be directly fixed based on the literature or on data. However, some other parameters cannot be directly measured and are not available in the literature. The values of these parameters have to be indirectly estimated via calibration, which means that the overall model has to be fitted to experimental data. Calibration can be done either manually or automatically. The latter method relies on the use of a search algorithm to identify the “optimal” values of the specified parameters based on the minimization of the distance between the model outputs and the observed data. Automatic calibration has a number of advantages over manual calibration. Indeed, automatic calibration is less subjective than manual calibration, as its success is less dependent on the experience of the modeller (Muleta and Nicklow, 2005). Furthermore, automatic calibration may allow the modeller to look for a distribution of likely values of the parameters instead of looking for an optimal parameter value. Bayesian calibration (Kennedy and O’Hagan, 2001 ; Higdon *et al.*, 2004 ; Bayarri *et al.*, 2007) is an automatic calibration method which also provides the modeller with a distribution of values of the parameters that are likely to reproduce the experimental data. The experimental data and the FSPM are linked in a statistical model from which a likelihood is then derived. A prior distribution, which may encode some expert or literature information on the parameter values, is chosen. Since the corresponding posterior distribution is not tractable, MCMC algorithms are run to obtain a sample in the posterior distribution. However, these algorithms resort to many runs of the model, which is why, for complex models, including FSPM, calibration is often done manually (see Lecarpentier *et al.*, 2019 and Gauthier *et al.*, 2020 for example). Indeed, FSPM runs are usually time-consuming and parallel

computing of each simulation is complicated because of the interactions between plants in these individual-based models. The solution proposed in the context of Bayesian calibration is to make recourse to a Gaussian Process (Currin *et al.*, 1991 ; Sacks *et al.*, 1989) metamodel, also known as Kriging metamodel, which provides not only a fast approximation of the FSPM but also a measure of uncertainty concerning the quality of the approximation of the FSPM. Metamodels are widely used to approximate complex models and have sometimes been used with FSPM to allow analyses that would require an important number of simulations (Perez *et al.*, 2018 ; Da Silva *et al.*, 2014). However, as metamodels are an approximation of the model, the calibration can be less precise and less efficient than with the actual FSPM. To address this problem, sequential methods aiming at improving the Kriging metamodel precision and based on the EGO algorithm (Jones *et al.*, 1998) can be implemented, as proposed in Damblin *et al.* (2018). These methods have been used for the calibration of complex models (see Carmassi *et al.*, 2019 for example) but have not yet, to the extent of our knowledge, been applied to FSPM.

Here, we present how a Kriging metamodel and an adaptive design can be used for the calibration of the WALTer FSPM.

MATERIAL AND METHODS

All the analyses were done using the R software (R Core Team, 2017).

WALTer: a 3D wheat model

WALTer (Lecarpentier *et al.*, 2019) is a functional structural plant model (FSPM) (Vos *et al.*, 2010 ; Godin & Sinoquet, 2005). This individual-based model simulates the development of the aerial architecture of winter wheat (*Triticum aestivum* L.) from sowing to maturity with a daily time step. In the model, the vegetative development of the plants follows a thermal time schedule and is based on a formalism derived from ADEL-Wheat (Fournier *et al.*, 2003). Thanks to a radiative model (CARIBU: Chelle and Andrieu, 1998), WALTer simulates the competition for light between plants within the field and the resulting plasticity of tillering (i.e. the branching ability of grasses). The regulation of tillering in the model is based on three simple rules. (i) Tillers emerge according to empirically fixed probabilities. Then, (ii) an early neighbour perception controls the cessation of tillering: plants stop emitting tillers when the surrounding Green Area Index^a reaches a critical value (GAI_c). Finally, (iii) some of the tillers that were emitted regress: a tiller regresses if the amount of photosynthetically active radiation (PAR) it intercepts falls below a threshold (PAR_t). During tiller regression, there is a protection period (Δ_{prot}) between the death of two successive tillers on a plant.

Based on this formalism, WALTer produces useful outputs, such as the tillering dynamics of each plant (i.e. the number of axes on a plant for each day of the simulation).

WALTer is not a completely deterministic model, as several elements in the model integrate some stochasticity. Indeed, the position of the plants, the orientation of the organs, the duration before plant emergence and the emergence of tillers are partly random. Furthermore, the final number of

a The GAI is the ratio of photosynthetic surface of a plant to ground area.

leaves on the main stem (N_{MS}^B) is defined at the scale of the field: it is a decimal number and the decimal part indicates the proportion of plants carrying round ($N_{MS}^B + 0.5$) leaves on the main stem.^b

Since its publication in 2019 (Lecarpentier *et al.*), WALTer has undergone some changes that are described in Chapter 1.

Parameter estimation

The values of five parameters of the model (Table 1) were estimated by fitting the tillering dynamics. These five parameters were selected based on the results of a global sensitivity analysis of WALTer (see Chapter 3) because they had an important impact on the variance of the tillering dynamics and/or because they could not be measured or estimated otherwise.

The parameters that were not estimated here were set according to the bibliography and the manual calibration described in Lecarpentier *et al.*, 2019 for the winter wheat cultivar 'Maxwell' (Appendix 1). For meteorological data, the original sequence of PAR available in Darwinkel (1978) for Lelystad (The Netherlands) and a sequence of daily temperatures obtained by averaging Lelystad data from 2004 to 2014 were used for the simulations.

Table 1. List of the 5 WALTer parameters to be estimated, definition, ranges of variation explored for the calibration, values used to generate the simulated test-dataset and units.

Parameter	Description	Range	Value for the simulated data	Unit
GAI_c	Green Area Index threshold above which the emission of tillers stops	0.25 - 1.25	0.75	-
PAR_t	PAR threshold below which a tiller does not survive	$10^5 - 10^6$	3×10^5	$\mu\text{mol} \cdot \text{cm}^{-2} \cdot \text{Cd}^{-1}$
Δ_{prot}	Thermal time interval during which two tillers of the same plant cannot die	10 - 100	25	$^{\circ}\text{Cd}$
N_{MS}^B	Final number of leaves on the main stem	8 - 16	12.5	-
L_{max}^B	Final length of the longest blade of the main stem	8 - 35	16.6	cm

The proposed calibration method is based on the Efficient Global Optimization (EGO) algorithm (Jones *et al.*, 1998) and involves the following six steps:

(1) First, an initial numerical design must be selected to sample the multidimensional parameter space. To ensure a good exploration of the parameter space, a five dimensions maximin Latin

^b For example, a field with $N_{MS}^B = 11.6$ would have 40% of plants with 11 leaves on their main stem and 60% of plants with 12 leaves.

Hypercube (LH) design (McKay *et al.*, 1979 ; Johnson *et al.*, 1990) of 100 parameter-sets was generated with the DiceDesign package (Dupuy *et al.*, 2015) using the ranges detailed in Table 1.

(2) The model must then be run over the initial design. All the simulations of the LHD were run with WALTER for 6 contrasted sowing densities (25, 50, 100, 200, 400 and 800 plants/m²), leading to a total of 600 simulations. For each simulation, the mean tillering dynamics of the field was computed by WALTER and a subset of 13 dates spread across the dynamics was extracted.

(3) The initial design must then be used to build a Kriging metamodel which approximates WALTER in the parameter space. For each sowing density and for each of the 13 dates, a metamodel was fitted by using the DiceKriging package (Roustant *et al.*, 2012). For each metamodel, the homogeneous nugget effect was set to WALTER's variance for the corresponding date and sowing density. WALTER's variance was computed thanks to 10 replicates of a reference simulation (Table 1). The accuracy of each Kriging metamodel was tested through cross validation by leave-one-out.

(4) The initial design could then be enriched by the addition of new points. The additional points are selected based on the Expected Improvement (EI) criterion (Damblin *et al.*, 2018), balancing the need to minimize the deviation of the metamodel prediction from the calibration data set, with the need to reduce the uncertainty of the surrogate model. For this step, we decided to enrich the design by the addition of a single point, corresponding to the highest EI criterion among a set of 10 000 points selected from the parameter space by an LH sampling.

(5) A Bayesian calibration method can then be applied to the model. For each input parameter, the posterior distribution was estimated using a Markov Chain Monte Carlo (MCMC) sampling. More precisely, a random walk Metropolis algorithm was used, as implemented in the MCMCpack package (Martin *et al.*, 2011). The log-likelihood was computed at each iteration as:

$$\log p(x|y) = \frac{1}{2} \times \sum_{i=1}^j \left(-\log(\text{var}_i + k_v) - \left(\frac{y_i - k_m(x)}{\sqrt{\text{var}_i + k_v(x)}} \right)^2 \right) \quad (\text{Equation 1})$$

With x the vector of model parameters, y the vector of observations, $\log p(x|y)$ the log-likelihood, j the number of outputs considered (number of elements in y). k_m is the function giving the Kriging mean, k_v is the function giving the Kriging variance, var is the vector of the observational variance of each output.

The observational variance of each output and the uncertainty of the Kriging metamodel were taken into account to avoid overfitting of the model.

(6) The goodness of fit can then be evaluated through cross-validation. For this step, we calibrated the model using only the data from one (200 plants/m²) sowing density and used the whole data set for the validation.

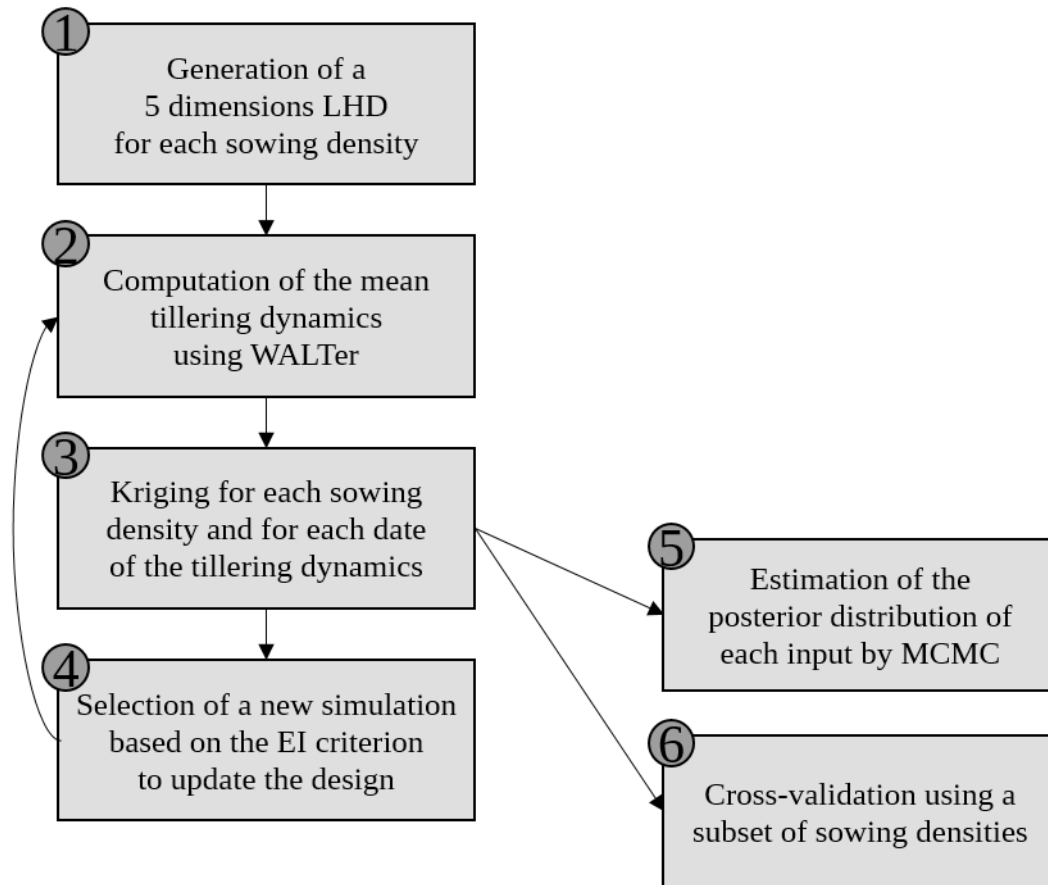


Fig. 1. Diagrammatic representation of the proposed calibration method

Application to synthetic data

At first, we applied our methodology to a simulated dataset. Using WALTer, we simulated tillering dynamics at six contrasted sowing densities (25, 50, 100, 200, 400 and 800 plants/m²) with a fixed set of parameters (Table 1). A Gaussian white noise was then added to the resulting tillering dynamics. For each date and at each sowing density, the noise had a standard deviation corresponding to WALTer's standard deviation between replicates at the same date and for the same sowing density. For each sowing density, the observational variance of each date was set to the value of WALTer's variance for the corresponding date at the same sowing density.

Application to experimental data

The same methodology was then applied to estimate the five parameter values that would allow the best fitting to experimental data. We used data from an experiment described by [Darwinkel \(1978\)](#), in which the winter wheat cultivar 'Lely' was sown in plots of 1 m² at seven contrasted sowing densities (5, 25, 50, 100, 200, 400 and 800 plants/m²) in Lelystad (The Netherlands). For each density, the number of shoots per m² was measured at 13 dates spread across the development period of the crop, thus giving a good estimation of the total tillering dynamics.

We decided not to use the data for the lowest sowing density (5 plants/m²) because simulations with WALTer for this sowing density would take a very long time. Furthermore, this sowing density

represents extreme conditions that would rarely be observed and is thus not the most interesting density to study. Therefore, we focused our calibration on the 6 other sowing densities (25, 50, 100, 200, 400 and 800 plants/m²).

For each sowing density, the observational variance, necessary to compute the likelihood (Eq. 1), was set to achieve a coefficient of variation of 20% for the first two dates, of 5% for the last two dates and of 10% for the other dates. These coefficients of variation were chosen to take into account the uncertainty associated with the experimental measurements. A more important coefficient of variation was selected for the first two dates to account for the relatively higher difficulty of measurement at the beginning of the tillering dynamics and for the temporal shift due to the discrepancy between the simulated and the experimental temperature sequences. The coefficient of variation for the last two dates of the tillering dynamics was set to a smaller value to account for the relative simplicity of measurement on mature plants and for the importance of the fit for these dates as they represent an important component of the yield (number of spikes/m²).

RESULTS

Application to synthetic data

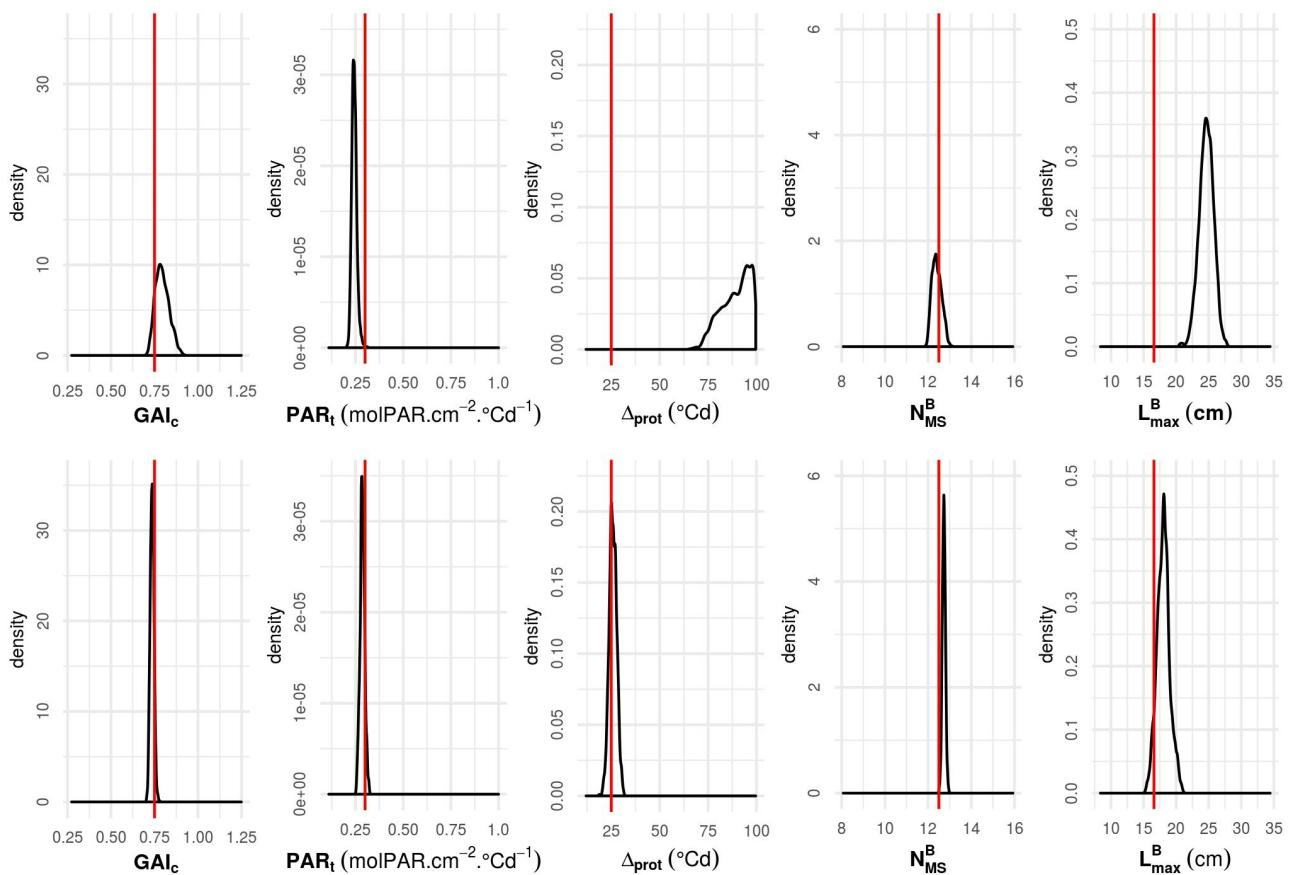


Fig. 2. Posterior distributions of WALTER parameters (GAI_c , PAR_t , Δ_{prot} , N_{MS}^B and L_{max}^B) using the simulated data from all densities and the initial LH design of 100 simulations (top row) or the enriched design after 15 loops of EGO (bottom row). Parameter values used to simulate the data are shown with a red vertical line.

The accuracy of the calibration is low when it is performed using the Kriging metamodel constructed from the initial design of 100 simulations (Fig 2, top row). Indeed, the posterior distributions of Δ_{prot} and L_{max}^B do not include the actual values used to generate the synthetic data. The posterior density for PAR_t has a small variance but is shifted from the actual parameter value and the posterior densities of GAI_c and N_{MS}^B include the actual parameter values but have rather important variances. On the other hand, the accuracy of the calibration is greatly improved by the use of the sequential design, even with an addition of only 15 points. Using this enriched design, the posterior distributions of all input parameters have small variances and include the values used to generate the synthetic data (Fig 2, bottom row). Thus, as shown in Figure 3, the enriched design allows to select a set of parameters that reproduces accurately the simulated tillering dynamics for all sowing densities.

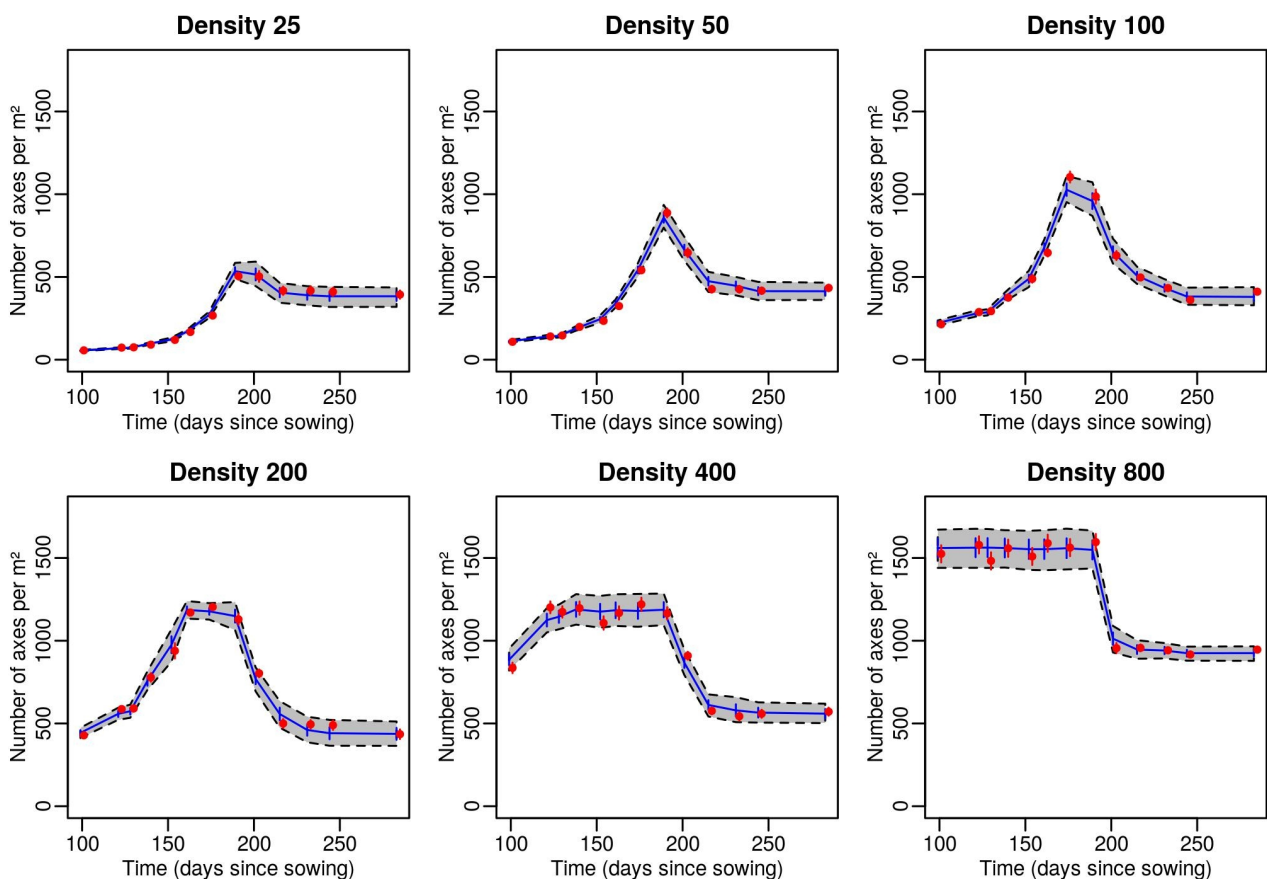


Fig 3. Number of axes per m^2 vs. time since sowing at six sowing densities (25, 50, 100, 200, 400 and 800 plants.m^{-2}). Red dots represent the synthetic data and red segments are the associated observational variance ; blue lines represent the mean Kriging prediction for the set of input parameters with the best loglikelihood and blue segments are the associated Kriging standard deviation. The grey area delimited by dotted lines is the 95% credibility interval taking into account the Kriging standard deviation. Red points and segments are shifted on the x-axis to avoid overlapping. The set with the best loglikelihood was selected by MCMC using the enriched design after 15 loops of EGO and simulated data from all sowing densities.

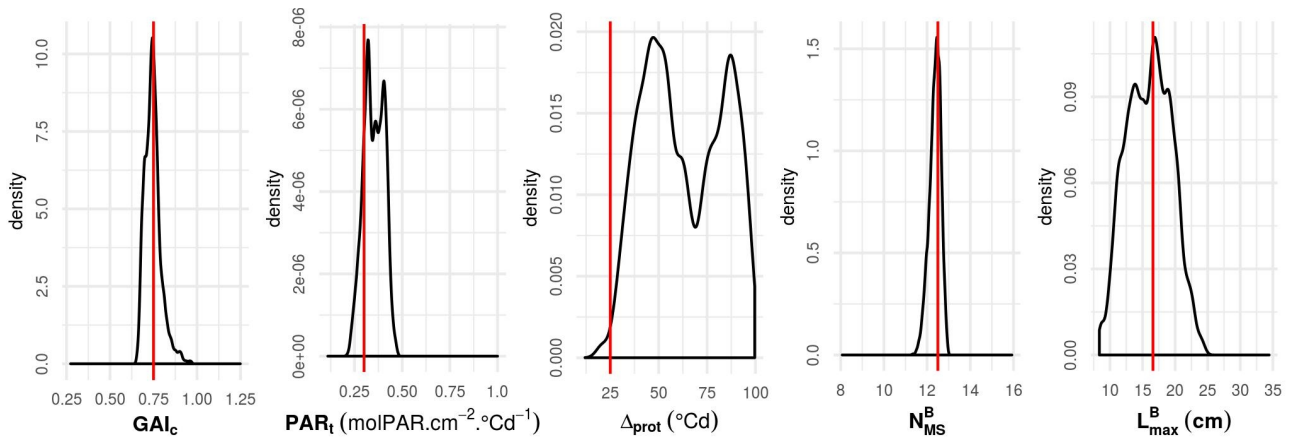


Fig. 4. Posterior distributions of WALTER parameters (GAI_c , PAR_t , Δ_{prot} , N_{MS}^B and L_{max}^B) using the enriched design after 15 loops of EGO and simulated data from density 200 plants/m². Parameter values used to simulate the data are shown with a red vertical line.

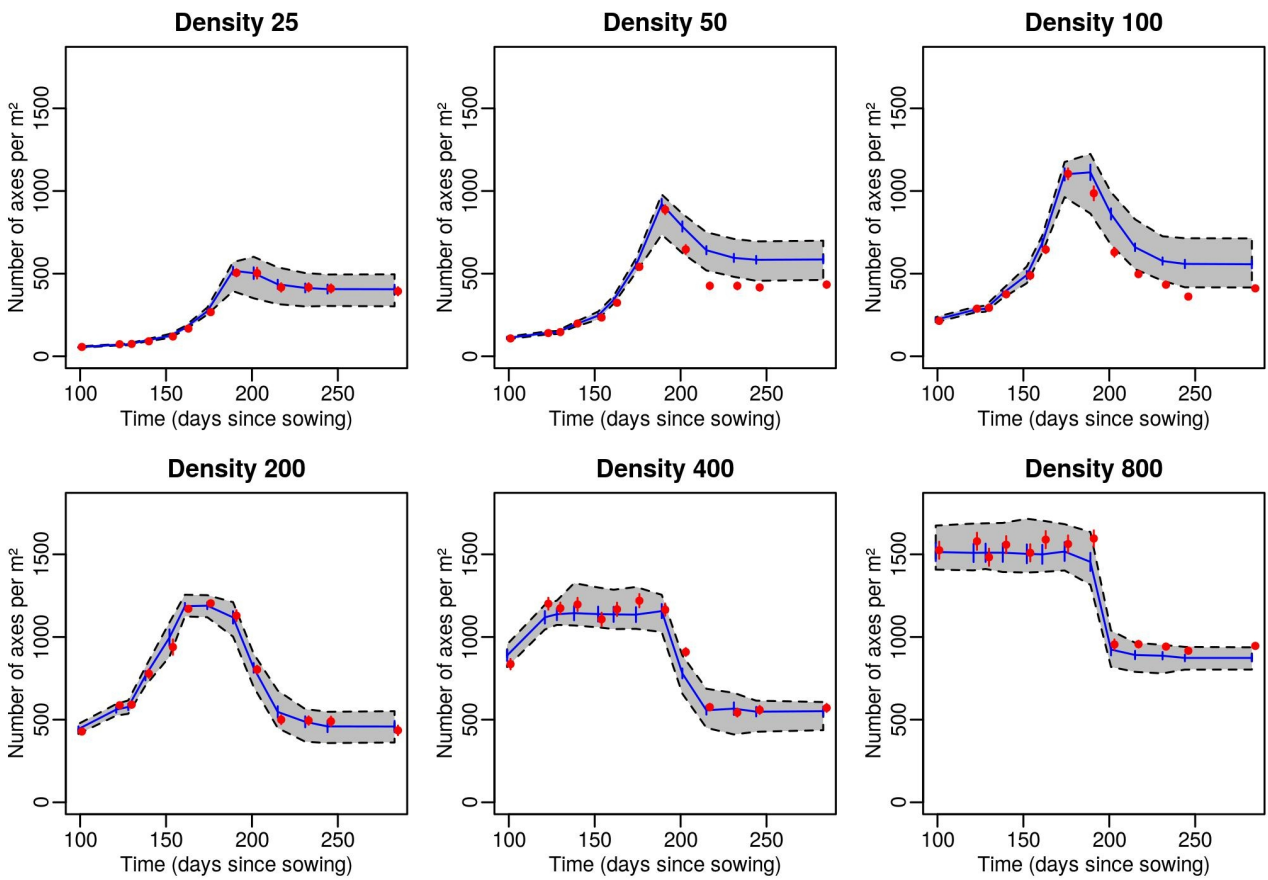


Fig 5. Number of axes per m² vs. time since sowing at six sowing densities (25, 50, 100, 200, 400 and 800 plants.m⁻²). Red dots represent the synthetic data and red segments are the associated observational variance ; blue lines represent the mean Kriging prediction for the set of input parameters with the best loglikelihood and blue segments are the associated Kriging standard deviation. The grey area delimited by dotted lines is the 95% credibility interval taking into account the Kriging standard deviation. Red points and segments are shifted on the x-axis to avoid overlapping. The set with the best loglikelihood was selected by MCMC using the enriched design after 15 loops of EGO and simulated data from density 200 plants/m².

However, even when using the design enriched with 15 simulations, the accuracy of the calibration is greatly deteriorated when it is carried out only with the data of a single sowing density (Fig 4). Indeed, the posterior distributions obtained using only the data at 200 plants/m² all include the actual parameter values but their variances are very large. This resulted in some of the simulated

data being outside the 95% credibility range of the Kriging metamodel when using this single density to select the 'optimal' set of parameters (Fig 5). However, the fit to the synthetic data still seems reasonably good, as the 95% credibility range includes the vast majority of the data.

Application to experimental data

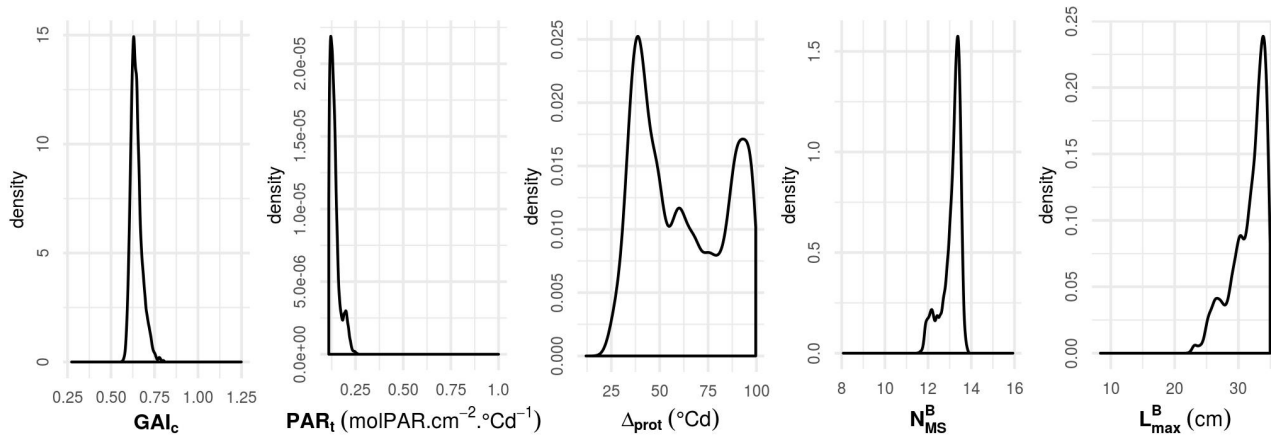


Fig. 6. Posterior distributions of WALTER parameters (GAI_c , PAR_t , Δ_{prot} , N_{MS}^B and L_{max}^B) using the experimental data from Darwinkel (1978) at all densities and the enriched design after 15 loops of EGO.

When the enriched design is used to calibrate WALTER on the experimental data from Darwinkel (1978), the posterior distributions (Fig 6) show contrasting results depending on the input parameter considered. The posterior density for N_{MS}^B and GAI_c , which are the 2 parameters with the most influence on the tillering dynamics according to the global sensitivity analysis (Chapter 3), only include intermediate values of the parameters and have a rather low variance. On the other hand, the posterior density for Δ_{prot} has a very large variance and includes almost all the range of values explored. As for the posterior density of PAR_t , its variance is low, but it only includes values that are close to the lower bound of the variation range. Similarly, the posterior density of L_{max}^B is clearly shifted towards the upper bound of the variation range, even though its variance is quite large.

With the 'optimal' set of parameters selected by MCMC, there is a rather good fit between the simulated tillering dynamics and the experimental dynamics of Darwinkel (1978) (Fig 7). However, some of the experimental data is outside the 95% credibility range of the Kriging metamodel, even when the observational variance is considered. In particular, for all sowing densities except 800 plants/m², the model fails to reproduce the number of axes per m² observed experimentally for the first 2 dates of the dynamics. Indeed, for these 2 dates, the number of axes per m² is always lower for the simulations than for the experimental data. For the highest sowing density (800 plants/m²), this discrepancy between simulated and experimental data is observed for the first date of the dynamics. However, for the second date of the dynamics, the experimental value is included in the 95% credibility range of the Kriging metamodel, when the observational variance is considered (red segments on Fig 7).

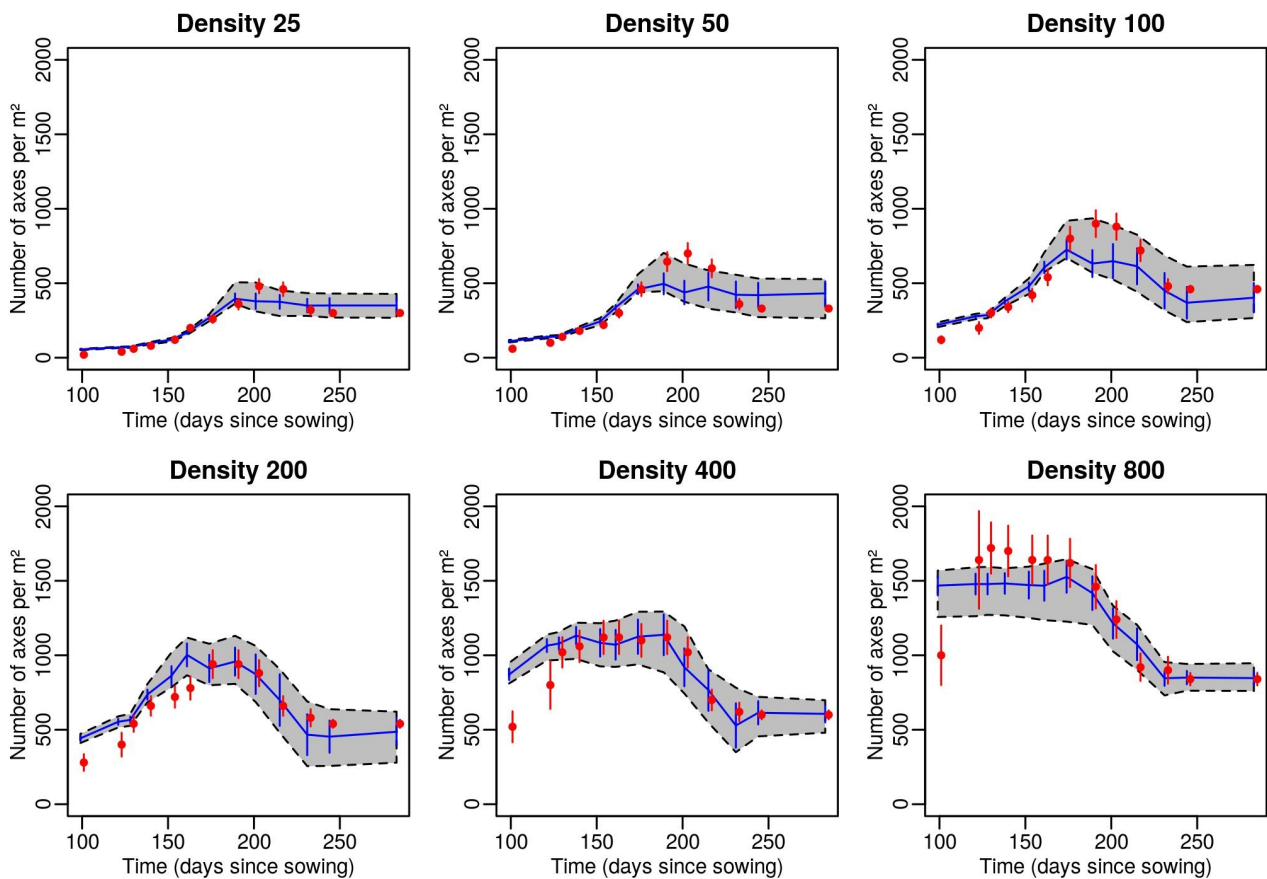


Fig 7. Number of axes per m^2 vs. time since sowing at six sowing densities (25, 50, 100, 200, 400 and 800 plants. m^{-2}). Red dots represent the experimental data from Darwinkel (1978) and red segments are the associated observational variance ; blue lines represent the mean Kriging prediction for the set of input parameters with the best loglikelihood and blue segments are the associated Kriging standard deviation. The grey area delimited by dotted lines is the 95% credibility interval taking into account the Kriging standard deviation. Red points and segments are shifted on the x-axis to avoid overlapping. The set with the best loglikelihood was selected by MCMC using the enriched design after 15 loops of EGO and experimental data of Darwinkel (1978) from all sowing densities.

DISCUSSION

The calibration of complex models, such as FSPM, involves several challenges, due in part to the high computational cost of the simulations, as well as the numerous parameters to consider. This study illustrates the interest of using a sequential design with a metamodeling approach to calibrate the WALTER FSPM for 5 critical parameters, based on data from the tillering dynamics. The method presented here performed well, especially when applied to synthetic data. The Kriging metamodel efficiently approximated the mean number of axes per m^2 simulated by WALTER for 13 dates in pure stands. This highlights the interest of using a Kriging metamodel with WALTER to limit the computational cost and thus allow to use methods that require a large number of runs, such as the MCMC. Importantly, a set of 100 simulations was not sufficient to ensure a quality of the metamodel that was high enough for the calibration. However, an enriched design with only 15 additional simulations (115 simulations total), obtained by the sequential method, allowed for a satisfactory fit of the metamodel outputs to both synthetic and experimental data. The use of an adaptive design thus proved to be a very efficient method to improve the quality of the metamodel, reducing the uncertainty in areas of the parameter space that are of interest for the fitting. Contrary to the manual calibration previously applied to WALTER (Lecarpentier *et al.*, 2019), the method

presented here provides the practitioner with a distribution of likely values for each parameter considered. Moreover, the Bayesian method presented allows for a more rigorous exploration of the parameter space than the previous manual calibration. However, the method presented here could be improved. For example, the choice of a maximin Latin Hypercube Sampling (LHS) as the initial design for the sequential method can be discussed. Indeed, [Zhang *et al.* \(2019\)](#) argue that other designs outperform maximin LHS in both static and sequential settings. Furthermore, the good performance of the method for the calibration on the experimental data of [Darwinkel \(1978\)](#) relies on assumptions regarding the observational variance. Indeed, the experimental dataset consisted only of the mean tillering dynamics of the plots and no variance was provided. The uncertainty associated with the experimental measurements was thus assumed to be rather high, but a different observational variance may have impacted the results of the fitting. It is also important to mention the potential divergence between the sequence of daily temperature used in this work and the real one, not available, that has an impact on the quality of the fitting and on its robustness. This highlights the importance of the quality of the data used for the fitting. For FSPM, parameter estimation often requires a lot of data at various scales and the issue of the limited availability of such appropriate datasets is a concern ([Louarn and Song, 2020](#)). Interestingly, our study also provides information regarding the type of data necessary for the calibration of WALTer. First of all, our results suggest that data on the tillering dynamics are sufficient to estimate the value of parameters controlling the regulation of tillering (GAI_c , PAR_t and Δ_{prot}), but also to estimate the final number of leaves on the main stem (N_{MS}^B) and the dimensions of the blades (L_{max}^B). Here, the value of these parameters were estimated with a set of 13 dates spread across the development period of the crop. However, since the parameters estimated here do not impact the rate of emission of the tillers, it is possible that the calibration could be done with a slightly lower measurement effort by reducing the number of dates at the beginning of the dynamics. Moreover, it is noteworthy that the final number of leaves on the main stem (N_{MS}^B) and the final length of the longest blade (L_{max}^B) can be measured experimentally. It might thus be possible to estimate the values of the parameters controlling the tillering dynamics (GAI_c , PAR_t and Δ_{prot}) with a dataset containing even fewer measurement dates, provided that N_{MS}^B and/or L_{max}^B are measured experimentally. In future work, we will explore these possibilities based on the synthetic data generated here, to identify the minimal dataset necessary to achieve a satisfactory calibration. Finally, our study has illustrated the importance of collecting data for several sowing densities for the calibration of WALTer, as the estimation of the parameters is deteriorated when it is based only on data collected at a 200 plants/m² sowing density.

The method presented here, based on an adaptive design and a Kriging metamodel, is an efficient approach for the calibration of WALTer and could be of interest for the calibration of other FSPM. Interestingly, the relatively low computational cost of this approach would make it possible to calibrate WALTer for a large number of genotypes, and the results of this study can help design the experiments needed to collect the necessary data for this task.

REFERENCES

- Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., ... & Tu, J. (2007).** A framework for validation of computer models. *Technometrics*, 49(2), 138-154.
- Bornhofen, S., Barot, S., & Lattaud, C. (2011).** The evolution of CSR life-history strategies in a plant model with explicit physiology and architecture. *Ecological Modelling*, 222(1), 1-10.
- Carmassi, M., Barbillon, P., Keller, M., Parent, E., & Chiodetti, M. (2019).** Bayesian calibration of a numerical code for prediction. *arXiv preprint arXiv:1801.01810*.
- Chelle, M., & Andrieu, B. (1998).** The nested radiosity model for the distribution of light within plant canopies. *Ecological Modelling*, 111(1), 75-91.
- Christensen, A. J., Srinivasan, V., Hart, J. C., & Marshall-Colon, A. (2018).** Use of computational modeling combined with advanced visualization to develop strategies for the design of crop ideotypes to address food security. *Nutrition reviews*, 76(5), 332-347.
- Currin, C., Mitchell, T., Morris, M., & Ylvisaker, D. (1991).** Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 86(416), 953-963.
- Damblin, G., Barbillon, P., Keller, M., Pasanisi, A., & Parent, É. (2018).** Adaptive numerical designs for the calibration of computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 6(1), 151-179.
- Darwinkel A. (1978).** Patterns of tillering and grains production of winter wheat at a wide range of plant densities. *Netherland Journal of Agricultural Science* 26: 383–398.
- Da Silva, D., Han, L., Faivre, R., & Costes, E. (2014).** Influence of the variation of geometrical and topological traits on light interception efficiency of apple trees: sensitivity analysis and metamodelling for ideotype definition. *Annals of botany*, 114(4), 739-752.
- Dupuy, D., Helbert, C., Franco, J. (2015).** DiceDesign and DiceEval: Two R Packages for Design and Analysis of Computer Experiments. *Journal of Statistical Software*, 65(11), 1-38. URL <http://www.jstatsoft.org/v65/i11/>.
- Evers, J. B., Letort, V., Renton, M., & Kang, M. (2018).** Computational botany: advancing plant science through functional–structural plant modelling. *Annals of Botany*, 121(5), 767–772.
- Fournier, C., Andrieu, B., Ljutovac, S., & Saint-Jean, S. (2003).** ADEL-Wheat: A 3D Architectural Model of wheat development. In *Plant Growth Modeling and Applications* (pp. 54-63). Springer Verlag.
- Gauthier, M., Barillot, R., Schneider, A., Chambon, C., Fournier, C., Pradal, C., ... & Andrieu, B. (2020).** A functional structural model of grass development based on metabolic regulations and coordination rules. *Journal of Experimental Botany*.
- Godin, C., Sinoquet, H. (2005).** Functional–structural plant modelling. *New phytologist*, 166(3), 705-708.

- Guo, Y., Fourcaud, T., Jaeger, M., Zhang, X., & Li, B. (2011).** Plant growth and architectural modelling and its applications. *Annals of botany*, 107(5), 723-727.
- He, D., Wang, E., Wang, J., & Robertson, M. J. (2017).** Data requirement for effective calibration of process-based crop models. *Agricultural and Forest Meteorology*, 234, 136-148.
- Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., & Ryne, R. D. (2004).** Combining field data and computer simulations for calibration and prediction. *SIAM Journal on Scientific Computing*, 26(2), 448-466.
- Johnson, M. E., Moore, L. M., & Ylvisaker, D. (1990).** Minimax and maximin distance designs. *Journal of statistical planning and inference*, 26(2), 131-148.
- Jones, D. R., Schonlau, M., & Welch, W. J. (1998).** Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4), 455-492.
- Kennedy, M. C., & O'Hagan, A. (2001).** Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), 425-464.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Louarn, G., & Song, Y. (2020).** Two decades of functional–structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany*, 126(4), 501-509.
- Martin, A. D., Quinn, K. M., & Park, J. H. (2011).** MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software*. 42(9): 1-21. URL <http://www.jstatsoft.org/v42/i09/>.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979).** Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239-245.
- Muleta, M. K., & Nicklow, J. W. (2005).** Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of hydrology*, 306(1-4), 127-145.
- Parsons, R. A., Mell, W. E., & McCauley, P. (2011).** Linking 3D spatial models of fuels and fire: effects of spatial heterogeneity on fire behavior. *Ecological Modelling*, 222(3), 679-691.
- Perez, R. P., Dauzat, J., Pallas, B., Lamour, J., Verley, P., Caliman, J. P., ... & Faivre, R. (2018).** Designing oil palm architectural ideotypes for optimal light interception and carbon assimilation through a sensitivity analysis of leaf traits. *Annals of botany*, 121(5), 909-926.
- R Core Team (2017).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Roustant, O., Ginsbourger, D., Deville, Y. (2012).** DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software*, 51(1), 1-55. URL <http://www.jstatsoft.org/v51/i01/>.

- Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989).** Design and analysis of computer experiments. *Statistical science*, 409-423.
- Sievänen, R., Godin, C., DeJong, T. M., & Nikinmaa, E. (2014).** Functional–structural plant models: a growing paradigm for plant studies. *Annals of botany*, 114(4), 599-603.
- Simpson, T. W., Poplinski, J. D., Koch, P. N., & Allen, J. K. (2001).** Metamodels for computer-based engineering design: survey and recommendations. *Engineering with computers*, 17(2), 129-150.
- Vidal, T., Gigot, C., de Vallavieille-Pope, C., Huber, L., & Saint-Jean, S. (2018).** Contrasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. *Annals of botany*, 121(7), 1299-1308.
- Villa-Vialaneix, N., Follador, M., Ratto, M., & Leip, A. (2012).** A comparison of eight metamodeling techniques for the simulation of N₂O fluxes and N leaching from corn crops. *Environmental Modelling & Software*, 34, 51-66.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- Wallach, D., Buis, S., Lecharpentier, P., Bourges, J., Clastre, P., Launay, M., ... & Justes, E. (2011).** A package of parameter estimation methods and implementation for the STICS crop-soil model. *Environmental Modelling & Software*, 26(4), 386-394.
- Zhang, B., Cole, D. A., & Gramacy, R. B. (2019).** Distance-distributed design for Gaussian process surrogates. *Technometrics*, 1-13.

CHAPITRE 5

Functional-structural plant modelling highlights how diversity in leaf dimensions and tillering capability could promote the efficiency of wheat cultivar mixtures

INTRODUCTION

Increasing the cultivated diversity has been identified as a major leverage for the agroecological transition as it can help improve the resilience of low input cropping systems (Malézieux, 2012 ; Isbell, 2015). At the scale of the plot, diversity can be increased by mixing species or by increasing the genetic diversity within a species (varietal mixture for example). For wheat, which is the most cultivated crop worldwide in terms of harvested area (FAO, 2018), the use of cultivar mixtures has been reported to present advantages for yield and quality as well as for diseases resistance, insect pests control, weed suppression, lodging limitation, exploitation of water and soil nutrients and yield stability (Borg *et al.*, 2018). Moreover, the use of wheat cultivar mixtures is spreading in several countries (Faraji, 2011). In France, for example, wheat blends represented less than 1% of the total wheat surface in 2010 but they represented more than 11% of the total wheat surface in 2019 (FranceAgriMer, 2019). To support the spread of cultivar mixtures, assembly rules have been developed, but they're almost exclusively focused on diseases resistance. Studies have seldom focused on establishing mixing rules based on the plants architecture and the few existing recommendations usually advocate the association of homogeneous cultivars, but the demonstration of these guidelines seems to be lacking (Borg *et al.*, 2018).

Yet, plant architecture and the overall canopy structure are critical for field performance, as they strongly influence light interception (Niinemets, 2010). Thus, many traits of the aerial architecture, such as height, leaf dimensions, leaf inclination and branching (*i.e.* tillering in cereals), have been identified for their impact on light interception. Indeed, leaf dimensions have a direct influence on the leaf area index (LAI) of the canopy, which is the most basic property affecting its light interception (Pugnaire and Valladares, 2007). Branching is also an important feature as it greatly impacts the total leaf area of the plant by changing its number of leaves (Valladares and Niinemets, 2008). On the other hand, plant height has a direct impact on the leaf area distribution in the canopy and on the overlap between neighbouring leaves, which is critical for competition for light, especially in heterogeneous canopies (Givnish, 1982). As for leaf inclination, it determines the penetration of light into the canopy: the more upright the leaves are, the deeper light can penetrate into the lower layers of the canopy (Long *et al.*, 2006). The impact of all of these traits has been considered for the definition of wheat ideotypes (*i.e.* ideal hypothetical plants for a given context) (Donald, 1968) and in breeding programs (Reynolds *et al.*, 2012), but mostly in the context of single plants or pure stands. However, the impact of these architectural traits would also be of interest for the establishment of assembly rules and the definition of mixture ideotypes (Litrico and Violle, 2015). To this end, it is necessary to better understand the interactions taking place in mixtures and the role played by architectural features in these interactions.

The very high number of trait combinations in wheat mixtures makes it difficult to conduct experimentations on this issue, which is why a modelling approach appears to be an appropriate solution. In particular, functional structural plant models (FSPM, Godin & Sinoquet, 2005 ; Vos *et al.*, 2010) have the advantage of being individual-based, allowing for the explicit consideration of inter-individual variations in the stand (Zhang and De Angelis, 2020). Most importantly, FSPM explicitly represent the 3D architecture of plants and its interactions with the ecophysiological processes controlling plant development. These models are thus particularly adapted to study the

impact of architectural traits on the performance of heterogeneous canopies (Evers *et al.*, 2019 ; Gaudio *et al.*, 2019). If past studies using FSPM were mainly focusing on sole crops (Sarlikioti *et al.* 2011 ; Chen *et al.*, 2014 ; Da Silva *et al.*, 2014 ; Streit *et al.*, 2016 ; Perez *et al.*, 2018), more and more work is now addressing mixtures of different species (Barillot *et al.*, 2014 ; Munz *et al.*, 2018 ; Louarn *et al.*, 2020) or cultivars (Barillot *et al.*, 2019). However, no FSPM study has yet focused on the impact of more than one trait on the performance of wheat cultivar mixtures.

In this study, we used an FSPM to simulate balanced binary wheat cultivar mixtures differing in four traits of their aerial architecture. Using a sensitivity analysis approach, we were able to identify the major traits impacting the performance of the simulated stands. This study improves our understanding of how differences between cultivars in key traits of the aerial architecture influence mixture performance.

MATERIAL AND METHOD

All the statistical analyses were done using the R software (R Core Team, 2018).

WALTer: a 3D wheat model

WALTer (Lecarpentier *et al.*, 2019) is an FSPM that simulates the development of the aerial architecture of winter wheat (*Triticum aestivum* L.) from sowing to maturity with a daily time step. In the model, the vegetative development of the plants follows a thermal time schedule and is based on a formalism derived from ADEL-Wheat (Fournier *et al.*, 2003). The geometry, size, developmental kinetics and the senescence of leaves and internodes are simulated thanks to deterministic equations. However, WALTer is not a completely deterministic model, as the position of the plants, the orientation of the organs, the duration before plant emergence and the emergence of tillers are partly random.

Most importantly, WALTer simulates the competition for light between plants within the field and the resulting plasticity of tillering (i.e. the branching ability of grasses), thanks to a radiative model (CARIBU: Chelle and Andrieu, 1998). The regulation of tillering in the model is based on three simple rules. (i) Empirically fixed probabilities control the emergence of new tillers. Then, (ii) the cessation of tillering is controlled by an early neighbour perception : plants stop emitting new tillers when the surrounding Green Area Index^a reaches a critical value (GAI_c). Finally, (iii) some of the tillers that were emitted regress: a tiller regresses if the amount of photosynthetically active radiation (PAR) it intercepts per unit area falls below a threshold (PAR_t).

Based on this formalism, WALTer produces meaningful outputs, such as the tillering dynamics of each plant (i.e. the number of axes on a plant for each day of the simulation) and the dynamics of light interception for each organ in the stand.

Since its publication in 2019 (Lecarpentier *et al.*), WALTer has undergone some changes aiming at reducing the computational cost of the simulations, enhancing its realism and improving its ability to simulate mixtures of varieties. These modifications of the model are detailed in Chapter 1.

a The GAI is the ratio of photosynthetic surface of a plant to ground area.

Sensitivity analysis

The impact of different architectural traits on the performance of wheat mixtures was evaluated through the simulation of balanced binary mixtures (mixtures of 2 varieties with 50% of each cultivar). To evaluate and hierarchise the influence of the different traits on the performance of the simulated mixtures, a sensitivity analysis was conducted with the variance-based Sobol method (Sobol, 1993). Because of the relatively high computational cost to run WALTer, a metamodelling approach was chosen to enable the computation of Sobol indices. The metamodel consists in a fast approximation of WALTer constructed from a limited number of model runs. We used a Gaussian process metamodel (Currin *et al.*, 1991 ; Sacks *et al.*, 1989), also known as Kriging metamodel, as in Marrel *et al.*, 2009.

Model outputs

Three key outputs of the model were selected for the analyses to reflect the performance of each simulated plot. First, the mean number of ears per plant at maturity (N_{ears}), which is a key component of the plot's yield, was selected. Second, the proportion of incident photosynthetically active radiation (PAR) intercepted by the stand (L_{perc}), was also taken into account. For each stand, the daily L_{perc} was averaged over the whole period of use of the radiative model (from around 100 days after sowing to the maturity of the stand). Third, the total fitness of the plot (F_{tot}) was used for the analyses. This last output is used as a proxy of the number of kernels produced by the plot and is computed by WALTer as the sum of the fitness of all the axes in the plot. The computation of the fitness of an axis is based on the following equation:

$$F_{\text{axis}} = \frac{PAR_{i45}}{T_{m45}} \quad (\text{Equation 1})$$

where F_{axis} is the fitness of the axis (expressed in molPAR.°Cd⁻¹) ; PAR_{i45} is the mean daily amount of PAR intercepted by the axis during the 45 days preceding flowering (in molPAR.d⁻¹) ; and T_{m45} is the mean temperature during the same period (in °C).

This equation is an adaptation of the photothermal quotient (Nix, 1976), which has been demonstrated to be linearly correlated to the number of grains produced in wheat (Midmore *et al.*, 1984 ; Fischer, 1985 ; Abbate *et al.*, 1995 ; Demotes-Mainard and Jeuffroy, 2001).

To allow the comparison of the performance of a mixture with the performance of its pure components, the overyielding was also computed for each of these three outputs, using the following equation:

$$OY_{\text{out}} = \frac{\text{out}_{\text{mix}}}{\frac{\text{out}_{\text{pure1}} + \text{out}_{\text{pure2}}}{2}} \quad (\text{Equation 2})$$

where OY_{out} is the overyielding for the considered output (N_{ears} , L_{perc} or F_{tot}) ; out_{mix} is the value of the considered output for the mixture ; and $\text{out}_{\text{pure1}}$ and $\text{out}_{\text{pure2}}$ are the values of the considered output for each of the 2 components of the mixture, when cultivated in pure stands.

Input factors

Five parameters associated with 4 key traits were considered (Table 1). For each of the 5 parameters, the impact of the differential between the 2 varieties in the mixture was considered, as well as the impact of the mean value in the plot. As a result, there were 2 input factors considered for each parameter, for a total of 10 input factors for the sensitivity analysis. The ranges of variation for all the input factors were selected to generate biologically realistic varieties with contrasted architectures. For example, the selected ranges allowed for the simulation of varieties with a final height between 40 and 160 cm (Table 1).

Table 1. List of parameters for the sensitivity analysis, definitions, traits impacted by the parameter, ranges of variation explored for the mean parameter value in the stand (Ref range) and for the difference of trait value between the two varieties of the mixture (Diff range), corresponding minimum and maximum values at the plant scale and units.

Parameter	Description	Trait	Ref range	Diff range	Min value	Max value	Unit
H_{MS}	Final height of the main stem	Plant height	65 ; 135	-50 ; 50	40	160	cm
φ_B	Insertion angle between the blade and the stem	Leaf inclination	30 ; 50	-40 ; 40	10	70	degree
L_{max}^B	Final length of the longest blade of the main stem ^b	Leaf dimensions	16 ; 27	-16 ; +16	8	35	cm
GAI_c	Green Area Index threshold above which the emission of tillers stops	Tillering	0.4 ; 0.7	-0.6 ; 0.6	0.1	1	-
PAR_t	PAR threshold below which a tiller does not survive	Tillering	0.3 ; 0.5	-0.2 ; 0.2	0.2	0.6	mol.cm ⁻² .°Cd ⁻¹

Initial design

To build the Kriging metamodel, we used a set of WALTER runs following a 10 dimensions maximin latin hypercube sampling (LHS ; McKay *et al.*, 1979 ; Johnson *et al.*, 1990) of 5000 simulations generated with the DiceDesign R package (Dupuy *et al.*, 2015) with the ranges detailed in Table 1. For each output (N_{ears} , L_{perc} and F_{tot}), a Kriging approximation was generated using the DiceKriging R package (Roustant *et al.*, 2012). The sensitivity R package (Iooss *et al.*, 2017) was then used to estimate the Sobol indices. To be able to compute the overyielding for every mixture, another set of WALTER runs was generated from a 5 dimensions maximin LHS of 1000 simulations. This second LHS only includes simulations of pure stands (no trait differential in the stand) and the ranges explored for each input parameter corresponded to the minimum and maximum values detailed in Table 1. Based on this second set of simulations, a Kriging model was constructed independently for each of the three outputs (N_{ears} , L_{perc} and F_{tot}). Thus, the first design allowed the approximation of the behaviour of mixtures while the second design allowed the approximation of the behaviour of the corresponding pure components. Using the Kriging metamodels from both designs, it was possible to compute sensitivity indices for the overyielding of each output.

^b L_{max}^B also has an indirect impact on the final length of the other blades and on their width. Thus, at the scale of the cultivar, the higher the value of L_{max}^B , the greater the blade area. See Appendix 3 for details.

The simulated plots were composed of 110 plants sown at a density of 200 plants/m². The plants were arranged in a grid pattern with plants equidistant from each other and the different cultivars were distributed randomly in the plot. Apart from the input factors of the sensitivity analysis described above and in Table 1, all genotypic parameters were identical for both varieties in the mixture and had values given in [Lecarpentier et al. \(2019\)](#) for the cultivar Maxwell. The climatic sequence used for the simulations was obtained by averaging the climatic sequences of 5 French locations over 10 years (from 2007 to 2017). The 5 locations listed in Appendix 2 were selected to represent contrasted climatic conditions. For the computation of light interception, only diffuse radiations were considered, according to the standard overcast sky ([Moon and Spencer, 1942](#)). Diffuse radiation was approximated using a set of 16 light sources according to a TURTLE sky ([den Dulk, 1989](#)).

Optimisation

From the Kriging metamodels, it is possible to seek for combinations of architectural traits maximising the performance of the stand. However, the uncertainty of the metamodels can affect the results. Therefore, we decided to enrich the initial LHS design with new runs of WALTer, based on an Efficient Global Optimization (EGO) algorithm ([Jones et al., 1998](#)). We decided to focus on the maximization of F_{tot} , as this output encompasses information about both N_{ears} and L_{perc} . We focused on the exploration of combinations of $L_{max_diff}^B$, $H_{MS-diff}$, and PAR_t-ref , which are the 3 input factors with the greatest influence on F_{tot} . These 3 input factors were identified from the sensitivity analysis, because their total sensitivity index for F_{tot} were greater than 0.1, *i.e* each of these input factors explained more than 10% of the total variance of F_{tot} (cf. following results). The other input factors were set to the mean value of their range of variation.

The initial LHS design of 5000 simulations was iteratively enriched with a new run of WALTer until 50 new runs were added. At each iteration, the parameters' values for the new run were selected based on the Expected Improvement (EI) criterion among a set of 10000 parameter combinations sampled from a 3 dimensions latin hypercube. The EI criterion was computed with the DiceOptim R package ([Roustant et al., 2012](#)). The EI criterion is balancing the need to maximize F_{tot} with the need to reduce the uncertainty of the Kriging metamodel.

RESULTS

Outputs from the 5000 initial simulations are highly variable (Appendix 3), which indicates that at least some of the parameters varying in the LHS design have a strong impact on the performance of the simulated stands. However, some outputs are much more variable than others. Thus N_{ears} is the most variable output with values ranging from 1.0 to 4.3 ears per plant, even though 98% of the simulations resulted in N_{ears} values below 3.2 ears per plant. On the other hand, L_{perc} is much less variable, with values ranging from 0.76 to 0.96. As for F_{tot} , the initial simulations gave results ranging from 0.73 to 1.5 mol.°Cd⁻¹. For the OY indicators, simulations gave values between 0.40 and 1.6 for N_{ears} , and values between 0.67 and 1.9 for F_{tot} . Finally, the OY for L_{perc} was less variable, with values ranging from 0.97 to 1.4.

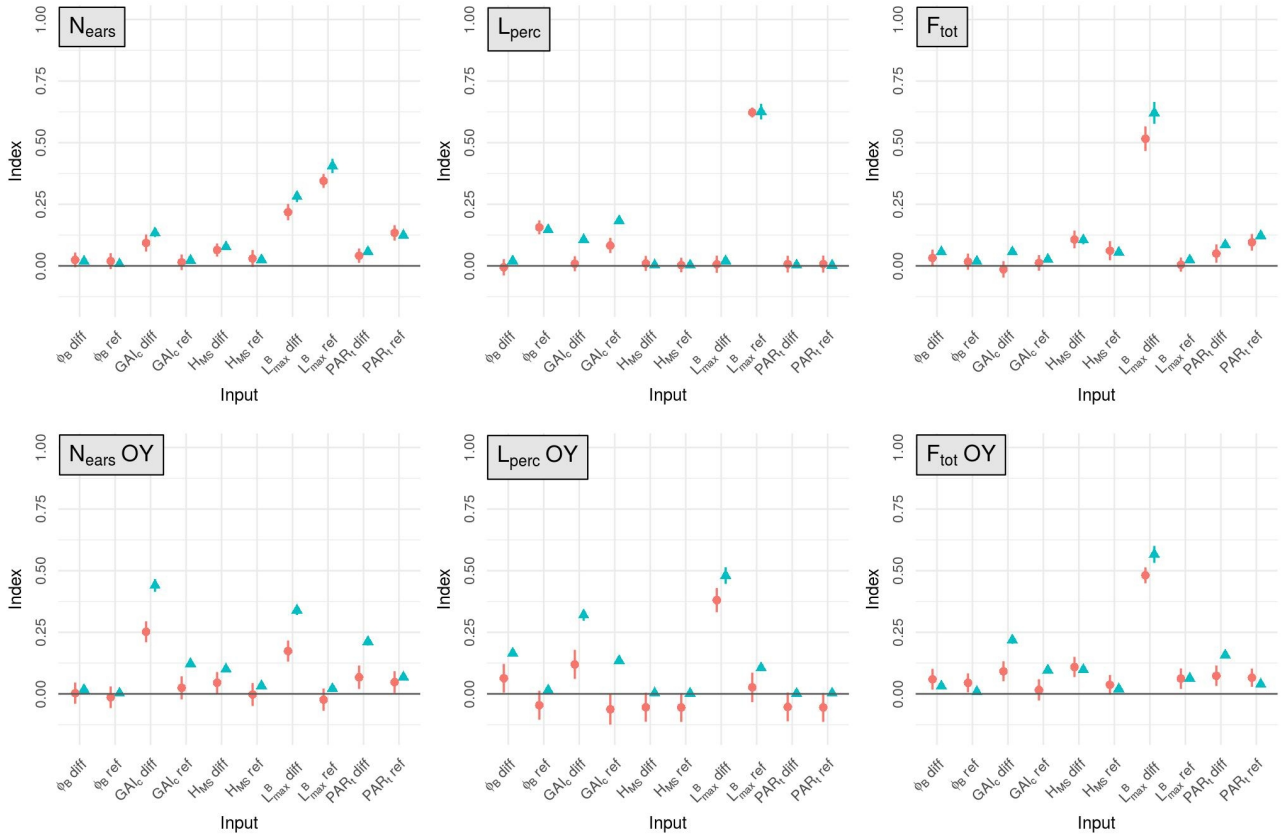


Fig. 1. Mean estimates for the total-order (blue triangles) and first-order (red circles) Sobol indices computed for N_{ears} , L_{perc} , F_{tot} and the OY of these three outputs, for the 10 input factors (x-axis). The standard-deviation of each index is represented with a vertical line.

The first-order sensitivity indices (MSI) provide information on the mean influence of the different inputs on each output and the total-order indices (TSI) aggregate mean influence and influence through interactions. The most important architectural traits regarding the performance of the stands are dependent on the considered output (Fig. 1.).

The mean number of ears per plant (N_{ears}) is mainly influenced by first-order effects. For this output the dimensions of the leaves is the most influential trait in the simulated plots as it explains more than 68% of the total variance of N_{ears} (taking into account the effect of the interactions). The mean value of L_{max}^B in the stand has more impact on N_{ears} than the difference of L_{max}^B between varieties (TSI = 0.41 and 0.28 respectively). The tillering capability of the plants in the stand also has an important impact on N_{ears} : $\text{GAI}_c\text{-diff}$ explains around 13% of the total variance of the output and $\text{PAR}_t\text{-ref}$ accounts for about 12%.

For L_{perc} , as for N_{ears} , the input factor with the most influence is $L_{\text{max}}^B\text{-ref}$. Indeed, the variations of $L_{\text{max}}^B\text{-ref}$ explain almost 63% of the total variance of L_{perc} . GAI_c is also an important parameter for light interception with $\text{GAI}_c\text{-ref}$ accounting for about 18% of the total variance of L_{perc} and $\text{GAI}_c\text{-diff}$ explaining more than 10% of its total variance. However, the impact of GAI_c on L_{perc} is mainly explained by interaction effects. Finally, the mean leaf insertion angle in the canopy has an important impact on L_{perc} with a total-order sensitivity index of 0.15 for $\phi_B\text{-ref}$. All the other input factors have very little influence on the light interception efficiency.

The input factor with the most impact on F_{tot} is $L_{\text{max}}^B\text{-diff}$ (TSI = 0.62). The mean value of PAR_t

in the mixture ($PAR_t\text{-ref}$) and the height differential between varieties ($H_{MS}\text{-diff}$) also have a lot of influence on F_{tot} as they explain 12% and 10% of the total variance of this output respectively.

For the OY outputs, the 2 most influential input factors are always $L_{max}^B\text{-diff}$ and $GAI_c\text{-diff}$, regardless of the output considered. Together, these 2 inputs always explain more than 77% of each OY output. Interestingly, interactions between input factors have an important impact on these outputs, particularly for the OY of N_{ears} and L_{perc} . Nevertheless, differences exist between the different OY outputs. For example, for the OY of N_{ears} , $GAI_c\text{-diff}$ has more impact than $L_{max}^B\text{-diff}$, whereas it is $L_{max}^B\text{-diff}$ that has the most influence on the OY of L_{perc} and F_{tot} .

To a lesser extent, other input factors also have a substantial influence on the OY outputs. $PAR_t\text{-diff}$, $GAI_c\text{-ref}$ and $H_{MS}\text{-diff}$ each explain more than 10% of the total variance of the OY of N_{ears} . The TSI of $PAR_t\text{-diff}$ for the OY of N_{ears} even reaches 0.21, but this effect is mainly explained by interactions as its MSI is only around 0.067. For the OY of L_{perc} , $GAI_c\text{-ref}$, $\varphi_B\text{-diff}$ and $L_{max}^B\text{-ref}$ are important factors, each with a TSI greater than 0.1. Finally, $PAR_t\text{-diff}$ explains almost 16% of the total variance of the OY of F_{tot} .

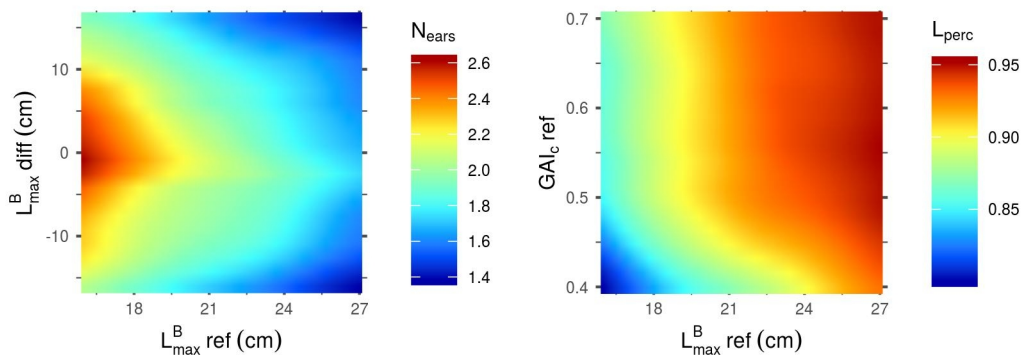


Fig. 2. Mean outputs predicted by the Kriging metamodells as a function of their 2 most influential input parameters. Predictions were made with all other parameters set to their mean values. Left: N_{ears} as a function of $L_{max}^B\text{-diff}$ and $L_{max}^B\text{-ref}$. Right: L_{perc} as a function of $GAI_c\text{-ref}$ and $L_{max}^B\text{-ref}$.

The sensitivity indices (Fig. 1.) only provide information about the relative importance of the different input parameters for the determination of the performance of the stand. Heatmaps representing the outputs of the Kriging metamodells as a function of their most influential parameters (Fig. 2, 3 and 4) give more information on how each parameter affects each output. For these figures however, it is important to keep in mind that the less influential parameters were set to the mean value of their range of variation, thus discarding most of the interaction effect that they might have with the other inputs. Another critical point when reading these heatmaps is the fact that, when the input variable is a parameter difference, it generates a plot that is symmetric on the axis $diff = 0$. And if the heatmap is a function of 2 parameters-diff, there is a central symmetry of the plot.

Figure 2 shows that the mean value of L_{max}^B in the stand ($L_{max}^B\text{-ref}$) is negatively correlated to N_{ears} , meaning that canopies with small blades tend to produce more spikes than stands with larger leaves. This heatmap also reveals that a differential of L_{max}^B between the cultivars of a mixture is

unfavourable for the ear yield in the simulated plots. Unlike N_{ears} , L_{perc} is positively correlated to $L_{\text{max}}^B - \text{ref}$: larger leaves allow for a higher light interception efficiency. L_{perc} is also positively correlated to $\text{GAI}_c - \text{ref}$, but mainly for values of $\text{GAI}_c - \text{ref}$ lower than 0.5 : above this value, the impact of $\text{GAI}_c - \text{ref}$ on L_{perc} is very low. This means that L_{perc} is higher for stands in which plants have a low sensitivity to the early competition (high GAI_c values) and can thus emit many tillers.

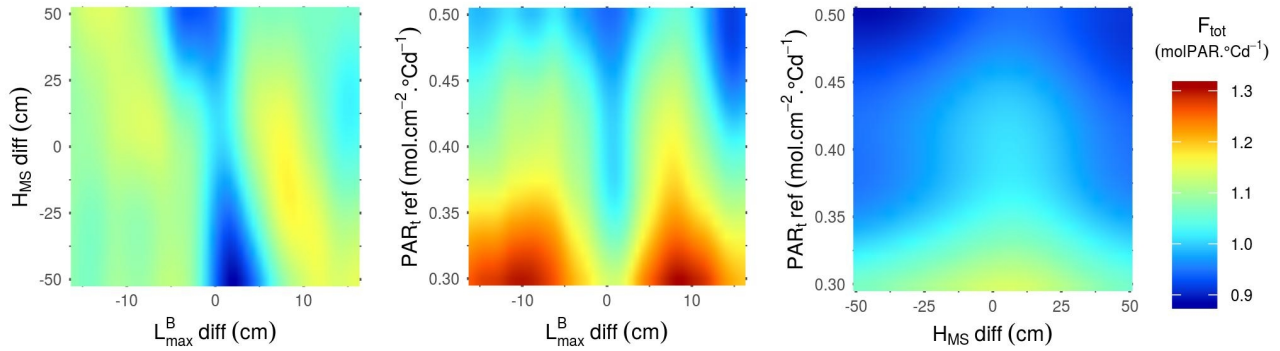


Fig. 3. F_{tot} predicted by the Kriging metamodel as a function of the 3 parameters that have the most impact on the variance of this output ($L_{\text{max}}^B - \text{diff}$, $H_{\text{MS}} - \text{diff}$ and $\text{PAR}_t - \text{ref}$). Predictions were made with all other parameters set to their mean values.

As for F_{tot} (Fig. 3.), unlike the results observed for N_{ears} (Fig. 2.), a differential of L_{max}^B between the cultivars of a mixture is favourable for the performance of the stand. An absolute differential of around 10 cm between the 2 varieties in the mixture leads to the highest values of F_{tot} . Moreover, the highest values of F_{tot} can only be reached with low values of $\text{PAR}_t - \text{ref}$. Indeed, there is a clear negative correlation between $\text{PAR}_t - \text{ref}$ and F_{tot} , which means that stands tend to be more productive when plants in the canopy do not need much light to prevent the regression of their tillers. Finally, even though the impact of $H_{\text{MS}} - \text{diff}$ on F_{tot} is low for the combinations of parameters presented in Fig.3., the results show that homogeneous plant height in the canopies lead to higher values of F_{tot} than the mixture of cultivars with a height differential. Interestingly, for mixtures of varieties with a height differential, F_{tot} is higher if the tallest variety is the one with the smallest leaves.

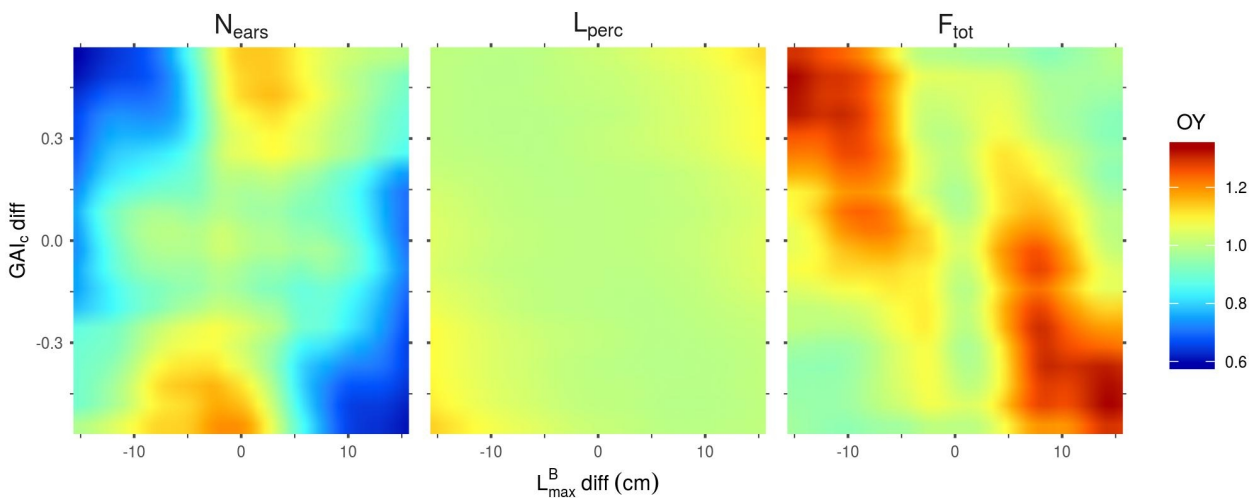


Fig. 4. Mean overyielding (OY) predicted by the Kriging metamodels for N_{ears} , L_{perc} and F_{tot} as a function of $\text{GAI}_c - \text{diff}$ and $L_{\text{max}}^B - \text{diff}$. Predictions were made with all other parameters set to their mean values.

For the OY outputs, L_{max}^B-diff and GAI_c-diff are always the most influential inputs (Fig. 1.). However, the effect of these 2 inputs, when the other inputs are fixed to their mean values, are very different depending on the output considered (N_{ears} , L_{perc} or F_{tot}) (Fig. 4.). Firstly, the combinations of parameters presented in Figure 4 lead almost exclusively to OY values greater than or equal to 1 for F_{tot} . This means that, for these combinations of parameters, mixtures are almost always at least as performant as the mean of their pure components. Conversely, OY values for N_{ears} may be as low as 0.6 and are rarely greater than 1. Finally, for L_{perc} , the OY is mostly around 1, with very little variation in its value. For this output, the highest values of OY are obtained for the extreme values of L_{max}^B-diff and GAI_c-diff by mixing plants with large leaves and high values of GAI_c with plants with small leaves and low GAI_c values. For N_{ears} , the highest values of OY are also obtained for the extreme values of GAI_c-diff , but for values of L_{max}^B-diff close to 0. For mixtures with a differential in both L_{max}^B and GAI_c , the OY of N_{ears} is more important when the cultivar with the largest leaves is also the one with the highest value of GAI_c , as is the case for the OY of L_{perc} . In contrast, the OY of F_{tot} is the highest when mixing plants with large leaves and low GAI_c values with plants with small leaves and high GAI_c values. Interestingly, despite the observed differences between the combinations of parameters leading to the highest OY for N_{ears} and F_{tot} , N_{ears} is positively correlated to F_{tot} in the overall dataset (data not shown). Finally, it is worth mentioning that the metamodels predict an OY of 1 for mixtures with no trait differential between cultivars (Fig. 4.), which means that they perform exactly like the corresponding pure stands. This is expected as all the plants in such canopies have identical parameter values. Thus, this result highlights the consistency of the metamodels' predictions.

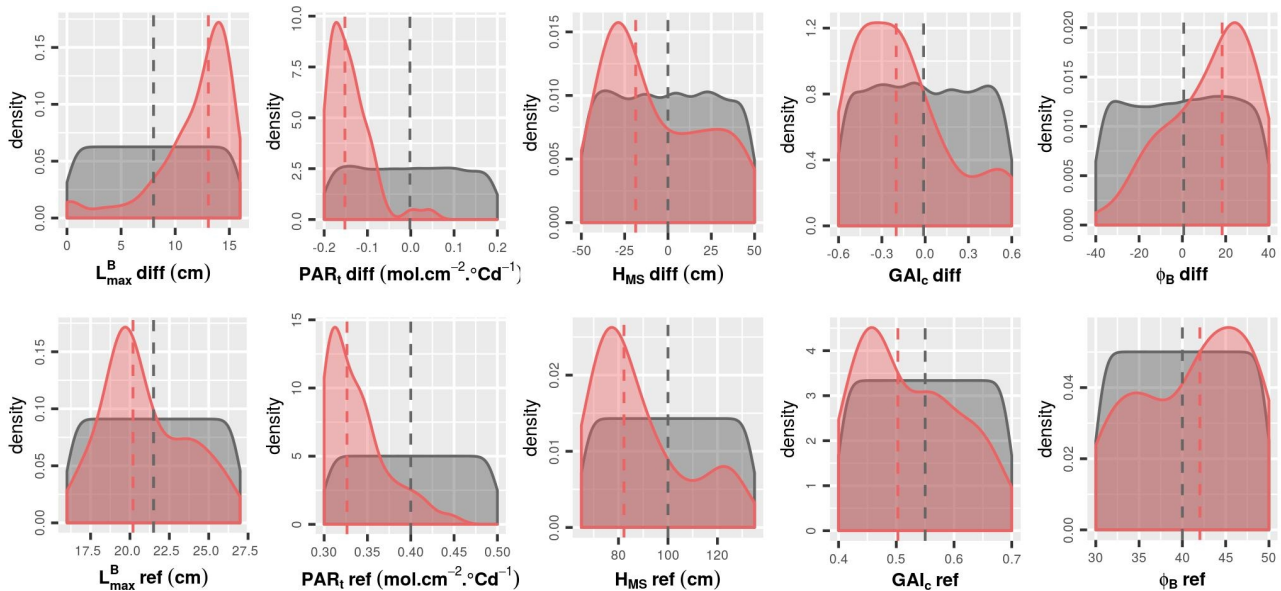


Fig. 5. Density of each input factor for the initial LHS (grey) and for the 50 runs with the highest F_{tot} values (red). The medians for each group is represented by a dashed line. To avoid redundancy in the combinations of parameters, results are represented only with positive values of L_{max}^B-diff .

In the initial design, different combinations of parameters can describe identical mixtures (as seen in the heatmaps). For example, considering only ϕ_B-diff and $H_{MS}-diff$, a stand with the minimum values of these 2 parameters is identical to a stand with the maximum values of the 2 parameters.

Indeed both combinations of parameters describe a mixture of tall planophile plants with short erectophile plants. To avoid this redundancy and simplify the analysis of the results, Figures 5, 6 and 8 have been represented only with positive values of $L_{max}^B - diff$ (all combinations are considered, but the ones with negative values of $L_{max}^B - diff$ have been reversed).

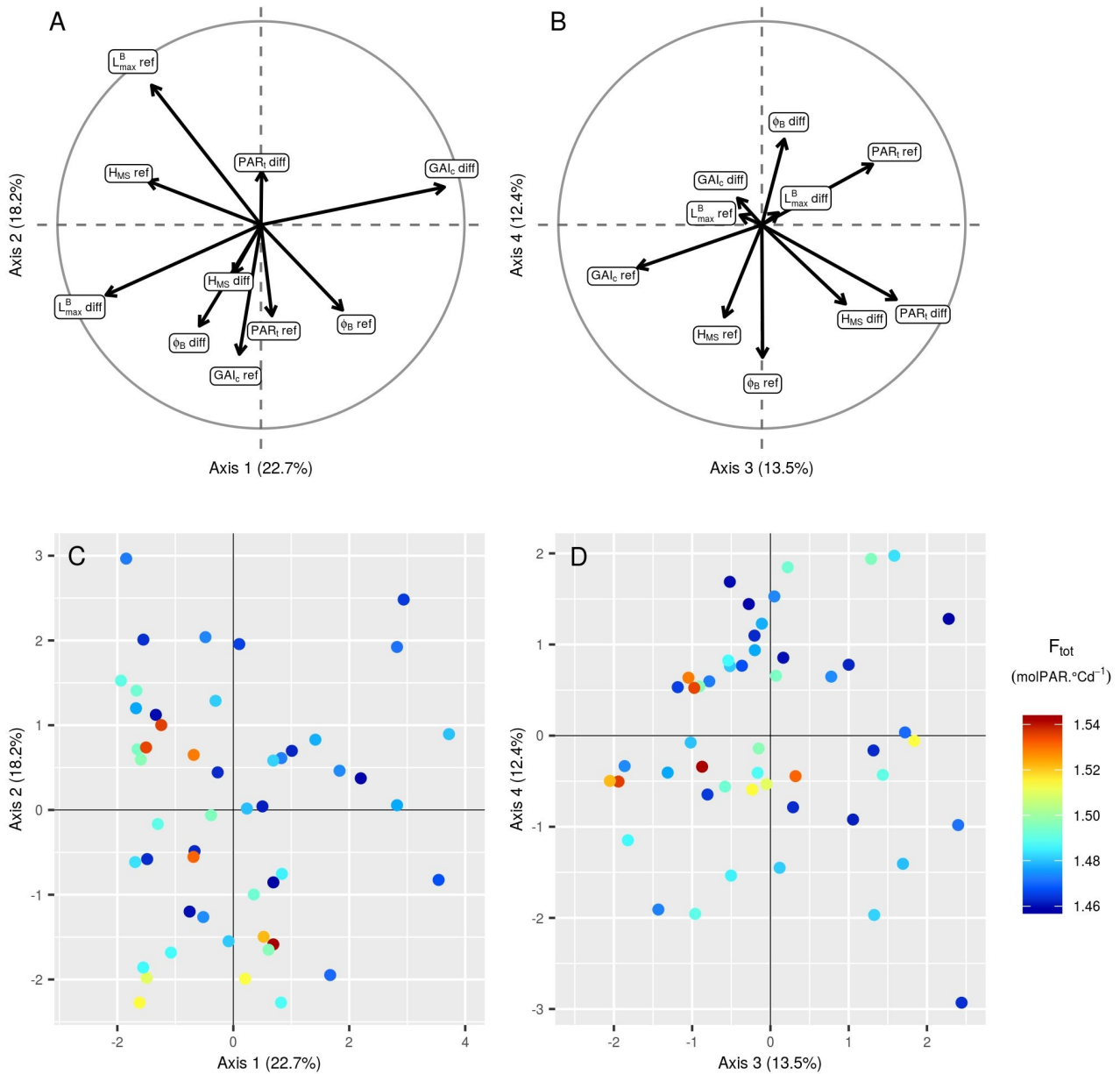


Fig. 6. Principal components analysis (PCA) of the 10 input parameters for the 50 simulations of the initial LHS with the highest F_{tot} . Projection of the 10 input parameters on the first two axes (A) and on the third and fourth axes of the PCA (B). Projection of the 50 points on the first two axes of the PCA (C) and on the third and fourth axes of the PCA (D). The points are coloured according to the corresponding value of F_{tot} . To avoid redundancy in the combinations of parameters, results are represented only with positive values of $L_{max}^B - diff$.

Focusing on the 50 runs of the initial design with the highest F_{tot} values (Fig. 5 and 6), it is possible to identify the trait combinations leading to the best performances. Simulations with high F_{tot} values usually combine high L_{max}^B-diff values with low (negative) values of PAR_t-diff . This combination of parameters describes mixtures in which the genotype with the largest leaves is the one with the lowest value of PAR_t , *i.e.*, the one less susceptible to tiller regression. However, there is a negative correlation between L_{max}^B-diff and GAI_c-diff (Fig. 6.), which means that it is preferable for the performance of the stand that the genotype with the largest leaves ceases tillering early in response to the neighbouring competition. Furthermore, mixtures with a differential of leaf insertion angle are more likely to have high F_{tot} values if the genotype with the most erected leaves is the one with the shortest leaves. Finally, the mean values of H_{MS} and PAR_t in the stand should be rather low to allow for high F_{tot} values. This means that canopies of short plants with low tiller regression tend to be the most performant ones.

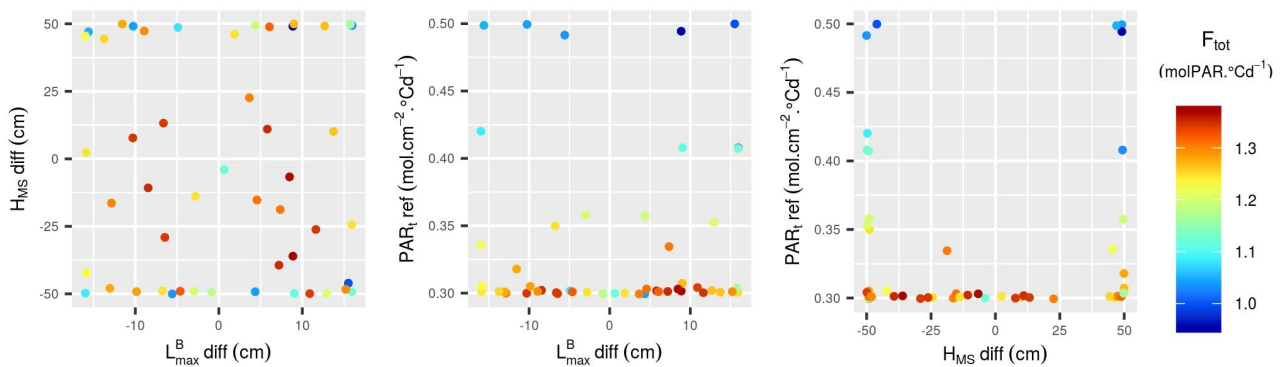


Fig. 7. Input parameter values for the 50 points added to the design via EGO. The points are coloured according to the corresponding value of F_{tot} . The additional points were selected in the whole range of variation for L_{max}^B-diff , $H_{MS}-diff$ and PAR_t-ref . The other input parameters were set to their mean values.

The addition of new runs of WALTER to the initial LHS design has confirmed the first results. In particular, for the new simulations as for the ones in the initial design, the highest values of F_{tot} are obtained for low values of PAR_t-ref (low tiller regression) and for mixtures of cultivars with different leaf dimensions (Fig. 7 and 8). Interestingly, the highest values of F_{tot} for the additional simulations are significantly lower than the ones for the initial design. This can be explained by the fact that the new simulations only explored combinations of L_{max}^B-diff , $H_{MS}-diff$ and PAR_t-ref and had the other parameters fixed to their mean values. The importance of the value of other parameters is thus highlighted by this result.

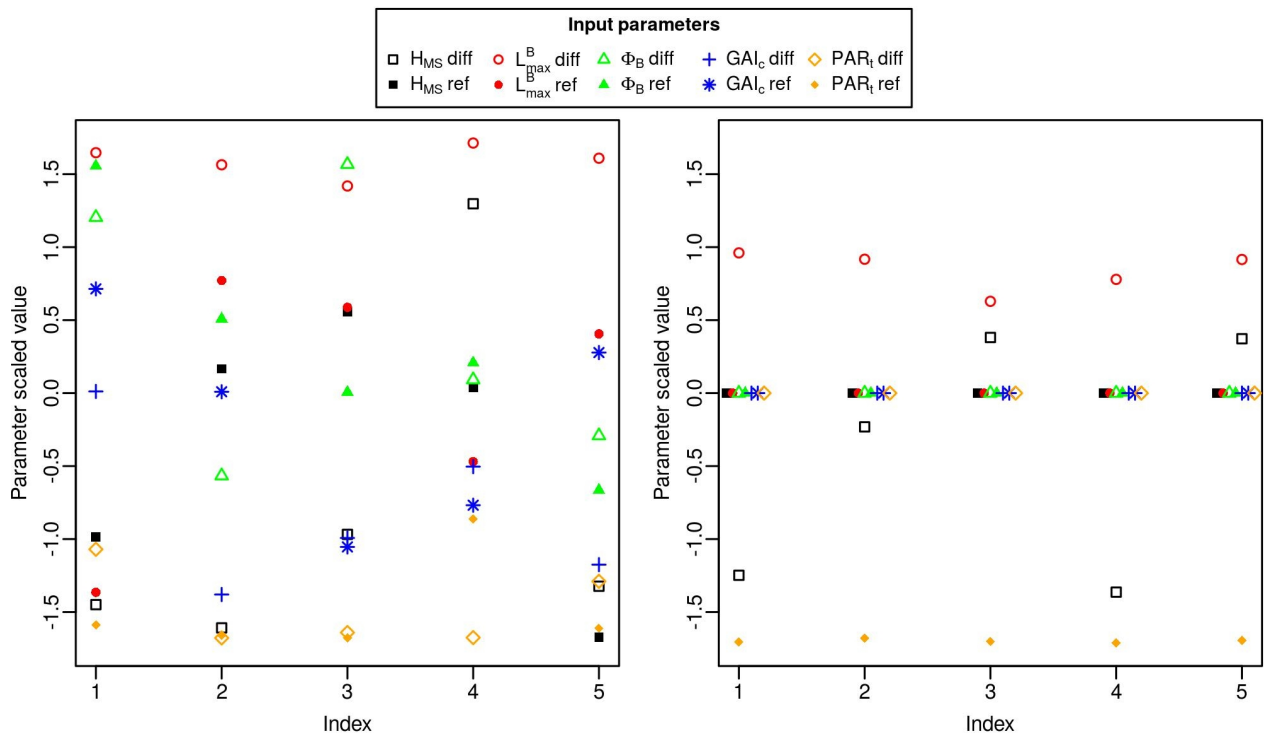


Fig. 8. Parameter values for the 5 simulations with the highest F_{tot} in the initial design (left) and in the 50 simulations added by EGO (Right). Parameter values were scaled and centred so they could all be represented on the same axis. To avoid redundancy in the combinations of parameters, results are represented only with positive values of $L_{max}^B - diff$.

This is also supported by the fact that the five combinations leading to the highest F_{tot} values all have very high values of $L_{max}^B - diff$ in the initial design (15 cm on average), whereas they have intermediate values of $L_{max}^B - diff$ in the 50 simulations added by EGO (8 cm on average) (Fig. 8). This is likely due to the fixation of $PAR_t - diff$ to its mean value, as there is a clear negative correlation between this parameter and $L_{max}^B - diff$. Interestingly, if the height differential between the mixed cultivars is always high in the 5 best simulations of the initial design, its value is highly variable in the 5 best combinations of the additional simulations. The influence of $H_{MS} - diff$ on F_{tot} thus probably depends on the values of other input parameters.

DISCUSSION

FSPM as a tool to study heterogeneous canopies

This study illustrates the interest of using FSPMs to explore the impact of architectural traits on the performance of heterogeneous canopies (Evers *et al.*, 2019 ; Gaudio *et al.*, 2019 ; Muller and Martre, 2019 ; Louarn and Song, 2020 ; Stomph *et al.*, 2020). Contrary to previous modelling work on wheat cultivar mixtures (Vidal *et al.*, 2018 ; Barillot *et al.*, 2019), this study focuses on multiple traits and on the entire development cycle of the plants. Thus, it provides a broader insight into the interactions taking place in wheat cultivar mixtures. Nevertheless, as for any modelling exercise (Passioura, 1996), our study presents some approximations that must be considered when interpreting the results. Indeed, even though WALTER allows for a realistic simulation of the tillering

dynamics in response to competition for light (Lecarpentier *et al.*, 2019), some of the underlying hypotheses of the model (detailed in Chapter 1) should be mentioned. For example, plasticity was considered only for tillering, as it is the most plastic trait in response to competition for light in wheat (Lecarpentier, 2017). However, considering plasticity in other traits, and especially in leaf dimensions, could lead to different results, although it would come at a higher computational cost. Furthermore, the results are also affected by the parameterization of the model. In particular, light interception was computed using only diffuse radiations, because the simulation of more realistic sky conditions (*i.e.* diffuse and direct radiations) would be significantly more time-consuming. Sky conditions can have an impact on our results and additional information could be provided by the comparison of our simulations with simulations using only direct radiations (Barillot *et al.*, 2019). Nonetheless, our preliminary studies (Appendix 1) showed that there were no significant differences between the results obtained with these 2 conditions when simulating a simple binary mixture. Finally, even though mixtures cultivated by farmers usually combine more than 2 cultivars, our study focuses on the simulation of binary mixtures, as these simple mixtures allow an efficient exploration of the relations between plant architecture and stand performance.

In contrast to these limitations, the strengths inherent in the modelling approach allowed us to conduct a study that would have been impossible to achieve experimentally. Contrary to experimental studies on wheat mixtures (Borg *et al.*, 2018 ; Montazeaud *et al.*, 2020), the effect of each architectural trait on the development of the stand could be isolated from any other factor. In addition, it was possible not to include correlations between architectural traits, such as the one between leaf dimensions and leaf inclination (Ma *et al.*, 2020), even though these biological constraints should be considered when interpreting the results. The sensitivity analysis combined with a metamodelling approach allowed us to take advantage of the complexity of an FSPM while limiting the computational time required. Thanks to the adaptive design, the uncertainty associated with the metamodelling approach was accounted for. Further conclusions could be drawn by extending the adaptive design to all the input factors.

Leaf dimensions and tillering are major drivers of mixture performance

Our study highlights the impact of the leaf dimensions and the tillering capability on the performance of the simulated mixtures. Indeed, for all the outputs considered, those 2 traits were the most influential ones among the 4 architectural traits explored. Leaf dimensions and tillering capability have an important impact on the plant's leaf area index (LAI) whereas leaf inclination and height don't have a direct impact on the LAI but rather on the spatial distribution of the leaf area and on its orientation. Thus, our results suggest that traits impacting the plants' LAI have more influence on the performance of the stand than traits impacting the arrangement of the leaves. This result is consistent with the well established effect of LAI on competition for light interception and productivity (Pugnaire and Valladares, 2007). Furthermore, it is in agreement with several studies using FSPMs to simulate trees (Da Silva *et al.*, 2014 ; Streit *et al.*, 2016 ; Perez *et al.*, 2018), or maize intercropped with soybean (Munz *et al.*, 2018), in which the traits related to leaf area were identified as the ones with the most influence on light interception. Finally, it is also consistent with the results of the experimental study of Montazeaud *et al.* (2020), in which the tillering capability was shown to have an important impact on the performance of binary wheat mixtures. The

relatively smaller impact of height and leaf inclination on stand performance is congruent with the findings of [Munz *et al.* \(2018\)](#) on intercropped maize and with the study of [Barillot *et al.* \(2019\)](#) on leaf inclination in wheat cultivar mixtures. However, even though their impact on mixture performance was less important than that of tillering and leaf dimensions, plant height and leaf inclination still had an important impact on some of the performance indicators. In particular, leaf inclination had an important impact on the light interception efficiency of the simulated canopies (L_{perc}) and on its OY. As for the differential of plant height between the cultivars in mixture, it was one of the most important factors influencing F_{tot} and the OY of N_{ears} .

Contrary to studies on pure stands or on single plants ([Sarlikioti *et al.*, 2011](#) ; [Da Silva *et al.*, 2014](#) ; [Streit *et al.*, 2016](#) ; [Perez *et al.*, 2018](#)), we were able to distinguish the effect of the trait differential between cultivars from the effect of the mean value of the trait in the stand, for each trait. Thus, for the leaf dimensions, the mean value in the canopy greatly impacts the light interception efficiency (L_{perc}) and the number of ears produced (N_{ears}), but it is mainly the differential between cultivars that influences the performance of the stand, with a major impact on all outputs except L_{perc} . As for the tillering capability, the differential between cultivars and the mean value in the stand both had an important impact on mixture performance. This is in line with the experimental study conducted by [Montazeaud *et al.* \(2020\)](#), in which mixture performance was impacted by both average trait values and trait differences. In particular, they showed an important effect of both the average tillering capability and its differential on the OY of grain yield.

It would be interesting to consider phenological traits, such as the number of days to heading or to maturity, in addition to the architectural traits studied here. Indeed, diversity in phenology could lead to temporal complementarity in addition to the spatial complementarity offered by differences in traits of the aerial architecture. Several studies have emphasized the impact of phenological traits on the performance of wheat mixtures ([Borg *et al.*, 2018](#) ; [Montazeaud *et al.*, 2020](#)). However, these traits were not yet considered in our study because it would require the integration of a specific module to WALTER to ensure that modifications in phenology and their implication on plant development are reliably simulated. Finally, future work could include the exploration of the effects of some agricultural practices. We simulated balanced mixtures (with 50% of each genotype) sown at a density of 200 plants/m². However, it is expected that both sowing density and the proportions of the 2 cultivars in mixture should have an impact on the results ([Grace and Tilman, 1990](#)). Moreover, the sowing density has been identified, as expected, as one of the most influential inputs in WALTER (Chapter 3 ; [Lecarpentier *et al.*, 2019](#)), supporting the interest in exploring variations of this parameter. FSPMs are particularly adapted to explore these agricultural practices ([Evers *et al.*, 2019](#) ; [Gaudio *et al.*, 2019](#)). WALTER could readily be used to consider the effects of variations in both sowing density and cultivar proportion as well as their interactions with architectural traits, albeit with an increased computational cost compared to the present study.

Towards the identification of assembly rules

Our results show that the performance of mixtures is very variable depending on the values of the explored architectural traits. In particular, the best performances were achieved by mixing cultivars with different architectures, which is in agreement with numerous studies linking the diversity of

functional traits in plant communities to their productivity (Reiss and Drinkwater, 2018). However, some of the worst performances were also achieved by mixing varieties differing in their aerial architecture and the OY indicators were less than 1 for a significant proportion of the simulated mixtures. This suggests that diversity is not a sufficient criterion to design efficient mixtures and that random-trait assembly will not necessarily lead to better performances than pure stands, as previously stated by Louarn *et al.* (2020) for grassland mixtures. Overall, these results highlight the importance of establishing assembly rules to design efficient mixtures.

When focusing on the trait configurations leading to the best performances, our results converge towards a single mixture ideotype for the total fitness of the plot (F_{tot}). Indeed, even if some parameters can have variable values in well performing stands, due to their low impact on F_{tot} , there is no multimodality in the distribution of the parameters with high impact on F_{tot} for the best performing simulations. In particular, regarding the mean values of traits in the stand, the best-performing canopies are composed of short plants with low levels of tiller regression. Interestingly, the fact that stands with short plants perform better than stands with taller plants is in line with the assumption of a trade-off between vegetative development and reproduction. As a matter of fact, significant yield gains have been achieved by breeding for wheat cultivars with reduced height, as a shorter stature limits lodging and improves the harvest index (Gale *et al.*, 1985). In WALTer, the negative correlation between plant height and stand performance is probably due to the formalism of tiller regression. Indeed, tiller regression depends on the amount of light intercepted per unit area, and taller tillers have a larger surface of internodes and sheaths, which makes them more likely to regress.

Regarding the differential of trait values between the 2 varieties in the mixture, the best-performing stands are the ones with an important diversity in both leaf dimensions and tiller regression. Our results tend to encourage the association of a very competitive genotype with a genotype with poor competitive ability, as the genotype with the largest leaves must also be the one less sensitive to tiller regression in order for the stand to perform well. This conclusion must however be qualified as the cultivar with the largest leaves should also be the one with the earliest cessation of tillering in response to the neighbouring competition. Thus, in the best-performing plots, the most productive genotype at maturity was usually not the one that produced the most tillers. Regardless, the high F_{tot} values offered by this trait combination is explained by the very important productivity gain of the cultivar with the largest leaves, which outweighs the loss of productivity of the other cultivar in the mixture. However, it is important to specify that no plant mortality is considered in WALTer and that all individuals produce at least one ear, which is not always the case in experimental settings. In the case of the best-performing configurations that were simulated with WALTer, the cultivar with the smallest leaves had a very low productivity that might have been even lower in field conditions because of potential plant death or individuals with null productivity. Under such experimental conditions, the gain of the cultivar with the largest leaves might not have compensated for the loss in productivity of the other genotype, thus calling for caution when mixing cultivars with significant differences in their competitive ability. Moreover, the results of Montazeaud *et al.* (2020) showed a negative effect of diversity in tillering capability on the performance of binary mixtures of spring durum wheat sown in rows, which is in contradiction with our results. However, even if the results of our simulations are not meant to be directly translated into assembly rules, they improve our understanding of the interactions taking place in mixtures.

Finally, the trait combinations identified in this study are the ones maximising F_{tot} , which is a proxy for the number of grains produced by the stand. However, cultivar mixtures can provide a wide variety of ecosystem services besides yield (Barot *et al.*, 2017). Thus, it would be interesting to consider multiple objectives as exemplified by the study of Louarn *et al.* (2020) in which the stability of grassland mixtures over time was also considered in addition to yield. Considering the stability of the composition of a mixture over time could also be of interest for wheat, as this crop is often resown from year to year. The mixtures that we identified to maximise F_{tot} present significant differences in performance between their 2 cultivars, which would lead to a drastic evolution of their composition and possibly to a decrease in stand performance over time. Further work with WALTER would be of high interest to identify trait configurations leading to both high performance and mixture stability over time.

REFERENCES

- Abbate, P. E., Andrade, F. H., & Culot, J. P. (1995).** The effects of radiation and nitrogen on number of grains in wheat. *The Journal of Agricultural Science*, 124(3), 351-360.
- Barillot, R., Chambon, C., Fournier, C., Combes, D., Pradal, C., & Andrieu, B. (2019).** Investigation of complex canopies with a functional–structural plant model as exemplified by leaf inclination effect on the functioning of pure and mixed stands of wheat during grain filling. *Annals of botany*, 123(4), 727-742.
- Barot, S., Allard, V., Cantarel, A., Enjalbert, J., Gauffreteau, A., Goldringer, I., ... & Porcher, E. (2017).** Designing mixtures of varieties for multifunctional agriculture with the help of ecology. A review. *Agronomy for sustainable development*, 37(2), 13.
- Borg, J., Kiær, L. P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., ... & Enjalbert, J. (2018).** Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*, 221, 298-313.
- Chelle, M., & Andrieu, B. (1998).** The nested radiosity model for the distribution of light within plant canopies. *Ecological Modelling*, 111(1), 75-91.
- Chen, T. W., Nguyen, T. M. N., Kahlen, K., & Stützel, H. (2014).** Quantification of the effects of architectural traits on dry mass production and light interception of tomato canopy under different temperature regimes using a dynamic functional–structural plant model. *Journal of Experimental Botany*, 65(22), 6399-6410.
- Currin, C., Mitchell, T., Morris, M., & Ylvisaker, D. (1991).** Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 86(416), 953-963.
- Da Silva, D., Han, L., Faivre, R., & Costes, E. (2014).** Influence of the variation of geometrical and topological traits on light interception efficiency of apple trees: sensitivity analysis and metamodelling for ideotype definition. *Annals of botany*, 114(4), 739-752.
- Demotes-Mainard, S., & Jeuffroy, M. H. (2001).** Incorporating radiation and nitrogen nutrition into a model of kernel number in wheat. *Crop Science*, 41(2), 415-423.
- Donald, C. T. (1968).** The breeding of crop ideotypes. *euphytica*, 17(3), 385-403.
- den Dulk, J. A. (1989).** The interpretation of remote sensing: a feasibility study (Doctoral dissertation, Den Dulk).
- Dupuy, D., Helbert, C., Franco, J. (2015).** DiceDesign and DiceEval: Two R Packages for Design and Analysis of Computer Experiments. *Journal of Statistical Software*, 65(11), 1-38. URL <http://www.jstatsoft.org/v65/i11/>.
- Evers, J. B., Van Der Werf, W., Stomph, T. J., Bastiaans, L., & Anten, N. P. (2019).** Understanding and optimizing species mixtures using functional–structural plant modelling. *Journal of experimental botany*, 70(9), 2381-2388.

- Faraji, J. (2011).** Wheat cultivar blends: A step forward to sustainable agriculture. *African Journal of Agricultural Research*, 6(33), 6780-6789.
- Fischer, R. A. (1985).** Number of kernels in wheat crops and the influence of solar radiation and temperature. *The Journal of Agricultural Science*, 105(2), 447-461.
- Food and Agriculture Organization of the United Nations. (2018).** FAOSTAT statistical database. Available at <http://www.fao.org/faostat/en/#data/QC> (accessed 25 August 2020).
- Fournier, C., Andrieu, B., Ljutovac, S., & Saint-Jean, S. (2003).** ADEL-Wheat: A 3D Architectural Model of wheat development. In *Plant Growth Modeling and Applications* (pp. 54-63). Springer Verlag.
- FranceAgriMer (2019).** Enquête répartition variétale des céréales.
- Gale, M. D., Youssefian, S., & Russell, G. E. (1985).** Dwarfing genes in wheat. *Progress in plant breeding*, 1, 1-35.
- Gaudio, N., Escobar-Gutiérrez, A. J., Casadebaig, P., Evers, J. B., Gerard, F., Louarn, G., ... & Barillot, R. (2019).** Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agronomy for sustainable development*, 39(2), 20.
- Givnish, T. J. (1982).** On the adaptive significance of leaf height in forest herbs. *The American Naturalist*, 120(3), 353-381.
- Godin, C., Sinoquet, H. (2005).** Functional–structural plant modelling. *New phytologist*, 166(3), 705-708.
- Grace, J. B., & Tilman, D. (Eds.). (1990).** *Perspectives on Plant Competition*. Academic Press.
- Iooss, B., Da Veiga, S., Janon, A., Pujol, G. (2017).** sensitivity: Global Sensitivity Analysis of Model Outputs. R package version 1.15.0. <https://CRAN.R-project.org/package=sensitivity>
- Isbell, F. (2015).** Agroecology: agroecosystem diversification. *Nature plants*, 1(4), 1-2.
- Johnson, M. E., Moore, L. M., & Ylvisaker, D. (1990).** Minimax and maximin distance designs. *Journal of statistical planning and inference*, 26(2), 131-148.
- Jones, D. R., Schonlau, M., & Welch, W. J. (1998).** Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4), 455-492.
- Lecarpentier, C. (2017).** Plasticité de l'architecture aérienne du blé en réponse à la compétition pour la lumière au sein de cultures pure ou d'associations variétales : caractérisation expérimentale et développement d'un modèle 3D. Thèse de doctorat, Université Paris-Saclay.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Litrico, I., & Violle, C. (2015).** Diversity in plant breeding: a new conceptual framework. *Trends in plant science*, 20(10), 604-613.

- Long, S. P., ZHU, X. G., Naidu, S. L., & Ort, D. R. (2006).** Can improvement in photosynthesis increase crop yields?. *Plant, cell & environment*, 29(3), 315-330.
- Louarn, G., Barillot, R., Combes, D., & Escobar-Gutiérrez, A. (2020).** Towards intercrop ideotypes: non-random trait assembly can promote overyielding and stability of species proportion in simulated legume-based mixtures. *Annals of Botany*.
- Louarn, G., & Song, Y. (2020).** Two decades of functional–structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany*, 126(4), 501-509.
- Ma, J., Tu, Y., Zhu, J., Luo, W., Liu, H., Li, C., ... & Mu, Y. (2020).** Flag leaf size and posture of bread wheat: genetic dissection, QTL validation and their relationships with yield-related traits. *Theoretical and Applied Genetics*, 133(1), 297-315.
- Malézieux, E. (2012).** Designing cropping systems from nature. *Agronomy for sustainable development*, 32(1), 15-29.
- Marrel, A., Iooss, B., Laurent, B., & Roustant, O. (2009).** Calculations of sobol indices for the gaussian process metamodel. *Reliability Engineering & System Safety*, 94(3), 742-751.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979).** Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239-245.
- Midmore, D. J., Cartwright, P. M., & Fischer, R. A. (1984).** Wheat in tropical environments. II. Crop growth and grain yield. *Field Crops Research*, 8, 207-227.
- Montazeaud, G., Violle, C., Roumet, P., Rocher, A., Ecarnot, M., Compan, F., ... & Fréville, H. (2020).** Multifaceted functional diversity for multifaceted crop yield: Towards ecological assembly rules for varietal mixtures. *Journal of Applied Ecology*.
- Moon, P., Spencer, D. (1942).** Illumination from a non-uniform sky, Transactions of the Illumination Engineering Society 37 (1942) 707712.
- Muller, B., Martre P. (2019).** Plant and crop simulation models: powerful tools to link physiology, genetics, and phenomics. *Journal of Experimental Botany*, 70(9), 2339–2344.
- Munz, S., Graeff-Hönninger, S., & Henke, M. (2018).** Functional-structural plant model for testing the effect of maize architecture on hourly light distribution in strip-intercropping systems. In *2018 6th International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA)* (pp. 115-119). IEEE.
- Nix, H. A. (1976).** Climate and crop productivity in Australia. Climate and rice. IRRI, Los Baños, The Philippines, 495-507.
- Niinemets, Ü. (2010).** A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecological Research*, 25(4), 693-714.
- Passioura, J. B. (1996).** Simulation models: science, snake oil, education, or engineering?. *Agronomy Journal*, 88(5), 690-694.

- Perez, R. P., Dauzat, J., Pallas, B., Lamour, J., Verley, P., Caliman, J. P., ... & Faivre, R. (2018).** Designing oil palm architectural ideotypes for optimal light interception and carbon assimilation through a sensitivity analysis of leaf traits. *Annals of botany*, 121(5), 909-926.
- Pugnaire, F., & Valladares, F. (Eds.). (2007).** *Functional plant ecology*. CRC press.
- R Core Team (2018).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reiss, E. R., & Drinkwater, L. E. (2018).** Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecological Applications*, 28(1), 62-77.
- Reynolds, M., Foulkes, J., Furbank, R., Griffiths, S., King, J., Murchie, E., ... & Slafer, G. (2012).** Achieving yield gains in wheat. *Plant, cell & environment*, 35(10), 1799-1823.
- Roustant, O., Ginsbourger, D., Deville, Y. (2012).** DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software*, 51(1), 1-55. URL <http://www.jstatsoft.org/v51/i01/>.
- Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989).** Design and analysis of computer experiments. *Statistical science*, 409-423.
- Sarlikioti, V., de Visser, P. H., Buck-Sorlin, G. H., & Marcelis, L. F. M. (2011).** How plant architecture affects light absorption and photosynthesis in tomato: towards an ideotype for plant architecture using a functional–structural plant model. *Annals of Botany*, 108(6), 1065-1073.
- Sobol, I. M. (1993).** Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4), 407-414.
- Stomph, T., Dordas, C., Baranger, A., de Rijk, J., Dong, B., Evers, J., ... & Wang, Q. (2020).** Designing intercrops for high yield, yield stability and efficient use of resources: Are there principles?. In *Advances in Agronomy* (Vol. 160, No. 1, pp. 1-50). Academic Press.
- Streit, K., Henke, M., Bayol, B., Cournède, P. H., Sievänen, R., & Kurth, W. (2016, November).** Impact of geometrical traits on light interception in conifers: analysis using an FSPM for Scots pine. In *2016 IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA)* (pp. 194-203). IEEE.
- Valladares, F., & Niinemets, Ü. (2008).** Shade tolerance, a key plant feature of complex nature and consequences. *Annual Review of Ecology, Evolution, and Systematics*, 39.
- Vidal, T., Gigot, C., de Vallavieille-Pope, C., Huber, L., & Saint-Jean, S. (2018).** Contrasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. *Annals of botany*, 121(7), 1299-1308.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional–structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, 61(8), 2101-2115.
- Zhang, B., & DeAngelis, D. L. (2020).** An Overview of Agent-Based Models in Plant Biology and Ecology. *Annals of Botany*, 126(4), 539–557.

CHAPITRE 6

**Étude de l'évolution de
peuplements hétérogènes de blé
par le couplage de WALTer et d'un
modèle de génétique des
populations**

INTRODUCTION

Une meilleure mobilisation de la diversité cultivée fait partie des leviers clefs de la transition agroécologique (Malézieux, 2012 ; Isbell, 2015). La diversité génétique intra-spécifique peut ainsi être mobilisée, à l'échelle de la parcelle, par la culture de peuplements hétérogènes. Pour le blé tendre (*Triticum aestivum* L.), il est possible de cultiver des variétés anciennes, des variétés population, ou des mélanges de variétés. Chacun de ces peuplements présente en effet une diversité génétique plus importante que celle d'une culture monovariétale de blé moderne (Khan *et al.*, 2020). La mobilisation de l'agro-biodiversité passe également par des méthodes de gestion de ces peuplements hétérogènes. En effet, il a été montré que la gestion dynamique des populations de blé permet d'augmenter leur diversité génétique et ainsi d'améliorer leur adaptabilité par rapport aux parcelles cultivées à partir de graines conservées ex-situ dans des banques de gènes (Khan *et al.*, 2020). La gestion dynamique consiste à ressemer une partie des grains produits par une parcelle d'une année sur l'autre. Les peuplements gérés de cette façon sont soumis aux forces évolutives, telles que la dérive génétique^a, les mutations et la sélection naturelle, et peuvent s'adapter aux conditions environnementales et aux pratiques de culture auxquelles ils sont soumis (Enjalbert *et al.*, 2011). La gestion dynamique peut-être mise en place à la ferme par les agriculteurs, ou dans des stations expérimentales.

En France, un important programme expérimental de gestion dynamique du blé tendre a ainsi été mis en place par l'INRA (Institut National de la Recherche Agronomique) en 1984. Le programme était basé sur 3 populations composites^b cultivées dans un réseau de 12 sites répartis sur le territoire métropolitain (Henry *et al.*, 1991). Parmi les 3 populations du programme, la population PA, qui provient d'un croisement entre 16 lignées parentales (Thomas *et al.*, 1991), a été ressemées pendant 20 années consécutives et a fait l'objet de nombreuses études, permettant d'améliorer les connaissances associées à la pratique de la gestion dynamique (Enjalbert *et al.*, 2011). L'évolution phénotypique et génétique de PA au cours des générations a été étudiée pour des traits adaptatifs majeurs de phénologie, de résistance aux maladies et d'architecture. Il a ainsi été montré que la population PA évoluait sous l'action de la sélection naturelle et pas seulement de la dérive génétique (Enjalbert *et al.*, 1999). Après plusieurs générations de resemis, une différenciation spatiale de PA, en fonction du site sur lequel elle était cultivée, a été mise en évidence pour la précocité de floraison (Goldringer *et al.*, 2006 ; Rhone *et al.*, 2008 ; Rhone *et al.*, 2010) ainsi que pour la résistance à l'oïdium (Le Boulc'h *et al.*, 1994 ; Paillard *et al.*, 2000a, b). En parallèle de cette adaptation locale, PA a subi une sélection directionnelle pour la hauteur dans tous les sites où elle était cultivée, conduisant à une augmentation importante de la taille des plantes au cours des générations (Goldringer *et al.*, 2001a ; Raquin *et al.*, 2008a).

Cependant, malgré les nombreuses connaissances apportées par ce type d'expérimentation, l'évolution des populations menées en gestion dynamique reste encore insuffisamment comprise (Khan *et al.*, 2020). La conservation de la diversité étant un des objectifs de la gestion dynamique, il est essentiel de comprendre le rôle des différentes forces évolutives dans le maintien ou la perte de diversité de ces populations dans le temps. Une part importante de la diversité présente dans la

a La dérive génétique correspond à l'effet de l'aléatoire sur l'évolution des populations.

b Les populations composites, "Composite Cross Populations" (CCP) en anglais, sont des populations hétérogènes qui résultent du croisement entre un nombre important de génotypes parentaux.

population initiale peut être perdue au cours des générations à cause de la dérive génétique, mais la sélection naturelle peut également entraîner des pertes de diversité importantes et peut même conduire à l'augmentation en fréquence d'allèles délétères par l'effet auto-stop^c (Goldringer *et al.*, 2001b). C'est pourquoi il est intéressant d'étudier l'impact de la diversité de traits architecturaux sur l'évolution des populations menées en gestion dynamique. Étant donné le rôle clef que joue l'architecture aérienne dans la compétition pour la lumière, on peut en effet s'attendre à ce qu'une sélection s'opère sur certains traits, comme cela a été montré pour la hauteur dans la population PA (Goldringer *et al.*, 2001a ; Raquin *et al.*, 2008a).

Afin d'étudier l'impact de l'architecture aérienne sur l'évolution de populations de blé menées en gestion dynamique, la modélisation peut s'avérer très utile, en complément des expérimentations. En effet, les expérimentations de gestion dynamique sont complexes et lourdes à mettre en place et à maintenir (Enjalbert *et al.*, 2011). La modélisation permet alors d'étudier l'évolution des populations à moindre coût et en contrôlant complètement les conditions testées. Les modèles de plante structure-fonction (FSPM ; Vos *et al.*, 2010 ; Godin et Sinoquet, 2005), comme WALTer (Lecarpentier *et al.*, 2019 ; Chapitre 1), permettent de simuler avec précision la compétition qui prend place entre les individus dans des peuplements avec des diversités d'architecture. Pour représenter l'évolution de ces peuplements lorsqu'ils sont resemés d'une année sur l'autre, le couplage à un modèle de génétique des populations est nécessaire. Ce chapitre présente ainsi le modèle de génétique des populations GenoWALT, et son couplage à WALTer. Afin d'illustrer le fonctionnement de ce couplage et de valider son bon fonctionnement, des simulations ont été réalisées en se basant sur l'expérience de gestion dynamique de PA, en se concentrant sur la diversité de hauteur dans la population.

MATÉRIEL ET MÉTHODES

La génétique du blé

Le blé tendre (*Triticum aestivum* L.) est un organisme allohexaploïde à 42 chromosomes qui provient d'événements d'hybridation et de polyploïdisation^d à partir de 3 ancêtres diploïdes ayant chacun 7 paires de chromosomes (Goldringer *et al.*, 2013). Le génome du blé tendre est donc divisé en 3 parties : A, B et D, correspondant aux 3 origines distinctes. Le génome A provient de *Triticum urartu* (Dvořák *et al.*, 1993), le génome B est proche de celui de l'espèce sauvage actuelle *Aegilops speltoides* (Sarkar et Stebbins, 1956 ; Kilian *et al.*, 2007) et le génome D provient de l'espèce sauvage *Aegilops tauschii* (Wang *et al.*, 2013). Cependant, il n'y a pas de crossing-over^e entre les chromosomes de ces 3 génomes, *i.e.* le blé a un comportement diploïde pendant la méiose. En effet, la divergence des chromosomes A, B et D est suffisante pour éviter leur appariement à la méiose. Le blé sera donc modélisé comme une espèce diploïde à 21 chromosomes dans ce chapitre (Figure 1). Le blé est autogame, ce qui signifie que la reproduction se fait par la fécondation d'un individu avec lui-même (autofécondation). Un faible pourcentage d'allogamie existe cependant chez cette espèce, ce qui veut dire que la fécondation a parfois lieu à partir de gamètes provenant d'individus

c L'effet auto-stop, "genetic hitchhiking" en anglais, est défini dans l'Annexe 4.

d Multiplication du nombre de chromosomes.

e Recombinaison intrachromosomique pendant la méiose

différents.

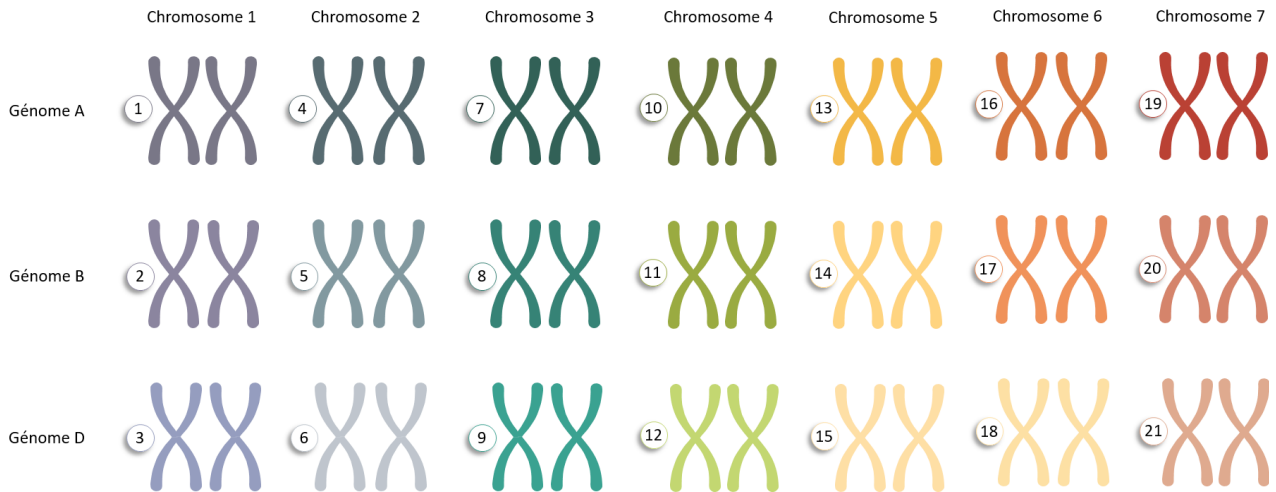


Figure 1. Caryotype du blé tendre (*Triticum aestivum*) avec la numérotation retenue dans la carte génétique (Tableau 1)

Le modèle GenoWALT

Un modèle de génétique des populations, appelé GenoWALT, a été développé dans le cadre du stage de M1 de Meije Gawinowski (2018). Le modèle a été développé dans l'environnement python simuPOP (Peng et Kimmel, 2005) et permet de simuler la reproduction d'une population et la ségrégation des marqueurs génétiques (gènes, QTL, marqueurs neutres) pour des régimes allant de l'autogamie à l'allogamie. Pour fonctionner, le modèle utilise une cartographie génétique renseignée par l'utilisateur et qui comporte : la position des marqueurs, le régime de dominance de leurs allèles et leur valeur d'effet sur le(s) trait(s) étudié(s). Dans sa version actuelle, le modèle permet seulement de considérer des marqueurs bi-alléliques. Une population initiale est également définie par l'utilisateur en précisant le génotype (état allélique pour chaque marqueur de la carte) et la fitness de chaque individu. Le nombre de descendants de la population doit également être renseigné, ce qui permet à GenoWALT de sélectionner les parents en se basant sur un tirage pondéré par la fitness des individus donnée en entrée du modèle. Le modèle génère ainsi les génotypes de la génération suivante en prenant en compte les effets de la sélection (via la fitness), de la recombinaison et de la dérive génétique (effet de l'aléatoire dans les populations de taille limitée). En revanche, GenoWALT ne considère pas de mutation ou de migration (qui pourrait correspondre, ici, à un apport de pollen ou de graines de l'extérieur). En se basant sur les valeurs d'effet des différents marqueurs, le modèle calcule la valeur génétique de chaque trait étudié pour chaque individu de la nouvelle génération. Le passage des valeurs d'effet des marqueurs aux valeurs de trait se fait par un modèle additif simple dans lequel la valeur génétique d'un trait pour un individu est obtenue en sommant les valeurs d'effet des différents marqueurs chez cet individu (cf. Annexe 4 - Figure 1). L'épistasie, *i.e.* l'effet des interactions entre marqueurs, n'est pas considérée par le modèle. La valeur d'effet d'un marqueur est calculée en fonction du génotype de l'individu à ce marqueur et du régime de dominance. Seules la dominance complète d'un allèle ou une co-dominance peuvent être simulées par GenoWALT dans sa version actuelle : des effets de dominance partielle ne peuvent pas être représentés. En revanche, la pléiotropie est prise en compte par le modèle : un même marqueur peut avoir un impact sur plusieurs traits.

Afin de s'assurer du bon fonctionnement de GenoWALT, de nombreux tests ont été réalisés dans le cadre du stage de Meije Gawinowski (résultats non présentés ici).

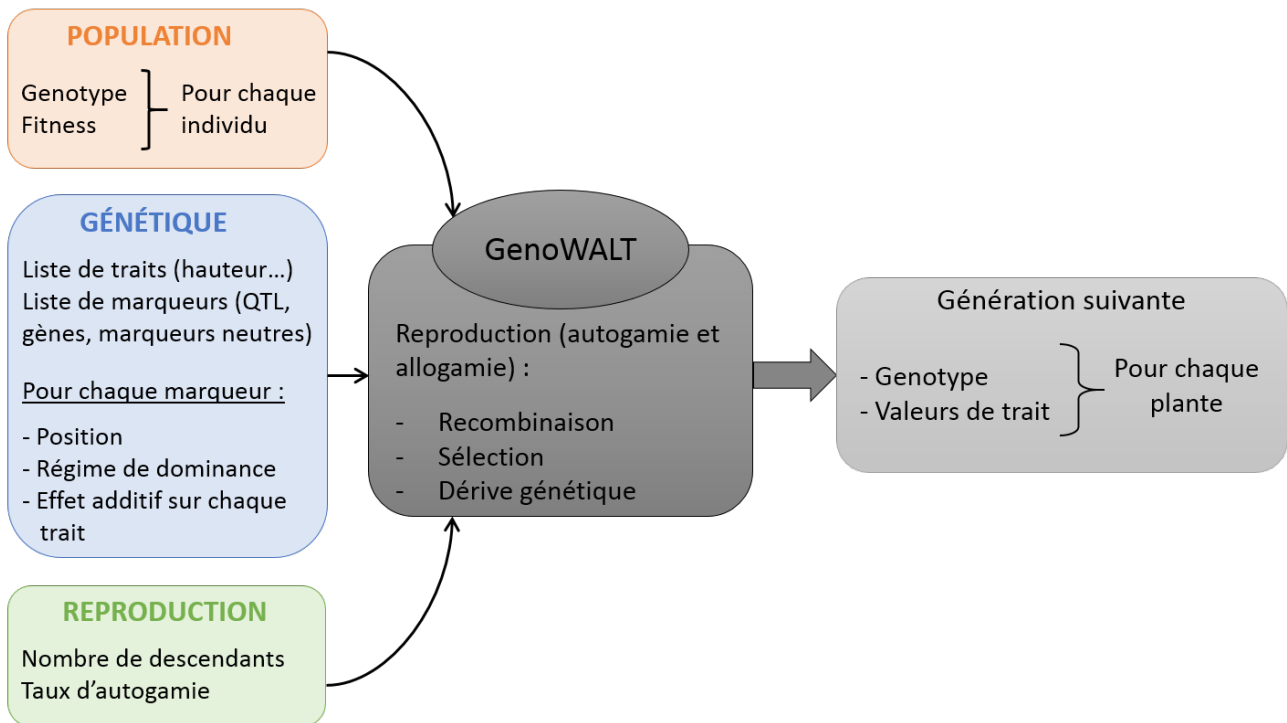


Figure 2. Représentation schématique du fonctionnement du modèle GenoWALT.

Couplage avec WALTer

Le modèle GenoWALT peut être couplé à WALTer pour simuler l'évolution d'une population de blé en prenant en compte l'impact de la compétition pour la lumière entre les individus. Avec ce couplage, WALTer fonctionne de la même manière que lorsqu'il est utilisé seul (c.f Chapitre 1), mais le nombre de génotypes à simuler et les valeurs de certains paramètres génotypiques (hauteur finale du brin maître par exemple) sont fournis via GenoWALT au lieu d'être définis directement par l'utilisateur. De la même manière, le fonctionnement de GenoWALT n'est pas impacté par le couplage, mais les valeurs de fitness utilisées par le modèle de génétique sont générées par WALTer pour chaque individu au lieu d'être renseignées par l'utilisateur.

Afin de s'assurer du bon fonctionnement du couplage entre les 2 modèles, le resemis d'un mélange binaire simple a été simulé et est présenté en Annexe 4.

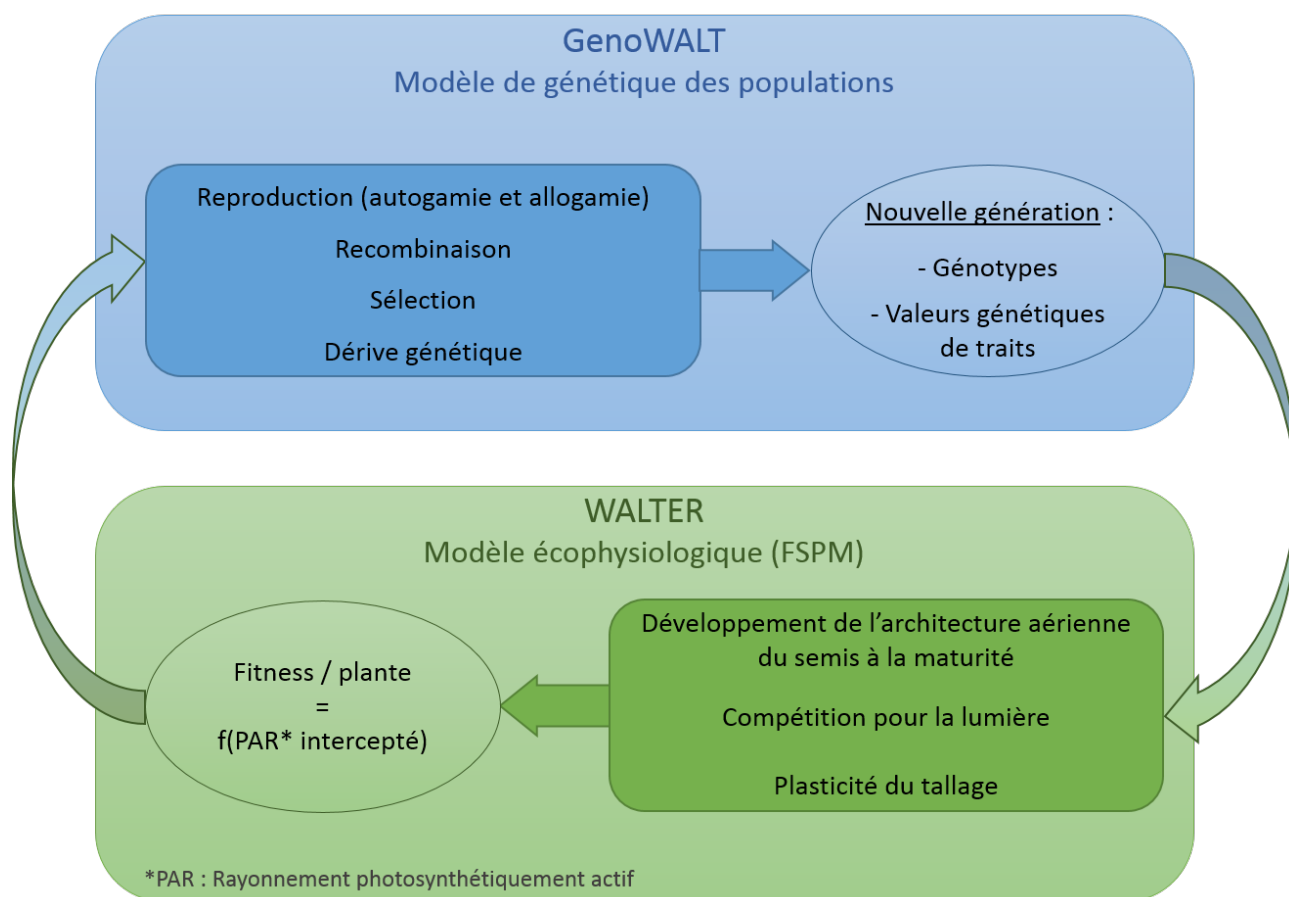


Figure 3. Représentation schématique du couplage entre WALTER et GenoWALT. Le couplage permet de simuler le resemis de populations de blé sur plusieurs années : WALTER simule le développement des parcelles du semis à la maturité et GenoWALT simule la phase de reproduction à chaque génération.

Expérience de gestion dynamique in silico^f

Le couplage entre le FSPM WALTER et le modèle de génétique des populations GenoWALT a été utilisé pour simuler le resemis, sur 20 générations, d'une population de blé diversifiée pour la hauteur.

Carte génétique

Afin de représenter le déterminisme génétique sous tendant les différences de hauteur entre les individus de la population, une carte génétique a été constituée. Chez le blé, la hauteur est contrôlée par un ensemble de gènes majeurs, notamment les gènes 'Rht' (pour '*reduced height*') (Gale *et al.*, 1985), ainsi que par des QTL (Quantitative Trait Loci) avec des effets plus ou moins importants sur la hauteur finale de la plante (Griffiths *et al.*, 2012 ; Würschum *et al.*, 2015).

^f Cette expérience a bénéficié de l'appui d'Antoine Payen dans le cadre de son stage de L3 (2019).

Tableau 1. Carte génétique utilisée pour les simulations. Pour chaque marqueur : nom, numéro du chromosome sur lequel le marqueur se trouve (cf. Figure 1), position sur le chromosome (en cM), effet additif du marqueur et effet additif mis à l'échelle (en cm), allèle dominant et référence bibliographique associée au marqueur

Nom	Chrom.	Position	Effet additif (cm)	Effet add. mis à échelle	Dominance	Référence
qPH1	1	42.7	1.13	1.28	haut	Griffiths <i>et al.</i> , 2012
qPH2	2	57.8	1.71	1.93	co-dominance	Griffiths <i>et al.</i> , 2012
qPH3	3	59.2	1.68	1.90	haut	Griffiths <i>et al.</i> , 2012
qPH4	4	30.0	2.94	3.32	co-dominance	Griffiths <i>et al.</i> , 2012
qPH5	4	111.0	2.49	2.82	haut	Griffiths <i>et al.</i> , 2012
qPH6	5	42.0	2.09	2.36	co-dominance	Griffiths <i>et al.</i> , 2012
Rht8	6	123.6	4.43	5.01	haut	Griffiths <i>et al.</i> , 2012
qPH7	7	50.4	3.01	3.40	co-dominance	Griffiths <i>et al.</i> , 2012
qPH8	8	75.7	1.97	2.23	haut	Griffiths <i>et al.</i> , 2012
Rht1	11	17.0	5.8	6.56	co-dominance	Raquin <i>et al.</i> , 2008a
qPH9	11	42.0	1.43	1.61	haut	Griffiths <i>et al.</i> , 2012
Rht2	12	12.0	5.18	5.85	co-dominance	Griffiths <i>et al.</i> , 2012
qPH10	13	60.3	1.50	1.70	haut	Griffiths <i>et al.</i> , 2012
qPH11	14	116.4	1.67	1.88	co-dominance	Griffiths <i>et al.</i> , 2012
qPH12	16	32.2	3.07	3.47	haut	Griffiths <i>et al.</i> , 2012
qPH13	17	35.8	1.33	1.50	co-dominance	Griffiths <i>et al.</i> , 2012
qPH14	18	70.3	2.06	2.33	haut	Griffiths <i>et al.</i> , 2012
miniqPH1	1	102.0	0.1	0.11	co-dominance	-
miniqPH2	2	102.0	0.2	0.23	haut	-
miniqPH3	3	102.0	0.3	0.34	co-dominance	-
miniqPH4	4	70.0	0.4	0.45	haut	-
miniqPH5	5	102.0	0.5	0.57	co-dominance	-
miniqPH6	6	50.0	0.6	0.68	haut	-
miniqPH7	7	102.0	0.7	0.79	co-dominance	-
miniqPH8	8	25.0	0.8	0.90	haut	-
miniqPH9	9	50.0	0.9	1.02	co-dominance	-
miniqPH10	10	50.0	0.1	0.11	haut	-
miniqPH11	11	102.0	0.2	0.23	co-dominance	-
miniqPH12	12	102.0	0.3	0.34	haut	-
miniqPH13	13	10.0	0.4	0.45	co-dominance	-
miniqPH14	14	50.0	0.5	0.57	haut	-
miniqPH15	15	50.0	0.6	0.68	co-dominance	-
miniqPH16	16	102.0	0.7	0.79	haut	-
miniqPH17	17	102.0	0.8	0.90	co-dominance	-
miniqPH18	18	22.0	0.9	1.02	haut	-
miniqPH19	19	50.0	0.1	0.11	co-dominance	-
miniqPH20	20	50.0	0.2	0.23	haut	-
miniqPH21	21	50.0	0.3	0.34	co-dominance	-

Pour constituer la carte, les 16 méta-QTL détectés chez des cultivars européens de blé d'hiver dans l'étude de [Griffiths et al. \(2012\)](#) ont été utilisés. La position et la valeur d'effet de chaque méta-QTL, incluant les gènes majeurs *Rht2* (*Rht-D1*) et *Rht8*, ont ainsi été intégrées à la carte génétique. Afin de compléter cette carte, le gène majeur *Rht1* (*Rht-B1*) y a été ajouté : la valeur d'effet sélectionnée pour ce marqueur est celle identifiée par [Raquin et al. \(2008a\)](#) et sa position correspond à celle sur la carte consensus de [Somers et al. \(2004\)](#), qui est utilisée dans l'étude de [Griffiths et al. \(2012\)](#). Un ensemble de 21 QTL à effet faible (effet additif < 1 cm) a également été ajouté à la carte génétique, à raison de 1 QTL par chromosome et en s'assurant de les positionner à distance des autres QTL de la carte. Une co-dominance a été choisie pour une moitié des marqueurs de la carte et une dominance de l'allèle "haut" a été sélectionnée pour l'autre moitié des marqueurs. Le régime de dominance choisi permet de produire des résultats réalistes, en accord avec les observations faites dans plusieurs études, dans lesquelles les plantes de blé hybrides sont régulièrement aussi hautes, voire plus hautes, que leur parent le plus haut ([Zhao et al., 2014](#) ; [Würschum et al., 2018](#)). Il a été considéré que tous les marqueurs de la carte étaient bi-alléliques, avec un allèle 'haut' et un allèle 'court'. Une mise à l'échelle des valeurs d'effet additif trouvées dans la bibliographie a été effectuée (Tableau 1), afin de permettre la simulation de plantes dont la hauteur finale est réaliste biologiquement (la gamme de variation de la hauteur était trop faible sans cette mise à l'échelle). La carte génétique permet ainsi de représenter des plantes dont la hauteur finale est comprise entre 40 cm et 160 cm. Enfin, des marqueurs neutres (sans effet sur le phénotype des individus) bi-alléliques ont également été ajoutés à cette carte, sur l'ensemble du génome, et en particulier à proximité des gènes majeurs *Rht-B1* (*Rht1*) et *Rht-D1* (*Rht2*). Pour chacun de ces 2 gènes majeurs, 10 marqueurs neutres ont ainsi été ajoutés tous les 0,5 cM^g dans une zone de 5 cM autour du gène. Les plantes simulées étant phénotypiquement identiques, à l'exception de leur différence de hauteur, aucun autre marqueur que ceux contrôlant la hauteur et les marqueurs neutres n'a été considéré ici.

Création de la population initiale

La population simulée reproduit la population *PA* décrite par [Thomas et al. \(1991\)](#) et qui provient d'un croisement pyramidal entre 16 parents. Afin de reproduire au mieux *PA*, les 16 parents de la population ont été simulés. À partir de la carte génétique décrite ci-dessus, les génomes des 16 parents ont été paramétrés de façon à ce qu'ils soient tous homozygotes pour tous les marqueurs. L'état allélique pour les gènes *Rht1* et *Rht2* étant renseigné pour chaque parent dans l'étude de [Thomas \(1986\)](#), il a été reproduit pour les génomes simulés. Une combinaison aléatoire de l'ensemble des autres QTL a été générée et sélectionnée pour chaque parent, pour que leur hauteur finale corresponde à celle renseignée par [Thomas \(1986 ; Tableau 2\)](#). D'autre part, l'état allélique des marqueurs neutres proches de *Rht1* et *Rht2* a été sélectionné afin de reproduire une décroissance du déséquilibre de liaison^h (DL) avec l'augmentation de la distance par rapport au gène majeur, tout en s'assurant d'obtenir des niveaux de DL proches de ceux mesurés expérimentalement autour de *Rht1* chez les 16 parents ([Raquin et al., 2008a](#) ; Figure 5). Pour les autres marqueurs neutres, l'état allélique a été choisi aléatoirement.

g Centimorgan (cM) : unité de mesure de la distance génétique entre 2 loci. La distance génétique est un indicateur du taux de recombinaison entre les 2 loci : une distance génétique de 1 cM entre 2 loci correspond à un taux de recombinaison de 1%.

h Le déséquilibre de liaison est une mesure de l'association entre 2 marqueurs (cf. Annexe 4).

À partir des 16 parents ainsi créés, les 2 premiers croisements du plan expérimental (Thomas *et al.*, 1991) ont été reproduits à l'aide de GenoWALT : les hybrides F1 ont tout d'abord été générés, puis 20 hybrides doubles ont été créés pour chaque type de croisement de F1 (Figure 4). Grâce à GenoWALT, les hybrides doubles ainsi obtenus ont ensuite été croisés en panmixieⁱ avec un taux d'allogamie de 100% pendant 2 générations afin de reproduire le brassage génétique obtenu par les 2 derniers croisements du plan expérimental (Figure 4, croisements 3 et 4). Pour les croisements 2, 3 et 4, les tailles de population ont été multipliées par 10 d'une génération à l'autre, afin d'obtenir une diversité génétique suffisante. Les 8000 individus obtenus à l'issue de ces 4 croisements constituent la population *PA* simulée. Afin de limiter le temps de simulation pour les prochaines étapes, 3 échantillons de 506 plantes ont été sélectionnés aléatoirement dans *PA*. Chacun de ces échantillon constitue la génération 0 (*PA*₀) d'un des 3 réplicats de l'expérience.

Tableau 2. Parents de la population *PA* – Nom, abréviation, origine, hauteur finale du brin maître (*H_F*) et état allélique pour les gènes *Rht1* et *Rht2*.

Nom	Abréviation	Origine d'obtention	<i>H_F</i> (cm)	<i>Rht1</i>	<i>Rht2</i>
Bizel	BIZ	SOGROUP (France)	88,4	haut	haut
Comtal	CTAL	RAGT (France)	95,5	haut	haut
Courtot	COURT	INRA (France)	56,5	court	court
Lutin	LUT	Ringot (France)	66,6	haut	haut
Maris Hobbit	MHO	PBI Cambridge (UK)	69,9	haut	haut
Zenith	ZEN	Inst Rech (Suisse)	95,2	haut	haut
C10-7-3	C1073	INRA (France)	68,8	haut	haut
(CxA)68.7.2.7	CA68	INRA (France)	81,4	haut	haut
(VPMxMoisson)4-1-2	VM4	INRA (France)	75,7	haut	haut
VPM1-1-1-2-R4	VP4	INRA (France)	116,9	haut	haut
(VPM1-1-1-2-R4 x C4261)C2429	C2429	INRA (France)	55,4	court	court
(VPM1-1-1-2-R4 x D65-5)11-3	V2D11	INRA (France)	81,3	haut	haut
C247-7-4	C247	INRA (France)	62,5	court	court
C248.8.1	C248	INRA (France)	66,4	court	court
C231-6-5	C231	INRA (France)	73,8	court	court
C6-2-4	C624	INRA (France)	80,9	haut	haut

Simulation du resemis

Le couplage entre WALTER et GenoWALT a ensuite été utilisé pour simuler le resemis de chaque *PA*₀ sur 20 générations, en conservant une taille de population constante (506 plantes) au cours des générations. À chaque génération, les 506 plantes ont été semées en maillage (distance inter-rang = distance entre plantes dans le rang) à une densité de 200 plantes/m². La densité de semis simulés (200 plantes/m²) est plus importante que la densité de semis utilisée expérimentalement (150 plantes/m²) car le coût computationnel diminue avec l'augmentation de la densité de semis, rendant irréalisable l'expérience à 150 plantes/m². Hormis leur différence de hauteur finale, les plantes composant le mélange étaient parfaitement identiques et leurs valeurs de paramètres génotypiques

i La panmixie correspond au croisement aléatoire des individus d'une population.

étaient celles estimées pour le génotype Maxwell (Annexe 1). La séquence climatique utilisée est une séquence moyenne obtenue à partir de données collectées sur 10 ans dans cinq localités de France, tel que décrit dans l'Annexe 2.

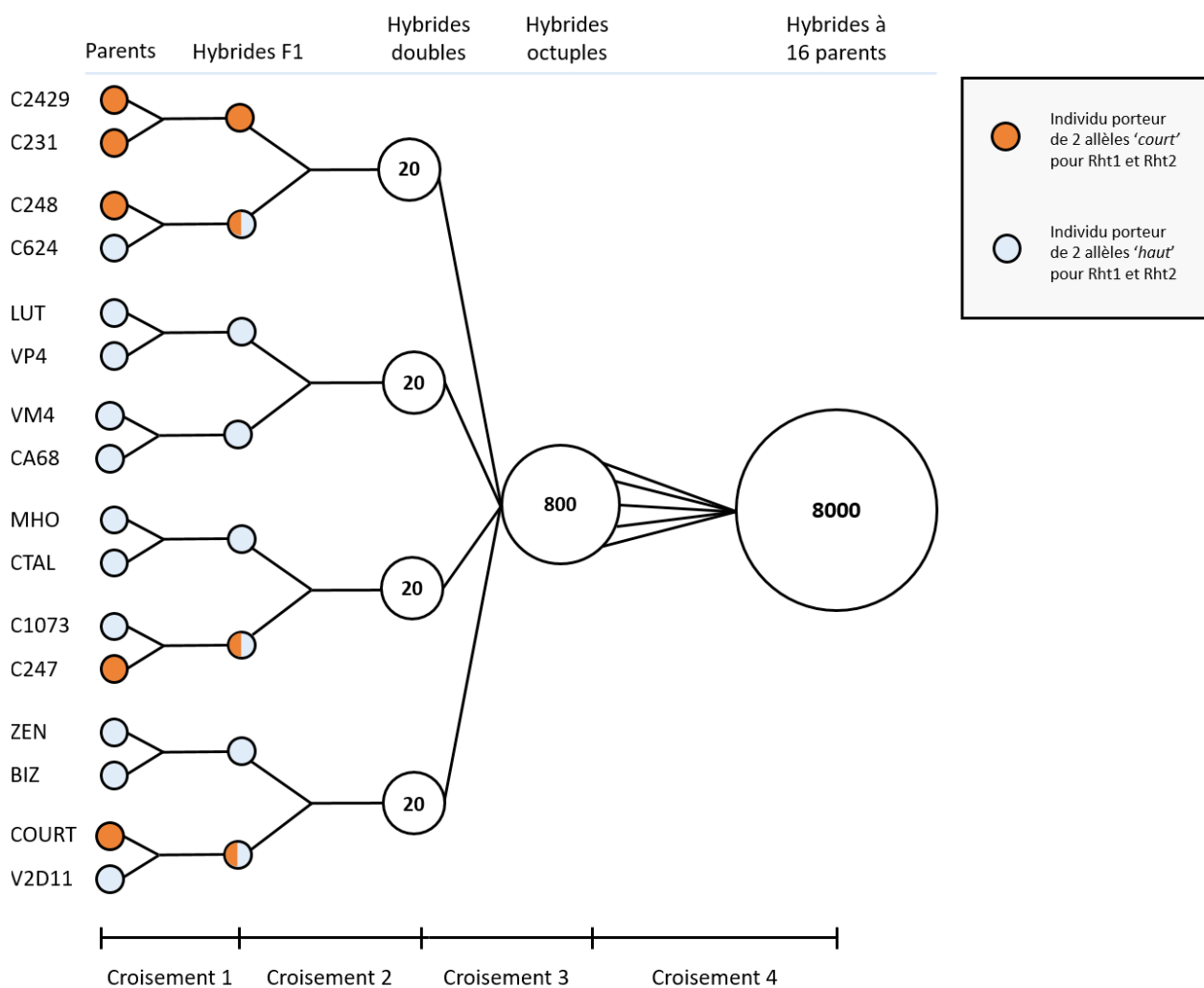


Figure 4. Schéma du croisement pyramidal permettant d'aboutir à la population PA (adapté de Thomas *et al.*, 1991)

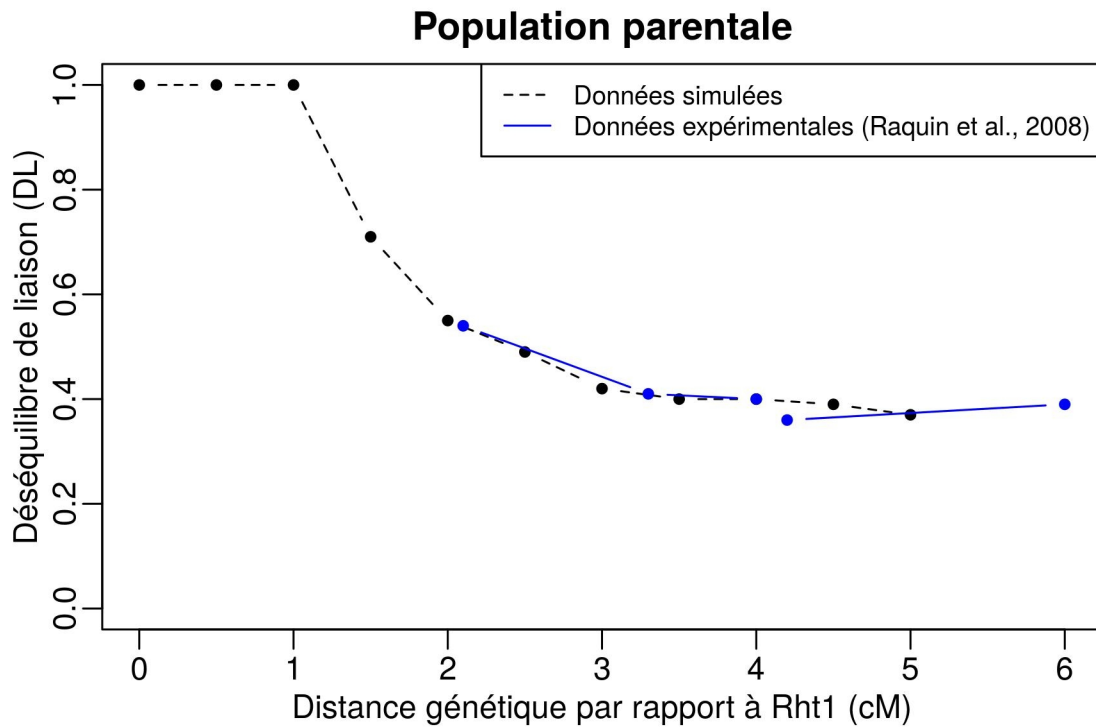


Figure 5. Déséquilibre de liaison chez les parents entre le gène *Rht1* (*Rht-B1*) et les marqueurs neutres adjacents, en fonction de leur distance génétique par rapport à *Rht1*. La courbe noire correspond aux données simulées et la courbe bleue correspond aux données expérimentales de Raquin *et al.*, 2008a. Les valeurs simulées pour les marqueurs neutres autour de *Rht2* (*Rht-D1*) sont identiques à celles simulées pour *Rht1*.

Analyses

Toutes les analyses ont été réalisées à l'aide du logiciel R (R Core Team, 2020). Les fréquences alléliques et le déséquilibre de liaison à chaque génération ont été calculés grâce au package Pegas (Paradis, 2010).

À partir des variations de fréquences alléliques à différents marqueurs, les effectifs efficaces ont été calculés pour différents intervalles de temps. L'effectif efficace d'une population correspond au nombre d'individus d'une population panmictique^j théorique qui présenterait les mêmes variations temporelles de fréquences alléliques que celles de la population observée, seulement à cause de la dérive génétique. Pour estimer l'effectif efficace, l'équation de Waples (1989) a été utilisée, en y retirant les termes d'échantillonnage car tous les individus de la population simulée sont utilisés pour le calcul des fréquences alléliques. Pour un intervalle de temps donné, l'effectif efficace de la population a ainsi été estimé via l'équation suivante :

$$N_e = \frac{t}{2F_c} \quad (\text{Équation 1}),$$

dans laquelle N_e correspond à l'effectif efficace estimé ; t correspond au nombre de générations dans l'intervalle de temps pour lequel N_e est calculé ; et F_c correspond à la variance moyenne de fréquence allélique parmi les différents loci utilisés, pour l'intervalle considéré.

Pour chaque locus utilisé, la variance de fréquence allélique a été calculée via l'équation suivante :

^j Une population panmictique est une population qui se reproduit en panmixie.

$$F_{cl} = \frac{1}{K} \sum_{i=1}^K \frac{(x_i - y_i)^2}{\frac{x_i + y_i}{2} - x_i y_i} \quad (\text{Équation 2}),$$

dans laquelle F_{cl} est la variance temporelle de fréquence allélique au locus l ; K est le nombre d'allèles pour le locus l ; x_i est la fréquence de l'allèle i au temps initial ; et y_i est la fréquence de l'allèle i au temps final.

RÉSULTATS

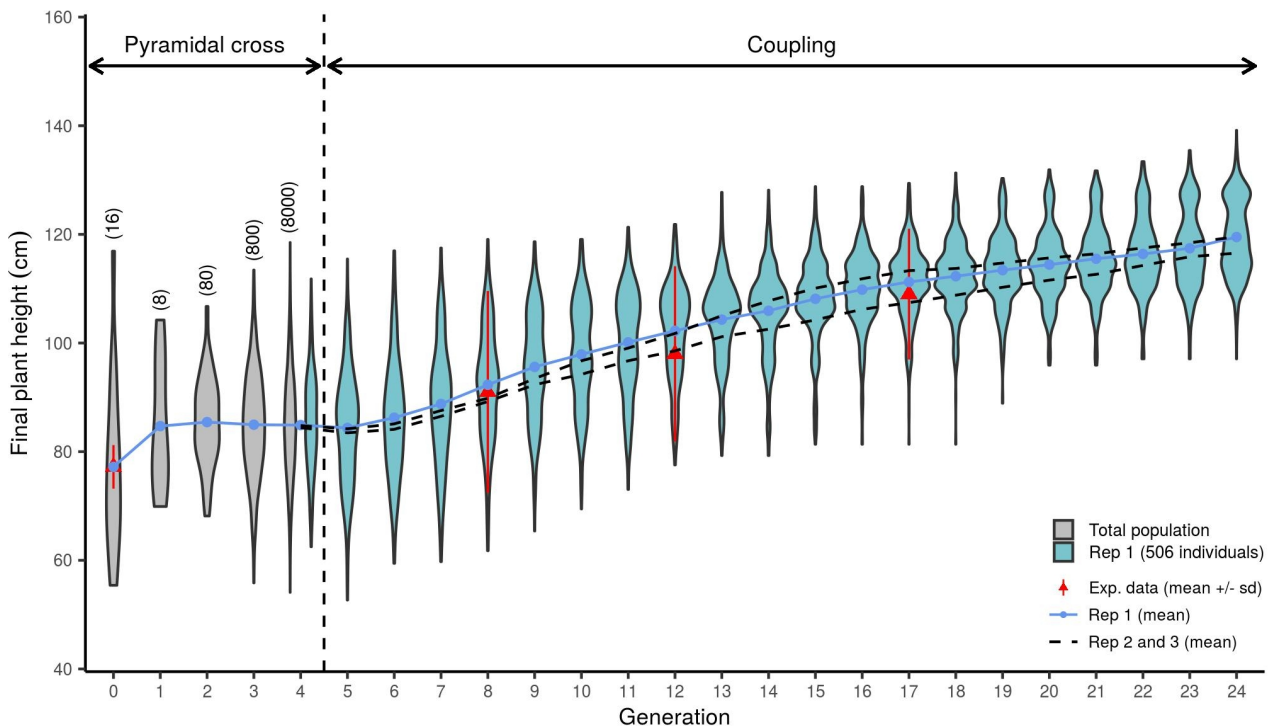


Figure 6. Evolution de la hauteur finale des plantes au cours des générations. La hauteur moyenne au cours des générations pour un des 3 réplicats simulés (rep 1) est représentée par une courbe bleue. La hauteur moyenne au cours des générations pour les 2 autres réplicats simulés est représentée par des courbes noires pointillées. Pour le croisement pyramidal, les distributions de hauteur pour chaque génération sont représentées en gris et le nombre d'individus pour chaque génération est indiqué entre parenthèses (cf. Figure 4). Les distributions de hauteur pour chaque génération du réplicat 1 sont représentées en bleu. Les données expérimentales disponibles dans l'étude de [Raquin et al. \(2008a\)](#) pour la hauteur moyenne et l'écart-type de hauteur sont représentés en rouge par des triangles et des segments respectivement.

Les simulations avec GenoWALT, pour le croisement pyramidal, puis avec le couplage, pour les générations de resemis, permettent d'étudier l'évolution de la hauteur au cours des générations (Figure 6). Le premier croisement (entre les générations G0 et G1) entraîne une augmentation importante (~7,5 cm) de la hauteur moyenne dans la population ; puis, dans la suite du croisement pyramidal, la hauteur moyenne de la population est relativement stable, autour de 85 cm. Pour ce qui est du couplage, les 3 réplicats donnent des résultats qualitativement identiques : après une très légère baisse lors de la première année de resemis (perte d'un peu moins de 1 cm entre les générations G4 et G5), la hauteur moyenne de PA augmente régulièrement au cours des générations pour atteindre près de 120 cm à la génération G24. Les distributions de hauteur au sein de la population pour chaque génération (Figure 6 et Annexe 4 – Figures 3 et 4) nous renseignent sur les

causes de l'augmentation de la hauteur moyenne au cours des générations. On note ainsi que les hauteurs maximales et minimales ainsi que la proportion de plantes hautes augmentent au sein de la population au cours des générations. L'augmentation de la hauteur moyenne au cours des générations s'accompagne d'une diminution de la variance de hauteur au sein de la population. Ces résultats sont en accord avec les observations faites dans la population expérimentale (Goldringer *et al.*, 2001a ; Raquin *et al.*, 2008a) : bien que les variances mesurées expérimentalement soient plus importantes que celles de la population simulée, la baisse de variance au cours des générations est du même ordre de grandeur pour l'expérimentation et pour les simulations. Pour ce qui est de la hauteur moyenne, les valeurs simulées correspondent tout à fait à celles mesurées expérimentalement.

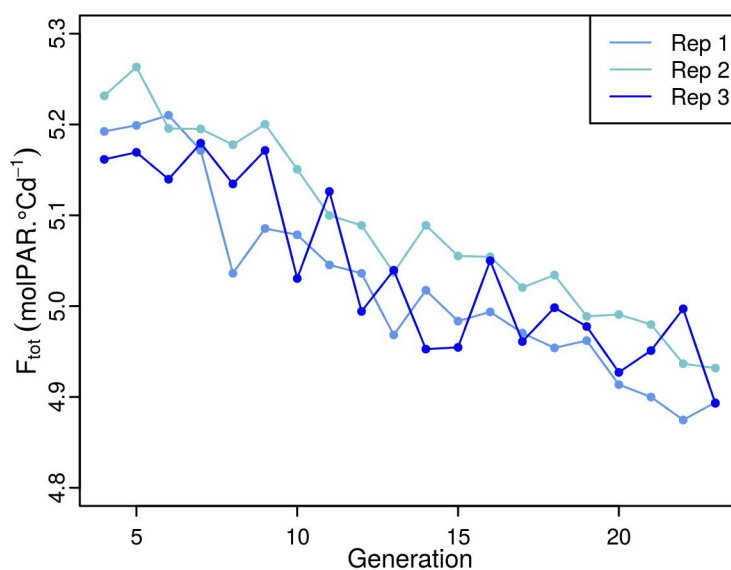


Figure 7. Evolution de la fitness totale de la parcelle (F_{tot}) au cours des générations. Les données simulées pour chaque réplicat sont représentées par des courbes à partir de la génération 4. WALTER n'étant pas utilisé pour le croisement pyramidal, F_{tot} n'est pas connu pour les générations précédentes.

L'utilisation de WALTER permet également d'étudier différents indicateurs de la performance de la population pour chaque génération. En particulier, la fitness totale de la population (F_{tot}) intègre des informations sur l'interception lumineuse de chaque axe pendant la période de détermination du nombre de grains (cf. Chapitre 1). Dans les simulations du resemis de PA, l'augmentation de la hauteur moyenne entre G5 et G23 s'accompagne d'une baisse de F_{tot} de 5,8% en moyenne (Figure 7). De la même manière, les différents indicateurs de performance mesurés dans la population expérimentale, et en particulier le nombre de grains par plante, ont diminué au cours des générations de resemis (Goldringer *et al.*, 2001a).

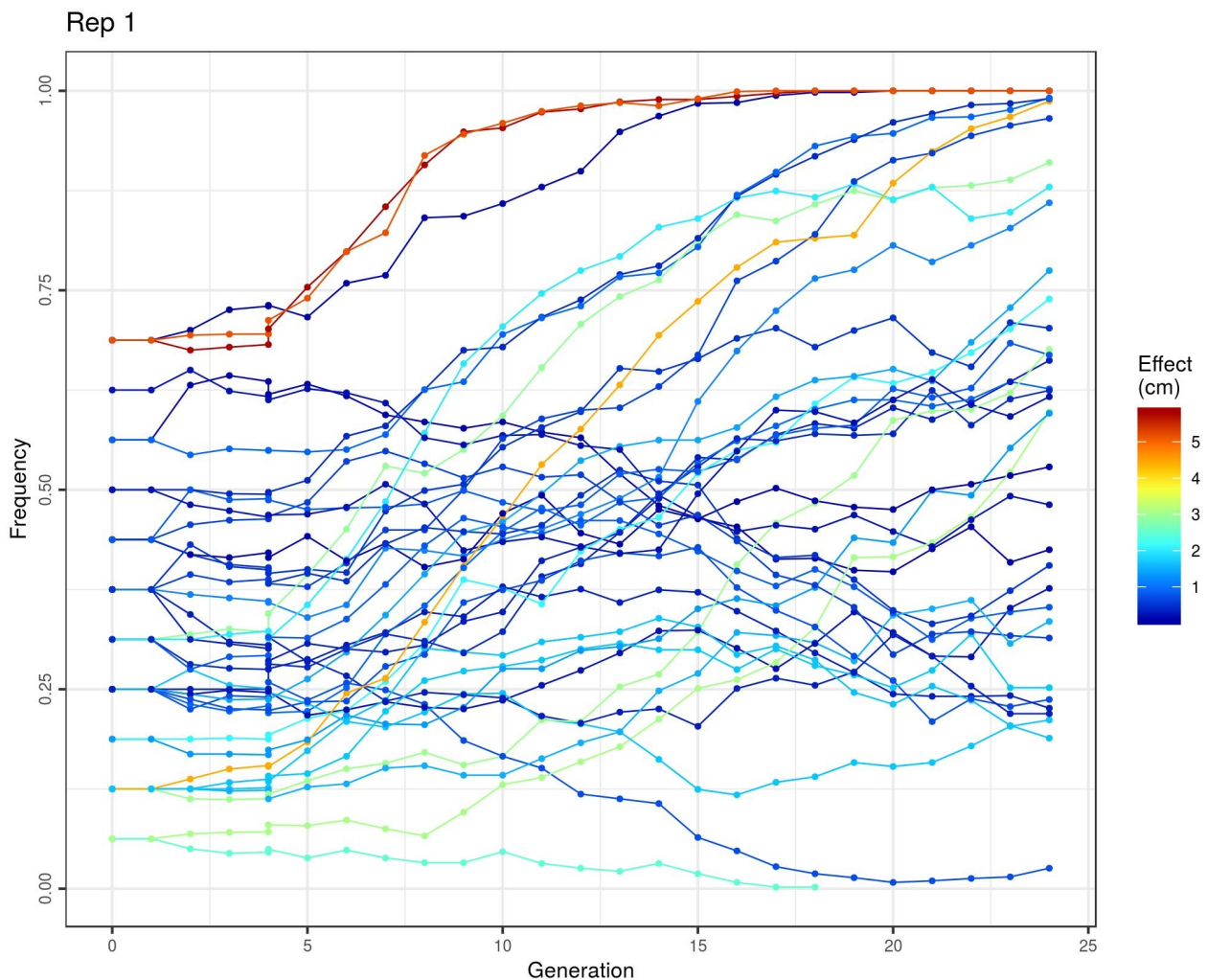


Figure 8. Evolution de la fréquence des allèles 'haut' de tous les gènes et QTL simulés au cours des générations pour un des 3 réplicats (Rep 1). La couleur des courbes indique l'effet de chaque marqueur sur la hauteur finale des plantes.

L'augmentation de hauteur observée pendant les générations de resemis (Figure 6) s'explique, à l'échelle génétique, par des variations importantes des fréquences alléliques pour les marqueurs contrôlant la hauteur finale des plantes (Figure 8 et Annexe 4 - Figures 5 et 6). Pour les marqueurs dont l'effet sur la hauteur est faible, l'évolution des fréquences alléliques ne suit pas de schéma particulier. En revanche, on note une augmentation importante de la fréquence des allèles 'haut' pour la plupart des marqueurs dont l'effet sur la hauteur est fort (effet > 2,5 cm). En particulier, la fréquence de l'allèle 'haut' du gène *Rht8* augmente fortement dans les 3 réplicats, passant de 0,15 environ à l'issue du croisement pyramidal (G4) à plus de 0,80 à la génération G24, et même à 0,99 pour l'un des 3 réplicats. La fréquence de l'allèle 'haut' augmente aussi pour les 2 gènes majeurs *Rht1* et *Rht2*, passant d'environ 0,69 en G4 à une fixation dans la population (fréquence de 1) pour tous les réplicats en G24. Cette fixation de l'allèle 'haut' intervient entre les génération G15 et G22 pour *Rht1*, selon les réplicats, et entre G16 et G23 pour *Rht2*. Bien que l'allèle 'haut' de *Rht1* augmente également en fréquence au cours des générations dans la population expérimentale jusqu'à être presque fixé à G24 (Raquin *et al.*, 2008a), on note que l'augmentation dans la population simulée est bien plus rapide que celle observée dans les données expérimentales (Figure 9).

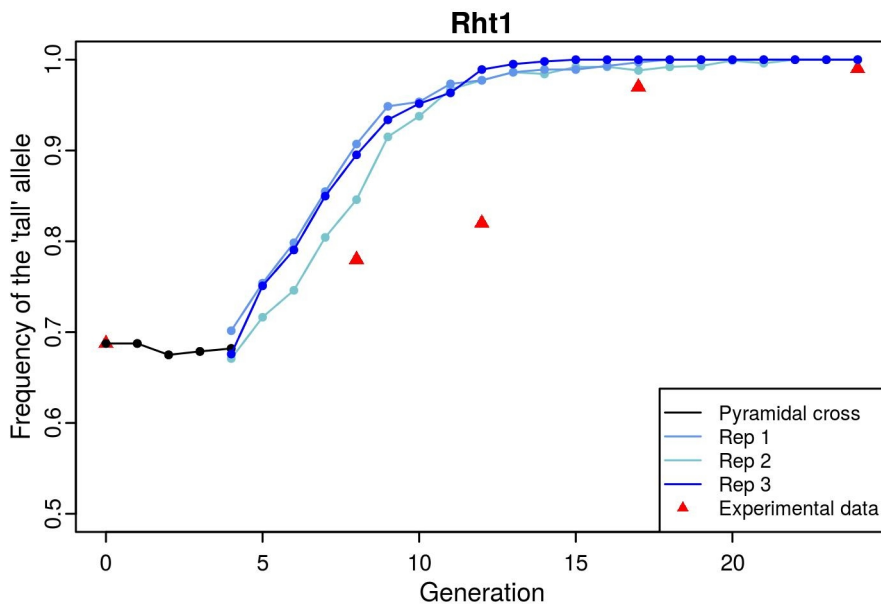


Figure 9. Evolution de la fréquence de l'allèle 'haut' du gène Rht1 au cours des générations. Les données simulées pour chaque réplikat sont représentées par des courbes et les données expérimentales sont représentées par des triangles.

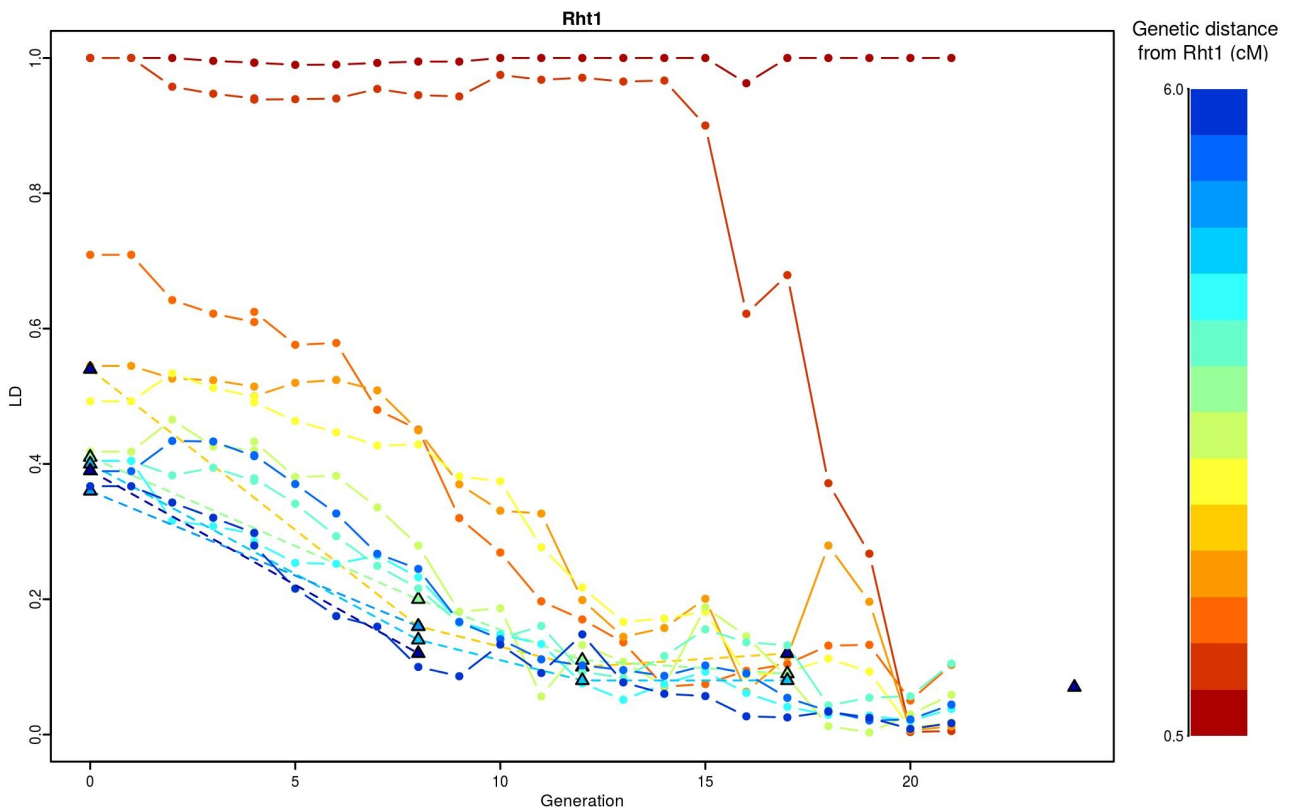


Figure 10. Déséquilibre de liaison moyen entre Rht1 et les marqueurs neutres adjacents en fonction de la génération. La distance par rapport à Rht1 est indiquée par la couleur. Les points reliés par des lignes continues correspondent au DL moyen pour les données simulées. Les triangles reliés par des lignes pointillées correspondent à des données expérimentales de Raquin *et al.* (2008a). À partir de la génération 22, l'allèle 'haut' du gène Rht1 est fixé dans la population pour les 3 réplikat, ce qui explique que le DL simulé ne puisse plus être calculé.

En plus des fréquences alléliques des marqueurs contrôlant la hauteur, le couplage de WALTER avec GenoWALT permet de suivre l'évolution des fréquences alléliques de marqueurs neutres (Annexe 4 – Figures 7, 8, 9). En suivant l'état allélique des marqueurs neutres proches de Rht1, il est ainsi possible de calculer le déséquilibre de liaison (DL) entre Rht1 et les marqueurs neutres adjacents

(Figure 10 et Annexe 4 – Figures 10, 11, 12). Le DL entre Rht1 et les marqueurs neutres situés entre 1,5 cM et 5 cM de ce gène majeur diminue entre G0 et G21, jusqu'à atteindre des valeurs proches de 0. À partir de la génération G22, l'allèle 'haut' de Rht1 étant fixé dans la population simulée pour les 3 réplicats, il n'est plus possible de calculer le DL entre Rht1 et les marqueurs adjacents. La diminution du DL observée dans les données simulées est en accord avec les données expérimentales (Raquin *et al.*, 2008a), pour lesquelles une diminution similaire du DL entre Rht1 et les marqueurs neutres adjacents a été observée. D'autre part, le DL entre Rht1 et les 2 marqueurs neutres les plus proches, initialement à 1 dans la population parentale, reste important au cours du croisement pyramidal et des générations de resemis. En particulier, le DL entre Rht1 et le marqueur neutre le plus proche, situé à 0,5 cM de Rht1, reste très proche de 1 pour toutes les générations. Le DL entre Rht1 et le marqueur neutre situé à 1 cM de ce gène majeur reste, quant à lui, supérieur à 0,9 tout au long des simulations pour 2 réplicats sur les 3. En revanche, pour le dernier réplicat, une diminution importante de ce DL est observée à partir de la génération G15.

Tableau 3. Effectifs efficaces calculés à partir des variations de fréquences alléliques de différents marqueurs, pour différentes périodes et pour chaque réplicat.

Période	Marqueurs	Rep 1	Rep 2	Rep 3
G0 - G8	50 marqueurs neutres répartis dans tous le génome	194	285	176
G8 - G12		66	118	92
G12 - G17		44	74	29
G17 - G24		81	60	76
G0 - G8	Rht1	14	30	16
G8 - G12		23	10	13
G12 - G17		82	360	115
G17 - G24		590	148	-
G0 - G8	Rht1 et 10 marqueurs neutres adjacents	33	75	36
G8 - G12		43	26	31
G12 - G17		37	55	21
G17 - G24		73	123	190

Les variations de fréquences alléliques pour les marqueurs neutres ou pour les marqueurs contrôlant la hauteur permettent également de calculer les effectifs efficaces (N_e) de la population (Tableau 3). Pour les deux premiers intervalles considérés (G0-G8 et G8-G12), les N_e estimés à partir de Rht1 (avec ou sans les marqueurs neutres adjacents) sont plus faibles que ceux estimés à partir des variations de fréquence alléliques des marqueurs neutres répartis sur tous le génome. En particulier, en moyenne, le N_e estimé à partir de 50 marqueurs neutres est dix fois plus élevé que celui estimé à partir de Rht1 et cinq fois plus élevé que celui estimé à partir de Rht1 et des marqueurs neutres adjacents pour la période allant de G0 à G8. Cela s'explique par les N_e très faibles estimés à partir de Rht1 (avec ou sans les marqueurs neutres adjacents) entre G0 et G12. Les N_e estimés entre G12 et G17 à partir de Rht1 et des marqueurs neutres adjacents sont également faibles (38 en moyenne). En revanche, pour cet intervalle (G12-G17), les N_e estimés à partir de Rht1 seul sont bien plus élevés (186 en moyenne) et sont même supérieurs à ceux estimés pour le même intervalle à partir de 50 marqueurs neutres (49 en moyenne). De même, les N_e estimés pour le dernier intervalle (G17-G24) sont plus élevés pour Rht1 (avec ou sans les marqueurs neutres adjacents) que pour les 50

marqueurs neutres répartis sur le génome.

DISCUSSION

Le couplage de WALTer avec le modèle de génétique des populations GenoWALT a permis de simuler l'évolution d'une population composite, paramétrée sur la base des caractéristiques de la population expérimentale PA (Thomas *et al.*, 1991), et présentant une importante diversité de hauteur. L'évolution phénotypique de la population simulée avec le couplage est comparable à celle de la population suivie expérimentalement. En effet, les simulations ont reproduit l'augmentation de hauteur moyenne ainsi que la baisse de variance de hauteur au sein de la population, au cours des générations (Goldringer *et al.*, 2001a ; Raquin *et al.*, 2008a). Au delà de cette tendance qualitative, il est intéressant de noter que l'augmentation de hauteur simulée est quantitativement comparable à celle mesurée expérimentalement, ce qui indique que le formalisme de fitness individuelle intégré à WALTer permet de bien retranscrire l'impact de différences de hauteur sur le succès reproducteur individuel, même dans des couverts très hétérogènes, comme c'est le cas ici. Étant donné que les plantes de la population expérimentale présentaient également des différences sur d'autres traits que la hauteur (résistance aux maladies, phénologie...), cette affirmation est cependant faite sous l'hypothèse que l'évolution de hauteur observée expérimentalement était due principalement à la compétition pour la lumière entre les individus et n'était pas, ou peu, impactée par une éventuelle sélection sur d'autres traits. D'autre part, dans les simulations, comme dans l'expérimentation, l'augmentation de hauteur au cours des générations s'accompagne d'une baisse de performance à l'échelle du peuplement (Goldringer *et al.*, 2001a). Cette observation est en accord avec la théorie de la tragédie des biens communs (TOC pour "*tragedy of the commons*" en anglais) (Anten et Vermeulen, 2016). Cette théorie repose sur l'idée qu'il existe un trade-off entre la performance individuelle et la performance à l'échelle du peuplement. Ce compromis a d'ailleurs été mis en évidence expérimentalement dans l'étude de Weiner *et al.* (2017) chez le blé. Ainsi, pour ce qui est de la hauteur, les plantes les plus hautes sont les plus performantes dans les parcelles hétérogènes mais, en culture pure, ces plantes sont moins performantes que des plantes plus courtes et moins compétitives. La concordance entre l'augmentation de la hauteur moyenne dans la parcelle simulée et la baisse de F_{tot} au cours des générations est également en accord avec les résultats des simulations du Chapitre 5, qui ont montré une corrélation négative entre la hauteur moyenne de la parcelle et sa fitness totale. Cette évolution de la hauteur est donc problématique pour la culture des populations composites, de part la baisse de performance associée, mais également de part la diminution de diversité qu'elle entraîne. Ces problématiques ont été soulevées par plusieurs auteurs (Le Boulc'h *et al.*, 1994 ; Goldringer *et al.*, 2001b ; Döring *et al.*, 2015 ; Knapp *et al.*, 2020) et des solutions comme la baisse de la densité de semis ou la sélection humaine ont été proposées. Le couplage entre WALTer et GenoWALT serait également un outil intéressant pour tester ces pratiques. En particulier, il est possible de faire varier la densité de semis dans les simulations afin d'identifier la densité optimale pour limiter la compétition pour la lumière entre les plantes et maintenir une diversité de hauteur, tout en préservant des rendements suffisants.

En plus de l'évolution phénotypique de la population, le couplage a permis de simuler son évolution génétique. Dans les simulations, l'évolution des fréquences alléliques pour les marqueurs contrôlant

la hauteur est conforme à ce qui est attendu sous l'hypothèse d'une sélection pour la hauteur, avec une augmentation de fréquence des allèles 'haut' d'autant plus forte que l'effet du marqueur est important. En effet, on s'attend à ce que la dérive génétique ait un fort impact sur les marqueurs à effet plus faible, ce qui explique l'évolution aléatoire de leurs fréquences alléliques. Pour la population simulée, l'augmentation importante de la fréquence de l'allèle 'haut' du gène majeur Rht1, jusqu'à sa fixation dans la population après 20 générations de resemis, est qualitativement conforme à ce qui a été observé expérimentalement (Raquin *et al.*, 2008a). Cependant, cette augmentation de fréquence est plus rapide dans la population simulée que dans la population expérimentale. Il semblerait ainsi que la pression de sélection sur Rht1 ait été plus forte dans nos simulations que dans l'expérimentation. On peut donc penser qu'un ou plusieurs autres gènes ou QTL ont subi une pression de sélection importante dans la population expérimentale pour expliquer l'évolution de hauteur mesurée au cours des générations. Cette information nous renseigne sur la structure génétique des individus composant la population expérimentale : il existe vraisemblablement des différences entre les génotypes simulés à partir de notre carte génétique hypothétique et les génotypes des individus de la population expérimentale, pour lesquels seuls les états alléliques aux gènes majeurs Rht1 et Rht2 étaient connus. En plus de l'évolution des fréquences alléliques pour les marqueurs contrôlant la hauteur, le couplage permet également de simuler l'évolution de marqueurs neutres. Il a ainsi été possible de mesurer le déséquilibre de liaison (DL) entre Rht1 et les marqueurs neutres adjacents. À l'exception du marqueur neutre le plus proche de Rht1, pour lequel le DL initial était maximum et la probabilité de recombinaison proche de 0, les simulations ont ainsi mis en évidence une diminution du DL au cours des générations de resemis. Ces résultats sont en accord avec les mesures effectuées dans la population expérimentale, qui indiquent qu'il n'y a pas d'effet auto-stop pour les marqueurs autour de Rht1. L'évolution de PA correspond ainsi à un balayage évolutif doux (Raquin *et al.*, 2008a), et seuls des marqueurs très proches de Rht1, pour lesquels le DL initial serait très fort, pourraient subir un effet auto-stop dans cette population. La capacité du couplage à simuler ce type d'éléments est très intéressante pour mieux comprendre la façon dont la sélection naturelle agit indirectement sur les zones génomiques neutres dans les populations composites et ainsi éviter d'éventuelles pertes de diversité ou la fixation d'allèles délétères par l'effet auto-stop (Goldringer *et al.*, 2001b). En plus de l'évolution du DL entre des marqueurs sous sélection et des marqueurs neutres adjacents, la simulation de l'évolution temporelle des fréquences alléliques pour les marqueurs neutres permet d'estimer l'effectif efficace de la population simulée. L'effectif efficace est un outil très utile pour estimer le potentiel évolutif d'une population et ainsi prédire l'évolution de sa diversité génétique. En effet, le risque qu'une population perde rapidement en diversité ou accumule des allèles délétères est d'autant plus important que son effectif efficace est faible (Bataillon et Kirkpatrick, 2000). De ce fait, l'estimation de l'effectif efficace est un élément central de l'étude des populations menées en gestion dynamique (Brumlop *et al.*, 2019). Dans la simulation de PA réalisée grâce au couplage, les effectifs efficaces estimés sont conformes à ce qui est attendu dans le cas d'une sélection directionnelle pour la hauteur, comme cela a été observé dans la population expérimentale (Raquin *et al.*, 2008a). En effet, les effectifs efficaces estimés à partir de marqueurs sous sélection sont plus faibles que ceux estimés à partir des variations de fréquence de marqueurs neutres pour les 12 premières générations de resemis. Pour les générations suivantes, les effectifs efficaces élevés, estimés à partir des variations de fréquence allélique de Rht1, sont en accord avec la fixation de l'allèle 'haut' de ce gène dans la population. Contrairement aux estimations basées sur Rht1, qui sont

faibles du fait de la sélection plus importante durant les premières années de resemis, les effectifs efficaces élevés estimés pour les premières générations à partir des variations de fréquence des marqueurs neutres rendent bien compte du brassage génétique important ayant eu lieu lors du croisement pyramidal. Les effectifs efficaces estimés à partir des simulations sont donc cohérents, ce qui confirme l'intérêt d'utiliser le couplage entre WALTer et GenoWALT pour estimer cet indice. En particulier, les simulations permettent de contrôler facilement la taille de la population (nombre d'individus) et sa variation dans le temps, qui sont des facteurs ayant beaucoup d'impact sur l'effectif efficace d'une population (Frankham, 1995). Ici, la taille constante de PA tout au long des générations de resemis permet ainsi de s'assurer que les faibles valeurs d'effectif efficace estimées sont bien liées à des effets de sélection. Ces faibles valeurs soulignent l'importance de cultiver des populations avec un nombre d'individus suffisant et de prendre en compte le contrôle de l'évolution de la hauteur et d'autres traits impliqués dans la compétition pour la lumière dans les programmes de gestion dynamique (Goldringer *et al.*, 2001a).

Le couplage a donc permis de simuler de façon réaliste l'évolution phénotypique et génétique d'une population hétérogène complexe menée en gestion dynamique, ce qui confirme son intérêt pour explorer les mécanismes prenant place dans ces populations et ainsi améliorer notre compréhension de leur évolution. Cependant, le nombre de réplicats réalisés (3) est faible et il serait intéressant de l'augmenter, d'autant que la taille des populations simulées est également faible (506 individus), ce qui donne plus de poids à la dérive génétique. D'autre part, en plus des limites de WALTer décrites dans le Chapitre 1, les limites de GenoWALT doivent être prises en considération lors de l'utilisation du couplage. Ainsi, il est intéressant de noter que les marqueurs génétiques simulés ne peuvent présenter que 2 allèles dans la version actuelle de GenoWALT, alors que les marqueurs neutres utilisés dans l'étude expérimentale de PA (microsatellites) présentaient un polymorphisme plus important (Raquin *et al.*, 2008a). De même, la carte génétique simulée ne permet pas de considérer des régimes de dominance partielle, bien que de tels régimes de dominances aient été mis en évidence dans le déterminisme de la hauteur chez le blé (Gale *et al.*, 1985). Cependant, le gain de précision apporté par l'utilisation d'un modèle prenant en compte des effets de dominance partielle serait probablement marginal (Liu *et al.*, 2016). Les effets d'épistasie (interactions entre marqueurs), connus pour jouer un rôle dans le déterminisme des traits quantitatifs comme la hauteur (Carlborg et Haley, 2004 ; Zhang *et al.*, 2008 ; Würschum *et al.*, 2015), ne sont pas non plus considérés dans le modèle. Bien que des taux de mutation aient été estimés pour la population expérimentale (Raquin *et al.*, 2008b), les mutations ne sont pas non plus simulées par GenoWALT. Enfin, les migrations ne peuvent pas être simulées par le modèle. Bien qu'aucune migration ne soit considérée dans l'étude présentée ici, il serait intéressant d'intégrer cette option au modèle car l'intérêt des migrations pour le maintien de la diversité dans les populations de gestion dynamique fait l'objet de travaux, notamment pour la population PA (Rousselle *et al.*, 2011).

Néanmoins, de nombreuses possibilités sont offertes par le couplage dans sa version actuelle. Il serait ainsi intéressant de comparer les résultats de l'étude présentée ici avec les résultats de simulations utilisant GenoWALT en association avec une fonction de fitness simple à la place de WALTer. En effet, l'évolution de la population simulée dépend de la compétition pour la lumière entre des individus qui présentent uniquement des différences de hauteur. Il est donc possible que cette évolution puisse être reproduite en utilisant une fonction de fitness plus simple, comme celle proposée par Montazeau *et al.* (2020) par exemple. Étant donné que la spécificité de WALTer est de

simuler la plasticité du tallage, la comparaison de nos résultats avec ceux obtenus à partir d'une fonction plus simple donnerait alors des informations sur l'importance de la plasticité du tallage dans les dynamiques observées. D'autre part, il serait particulièrement intéressant d'utiliser le couplage de GenoWALT et WALTer pour simuler des populations présentant une diversité sur plusieurs traits d'architecture aérienne (hauteur, tallage, dimensions et angle foliaire...). En effet, l'utilisation d'un FSPM est particulièrement adaptée pour simuler ce type de couverts complexes. Il serait alors possible de tester si certaines combinaisons de traits permettent de maintenir une diversité au cours des générations en limitant les effets de sélection directionnelle. Enfin, il serait intéressant d'utiliser le couplage pour étudier l'évolution de populations hétérogènes dans lesquelles le taux d'allogamie est plus important. En effet, bien que le blé soit principalement autogame, l'allogamie joue un rôle important dans l'évolution des populations composites (Enjalbert *et al.*, 1998). C'est pourquoi plusieurs programmes se sont intéressés à l'évolution de populations expérimentales dans lesquelles un gène de stérilité mâle permettait d'augmenter le taux d'allogamie et ainsi d'augmenter encore la diversité génétique intra-parcelle (Döring *et al.*, 2015 ; Goldringer *et al.*, 2001b). Le couplage entre WALTer et GenoWALT permet de simuler ce type de population et ainsi d'améliorer notre compréhension de leur évolution.

RÉFÉRENCES

- Anten, N. P., & Vermeulen, P. J. (2016).** Tragedies and crops: understanding natural selection to improve cropping systems. *Trends in Ecology & Evolution*, 31(6), 429-439.
- Bataillon, T., & Kirkpatrick, M. (2000).** Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genetics research*, 75(1), 75-81.
- Brumlop, S., Weedon, O., Link, W., & Finckh, M. R. (2019).** Effective population size (N_e) of organically and conventionally grown composite cross winter wheat populations depending on generation. *European Journal of Agronomy*, 109, 125922.
- Carlborg, Ö., & Haley, C. S. (2004).** Epistasis: too often neglected in complex trait studies?. *Nature Reviews Genetics*, 5(8), 618-625.
- Döring, T. F., Annicchiarico, P., Clarke, S., Haigh, Z., Jones, H. E., Pearce, H., ... & Wolfe, M. S. (2015).** Comparative analysis of performance and stability among composite cross populations, variety mixtures and pure lines of winter wheat in organic and conventional cropping systems. *Field Crops Research*, 183, 235-245.
- Dvořák, J., Terlizzi, P. D., Zhang, H. B., & Resta, P. (1993).** The evolution of polyploid wheats: identification of the A genome donor species. *Genome*, 36(1), 21-31.
- Enjalbert, J., Dawson, J. C., Paillard, S., Rhoné, B., Rousselle, Y., Thomas, M., & Goldringer, I. (2011).** Dynamic management of crop diversity: From an experimental approach to on-farm conservation. *Comptes rendus biologies*, 334(5-6), 458-468.
- Enjalbert, J., Goldringer, I., David, J., & Brabant, P. (1998).** The relevance of outcrossing for the dynamic management of genetic resources in predominantly selfing *Triticum aestivum* L. (bread wheat). *Genetics Selection Evolution*, 30(1), 1-15.
- Enjalbert, J., Goldringer, I., Paillard, S., & Brabant, P. (1999).** Molecular markers to study genetic drift and selection in wheat populations. *Journal of experimental botany*, 50(332), 283-290.
- Frankham, R. (1995).** Effective population size/adult population size ratios in wildlife: a review. *Genetics Research*, 66(2), 95-107.
- Gale, M. D., Youssefian, S., & Russell, G. E. (1985).** Dwarfing genes in wheat. *Progress in plant breeding*, 1, 1-35.
- Godin, C., Sinoquet, H. (2005).** Functional–structural plant modelling. *New phytologist*, 166(3), 705-708.
- Goldringer, I., Bonnin, I., & David, J. (2013).** Rôle de la sélection dans la dynamique de la diversité des Blés cultivés: de la domestication à la sélection moderne au 20ème siècle. *Innovations Agronomiques*, 29, 1-15.
- Goldringer, I., Enjalbert, J., David, J., Paillard, S., Pham, J. L., & Brabant, P. (2001b).** Dynamic management of genetics resources: a 13-year experiment on wheat. *Broadening the genetic base of crop production. IPGRI/FAO*, 245-260.

- Goldringer, I., Enjalbert, J., Raquin, A. L., & Brabant, P. (2001a).** Strong selection in wheat populations during ten generations of dynamic management. *Genetics Selection Evolution*, 33(1), S441-S463.
- Goldringer, I., Prouin, C., Rousset, M., Galic, N., & Bonnin, I. (2006).** Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. *Annals of Botany*, 98(4), 805-817.
- Griffiths, S., Simmonds, J., Leverington, M., Wang, Y., Fish, L., Sayers, L., ... & Snape, J. (2012).** Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. *Molecular Breeding*, 29(1), 159-171.
- Henry, J. P., Pontis, C., David, J., & Gouyon, P. H. (1991).** An experiment on dynamic conservation of genetic resources with metapopulations. In *Species conservation: a population-biological approach* (pp. 185-198). Birkhäuser, Basel.
- Isbell, F. (2015).** Agroecology: agroecosystem diversification. *Nature plants*, 1(4), 1-2.
- Khan, A. R., Goldringer, I., & Thomas, M. (2020).** Management Practices and Breeding History of Varieties Strongly Determine the Fine Genetic Structure of Crop Populations: A Case Study Based on European Wheat Populations. *Sustainability*, 12(2), 613.
- Kilian, B., Özkan, H., Deusch, O., Effgen, S., Brandolini, A., Kohl, J., ... & Salamini, F. (2007).** Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. *Molecular Biology and Evolution*, 24(1), 217-227.
- Knapp, S., Döring, T. F., Jones, H. E., Snape, J., Wingen, L. U., Wolfe, M. S., ... & Griffiths, S. (2020).** Natural selection towards wild-type in composite cross populations of winter wheat. *Frontiers in Plant Science*, 10, 1757.
- Le Boulc'h, V., David, J. L., Brabant, P., & de Vallavieille-Pope, C. (1994).** Dynamic conservation of variability: responses of wheat populations to different selective forces including powdery mildew. *Genetics Selection Evolution*, 26(1), 1-20.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Liu, G., Zhao, Y., Gowda, M., Longin, C. F. H., Reif, J. C., & Mette, M. F. (2016).** Predicting hybrid performances for quality traits through genomic-assisted approaches in central European wheat. *PLoS One*, 11(7), e0158635.
- Malézieux, E. (2012).** Designing cropping systems from nature. *Agronomy for sustainable development*, 32(1), 15-29.
- Montazeaud, G., Rousset, F., Fort, F., Violle, C., Fréville, H., & Gandon, S. (2020).** Farming plant cooperation in crops. *Proceedings of the Royal Society B*, 287(1919), 20191290.
- Paillard, S., Goldringer, I., Enjalbert, J., Doussinault, G., de Vallavieille-Pope, C., & Brabant, P. (2000a).** Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. I—Is specific seedling resistance selected?. *Theoretical and Applied Genetics*, 101(3), 449-456.

- Paillard, S., Goldringer, I., Enjalbert, J., Trottet, M., David, J., de Vallavieille-Pope, C., & Brabant, P. (2000b).** Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. II. Adult plant resistance. *Theoretical and Applied Genetics*, *101*(3), 457-462.
- Paradis, E. (2010).** pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419-420.
- Peng, B., & Kimmel, M. (2005).** simuPOP: a forward-time population genetics simulation environment. *Bioinformatics*, *21*(18), 3686-3687.
- R Core Team (2020).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raquin, A. L., Brabant, P., Rhoné, B., Balfourier, F., Leroy, P., & Goldringer, I. (2008a).** Soft selective sweep near a gene that increases plant height in wheat. *Molecular Ecology*, *17*(3), 741-756.
- Raquin, A. L., Depaulis, F., Lambert, A., Galic, N., Brabant, P., & Goldringer, I. (2008b).** Experimental estimation of mutation rates in a wheat population with a gene genealogy approach. *Genetics*, *179*(4), 2195-2211.
- Rhone, B., Remoue, C., Galic, N., Goldringer, I., & Bonnin, I. (2008).** Insight into the genetic bases of climatic adaptation in experimentally evolving wheat populations. *Molecular Ecology*, *17*(3), 930-943.
- Rhoné, B., Vitalis, R., Goldringer, I., & Bonnin, I. (2010).** Evolution of flowering time in experimental wheat populations: a comprehensive approach to detect genetic signatures of natural selection. *Evolution: International Journal of Organic Evolution*, *64*(7), 2110-2125.
- Rousselle, Y., Thomas, M., Galic, N., Bonnin, I., & Goldringer, I. (2011).** Inbreeding depression and low between-population heterosis in recently diverged experimental populations of a selfing species. *Heredity*, *106*(2), 289-299.
- Sarkar, P., & Stebbins, G. L. (1956).** Morphological evidence concerning the origin of the B genome in wheat. *American Journal of Botany*, 297-304.
- Somers, D. J., Isaac, P., & Edwards, K. (2004).** A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and applied genetics*, *109*(6), 1105-1114.
- Thomas, G. (1986).** Thèse de doctorat, Université de Rennes 1.
- Thomas, G., Rousset, M., Pichon, M., Trottet, M., Doussinault, G., & Picard, E. (1991).** Méthodologie de l'amélioration de blé tendre (*Triticum aestivum* L.). I. Création par croisements et analyse d'une population artificielle à 16 parents, base de cette étude méthodologique. *Agronomie*, *11*(5), 359-368.
- Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & De Visser, P. H. (2010).** Functional-structural plant modelling: a new versatile tool in crop science. *Journal of experimental Botany*, *61*(8), 2101-2115.

- Wang, J., Luo, M. C., Chen, Z., You, F. M., Wei, Y., Zheng, Y., & Dvorak, J. (2013).** Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New phytologist*, 198(3), 925-937.
- Waples, R. S. (1989).** A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, 121(2), 379-391.
- Weiner, J., Du, Y. L., Zhang, C., Qin, X. L., & Li, F. M. (2017).** Evolutionary agroecology: individual fitness and population yield in wheat (*Triticum aestivum*). *Ecology*, 98(9), 2261-2266.
- Würschum, T., Langer, S. M., & Longin, C. F. H. (2015).** Genetic control of plant height in European winter wheat cultivars. *Theoretical and Applied Genetics*, 128(5), 865-874.
- Würschum, T., Liu, G., Boeven, P. H., Longin, C. F. H., Mirdita, V., Kazman, E., ... & Reif, J. C. (2018).** Exploiting the Rht portfolio for hybrid wheat breeding. *Theoretical and Applied Genetics*, 131(7), 1433-1442.
- Zhang, K., Tian, J., Zhao, L., & Wang, S. (2008).** Mapping QTLs with epistatic effects and QTL× environment interactions for plant height using a doubled haploid population in cultivated wheat. *Journal of Genetics and Genomics*, 35(2), 119-127.
- Zhao, Y., Mette, M. F., Gowda, M., Longin, C. F. H., & Reif, J. C. (2014).** Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity*, 112(6), 638-645.

DISCUSSION GÉNÉRALE

Apports de la thèse pour la modélisation structure-fonction des peuplements hétérogènes de blé

Les différents chapitres de cette thèse présentent le développement et l'utilisation d'outils et de méthodes de modélisation structure-fonction pour étudier l'impact de l'architecture aérienne sur la compétition pour la lumière dans des peuplements hétérogènes de blé et les conséquences de cette compétition sur l'évolution des peuplements. La modélisation structure-fonction est particulièrement adaptée à l'étude des peuplements hétérogènes, mais cette approche est encore peu développée à ce jour (Gaudio *et al.*, 2019 ; Evers *et al.*, 2019 ; Louarn et Song, 2020). Pour le blé, le FSPM ADEL-blé (Evers *et al.*, 2010) peut être utilisé pour simuler la compétition pour la lumière et la photosynthèse dans des mélanges de variétés. Le couplage entre CN-Wheat et ADEL-blé (Barillot *et al.*, 2019) permet de prendre en compte les interactions entre le métabolisme du carbone et de l'azote, l'architecture des plantes et de nombreux facteurs environnementaux, dont la lumière. Si ce couplage simule uniquement la période post-floraison, il a été étendu par Gauthier *et al.* (2020) pour simuler le développement du semis à la montaison. Les travaux présentés dans cette thèse ont permis d'adapter le FSPM WALTer (Lecarpentier *et al.*, 2019) avec des objectifs distincts et complémentaires des modèles listés ci-dessus. L'utilisation de règles de développement simplifiées à l'origine du modèle (Lecarpentier *et al.*, 2019) et les améliorations apportées à WALTer dans cette thèse (Chapitre 1) rendent possible son utilisation pour simuler des peuplements hétérogènes très complexes, dans lesquels le nombre de génotypes est illimité et tous les paramètres génotypiques peuvent différer entre génotypes. À notre connaissance, WALTer est ainsi le seul FSPM permettant de simuler les 3 étapes du tallage (émission, arrêt et régression) et leur plasticité en réponse à la compétition pour la lumière dans des peuplements hétérogènes de blé. Grâce aux améliorations apportées à WALTer pour réduire le coût computationnel de ses simulations, le modèle est d'autant plus complémentaire des autres FSPM de blé que son temps de simulation est bien moins important. Ainsi, la simulation de la période post-floraison pour un épi prend environ 30 minutes avec le couplage proposé par Barillot *et al.* (2019). Avec la nouvelle version de WALTer, une simulation de 30 minutes permet de représenter le développement de 100 plantes du semis à la maturité, à une densité de 200 plantes/m² (Chapitre 2). Ces différences importantes de temps de simulation entre les 2 modèles s'expliquent par la plus grande complexité du modèle de Barillot *et al.* (2019), qui permet de simuler des processus plus poussés que WALTer (notamment le métabolisme du carbone et de l'azote), mais elles soulignent la complémentarité des 2 modèles, qui sont destinés à des utilisations différentes. Ainsi, il est notamment possible de simuler des populations composites de plusieurs centaines d'individus, tous différents, avec WALTer (Chapitre 6).

Les améliorations apportées à WALTer dans le cadre de cette thèse permettent donc de proposer un nouvel outil de modélisation structure-fonction pour l'étude des peuplements hétérogènes de blé. De plus, la compréhension du fonctionnement de WALTer a été améliorée par l'étude de l'impact de l'aléatoire dans le modèle (Chapitre 2) et son analyse de sensibilité dans le cadre de la simulation de peuplements purs (Chapitre 3). Ces deux étapes nous renseignent notamment sur l'importance des différents paramètres du modèle et sur les temps de simulation. Elles sont des étapes essentielles pour une utilisation éclairée de WALTer, en particulier pour la simulation de peuplements hétérogènes complexes, pour lesquels l'analyse des résultats requiert une bonne compréhension du

fonctionnement du modèle. De plus, les Chapitres 4 et 6 ont permis de confronter les simulations de WALTer, et du couplage de WALTer et GenoWALT, à des données expérimentales sur la dynamique de tallage de peuplements purs et sur l'évolution génétique de populations composites. Ces comparaisons sont très informatives et ont mis en avant que les simulations reproduisent correctement les données réelles, indiquant un bon fonctionnement du modèle, utilisé seul ou en couplage. Il est ainsi possible de s'appuyer sur les éléments de compréhension acquis dans cette thèse pour toutes les futures simulations de WALTer, et pour d'éventuels développements du modèle. Les futurs développements de WALTer, sa maintenance et sa diffusion sont également simplifiés par l'intégration du modèle à l'incubateur OpenAlea (Pradal *et al.*, 2008). Cette intégration est d'autant plus intéressante que ces différents éléments (développement, maintenance, diffusion) sont généralement plus lourds pour des modèles complexes comme les FSPM (Louarn et Song, 2020).

Si les simulations de WALTer sont relativement rapides pour un FSPM, la problématique du coût computationnel élevé des FSPM, soulevée par Louarn et Song (2020), demeure et a pu être traitée par plusieurs moyens dans cette thèse. Tout d'abord, le temps de simulation a été réduit grâce aux améliorations apportées au modèle (Chapitre 1). De plus, le nombre de plantes à simuler a été choisi pour limiter la durée des simulations tout en assurant la robustesse des résultats (Chapitre 2). L'étude de l'aléatoire du Chapitre 2 a également montré que l'on pouvait se passer de réplicats pour les Chapitres 4 et 5, limitant ainsi beaucoup le temps de simulation. Ce dernier a également été limité par l'utilisation de méthodes et de plans d'expérience adaptés aux différents objectifs, comme le plan factoriel fractionnaire utilisé pour l'analyse de sensibilité du Chapitre 3. Enfin, nos travaux ont montré l'utilité et l'efficacité de la construction de métamodèles pour approcher les sorties de WALTer. L'utilisation d'un métamodèle par krigeage a ainsi permis d'utiliser des méthodes qui demandent un nombre important d'appels au modèle et qui n'auraient donc pas pu être utilisées directement avec un FSPM (Chapitres 4 et 5). La construction de métamodèles a parfois été utilisée avec des FSPM pour la simulation de plantes isolées ou de couverts homogènes dans des approches d'analyse de sensibilité (Da Silva *et al.*, 2014 ; Perez *et al.*, 2018). Pour WALTer, les travaux présentés dans cette thèse ont montré l'intérêt du krigeage dans le cadre de la simulation de couverts homogènes mais également pour la simulation de mélanges binaires équilibrés présentant une diversité d'architecture importante. La simulation des mélanges étant plus longue que la simulation de peuplements purs (du fait du plus grand nombre de plantes requis), la bonne performance du krigeage pour approximer les sorties de WALTer pour des mélanges binaires est particulièrement intéressante. Les Chapitres 4 et 5 proposent également une approche couplant le krigeage avec un plan d'expérience séquentiel pour la calibration de WALTer ou l'optimisation de la performance de mélanges binaires. Bien que ces méthodes soient utilisées avec des modèles complexes (Carmassi *et al.*, 2018), elles n'avaient pas, à notre connaissance, été appliquées à des FSPM auparavant. Ces deux chapitres illustrent ainsi l'intérêt de ces méthodes qui peuvent être utilisées pour répondre à une large gamme de problématiques, avec WALTer ou avec d'autres FSPM.

Modélisation de l'évolution de peuplements hétérogènes

En plus des développements de WALTer permettant de simuler des peuplements hétérogènes à

l'échelle d'une saison de culture, cette thèse présente également des développements permettant d'étendre la simulation à des échelles de temps plus larges. L'utilisation de FSPM pour simuler plusieurs générations successives est intéressante pour l'étude des processus évolutifs prenant place dans des peuplements hétérogènes de plantes (Renton, 2013), mais est encore peu développée. Quelques auteurs ont cependant travaillé sur le couplage entre FSPM et algorithmes évolutionnistes dans le but de simuler plusieurs générations de plantes et d'améliorer la compréhension de leur évolution (Bornhofen *et al.*, 2011 ; Renton et Poot, 2014 ; de Vries *et al.*, 2020). Les algorithmes évolutionnistes sont une famille d'algorithmes qui sont inspirés de la théorie de l'évolution et sont utilisés en optimisation. En association avec des FSPM, ces algorithmes peuvent apporter une perspective évolutive à l'exploration de l'architecture des plantes et de leur développement (Utama, 2015). Dans cette approche il n'y a pas de modélisation explicite des gènes : certains paramètres du FSPM représentent l'information génétique des individus et les valeurs de ces paramètres dans la population évoluent au cours des générations par la sélection des individus les plus performants et par des mutations aléatoires. Une modélisation génétique plus explicite est proposée par certains auteurs. Combes *et al.* (2013, 2015) ont ainsi proposé de simuler des générations successives de ray-grass à partir du couplage entre un FSPM et un modèle de génétique. Le modèle de génétique considère les effets additifs des allèles de plusieurs loci pour déterminer la valeur d'un des paramètres du FSPM qui influence le phénotype de la plante. Grâce à cette approche, il est possible d'étudier l'évolution de la composition génétique des populations en plus de leur évolution phénotypique. En outre, le couplage entre FSPM et modèle de génétique est une approche prometteuse pour simuler les interactions génotype-environnement et pourrait être utile pour la sélection (Struik, 2016 ; Louarn et Song, 2020). Le couplage entre WALTer et GenoWALT, présenté dans cette thèse, correspond à ce type d'approche. Il permet ainsi d'augmenter le niveau de détail génétique pris en compte dans WALTer. En effet, d'après White et Hoogenboom (2003), il est possible de distinguer 6 catégories de modèles écophysiologicals en fonction du niveau de détail génétique pris en compte :

- 1) Modèle générique sans référence à l'espèce considérée
- 2) Modèle spécifique à une espèce mais ne distinguant pas de génotypes
- 3) Modélisation des différences génétiques via des paramètres génotypiques (paramètres dont la valeur est spécifique à chaque génotype)
- 4) Modélisation des différences génétiques par l'action de gènes avec un effet de leurs allèles sur les valeurs de paramètres
- 5) Modélisation des différences génétiques par l'action de gènes en prenant en compte la régulation de l'expression des gènes et l'effet des produits des gènes
- 6) Modélisation des différences génétiques par l'action de gènes en prenant en compte les interactions entre les régulateurs de l'expression des gènes, les produits des gènes, et d'autres métabolites

D'après cette classification, WALTer, utilisé seul, appartient à la catégorie 3. En revanche, le couplage de WALTer et GenoWALT appartient à la catégorie 4 et permet donc une modélisation plus détaillée.

En contrôlant les valeurs de certains paramètres de WALTer par les effets de gènes explicitement modélisés par GenoWALT, le couplage proposé permet ainsi l'intégration de plusieurs échelles (Figure 1). En effet, pour chaque individu, les gènes modélisés dans GenoWALT déterminent certaines caractéristiques des organes (limbes, entre-nœuds, etc...) et impactent donc le phénotype des plantes simulées dans WALTer. Dans WALTer, la fitness de chaque individu dépend de son phénotype et du phénotype des plantes voisines, en interaction avec l'environnement lumineux. Au cours des générations de resemis, la fitness des individus impacte la composition génétique et phénotypique de la population, et donc sa performance. L'intégration de ces différentes échelles dans les FSPM, bien que prometteuse, est encore peu développée (Evers *et al.*, 2018 ; Louarn *et Song*, 2020). Le couplage proposé entre WALTer et GenoWALT apparaît donc cohérent face aux défis actuels de la modélisation structure-fonction.

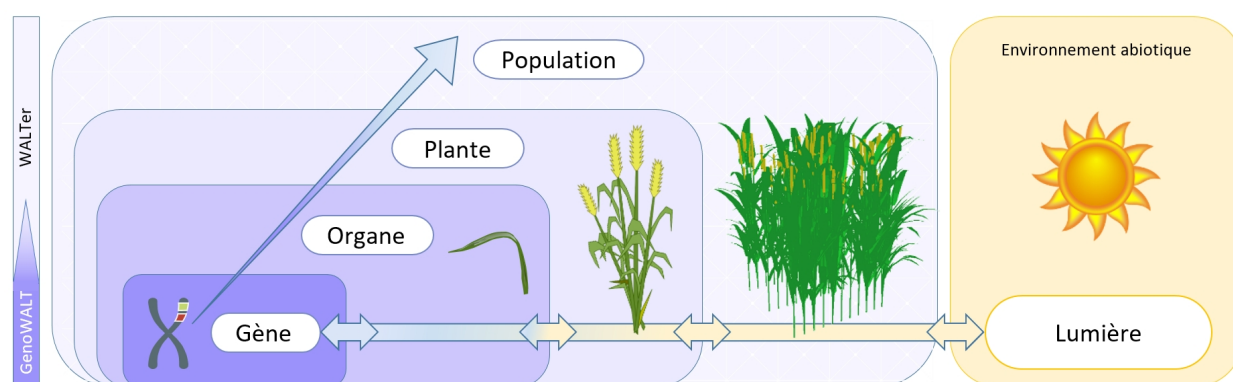


Figure 1. Schéma adapté de Louarn *et Song* (2020) et présentant l'intégration de différentes échelles (du gène à la population) qui découle du couplage entre WALTer et GenoWALT.

Apports de la thèse pour la conception et la gestion des peuplements hétérogènes de blé

Cette thèse constitue un des rares travaux visant à étudier spécifiquement l'impact de l'architecture aérienne sur la performance et l'évolution de peuplements hétérogènes de blé. Elle apporte des éléments de compréhension sur l'impact de l'architecture aérienne sur la compétition pour la lumière en prenant en compte la plasticité de tallage. En effet, l'approche par simulation permet d'isoler l'impact de différents traits d'architecture sur la compétition pour la lumière en ne considérant aucun autre impact (impact sur l'utilisation de l'eau et des nutriments, ou impact sur les maladies par exemple) et en s'affranchissant des trade-offs biologiques entre traits. En ce sens, l'étude proposée ici apporte des informations complémentaires à celles fournies par les approches expérimentales existantes pour l'étude des peuplements hétérogènes de blé (Borg *et al.*, 2018 ; Montazeaud *et al.*, 2020 ; Goldringer *et al.*, 2001 ; Raquin *et al.*, 2008 ; Döring *et al.*, 2015 ; Knapp *et al.*, 2020). Les travaux présentés sont également complémentaires des études de modélisation structure-fonction déjà réalisées sur le sujet (Vidal *et al.*, 2018 ; Barillot *et al.*, 2019) car ils se concentrent sur l'impact de la compétition pour la lumière et considèrent la plasticité du tallage et le développement des plantes pendant l'intégralité de leur cycle de développement. À notre connaissance, la prise en compte de plusieurs traits d'architecture et de leurs interactions (Chapitre 5) ou de l'évolution sur

plusieurs générations (Chapitre 6) n'avaient pas non plus été développées dans les études précédentes de modélisation structure-fonction sur les peuplements hétérogènes de blé.

Le Chapitre 5 a ainsi permis d'identifier que les dimensions des limbes et la capacité de tallage des plantes sont les traits de l'architecture aérienne qui impactent le plus la performance des mélanges via la compétition pour la lumière et la plasticité du tallage. Il est intéressant de noter que ces traits sont également les plus importants pour la performance des peuplements purs dans les simulations. Ces résultats n'impliquent pas que ces traits soient les principaux déterminants de la performance des peuplements de blé en conditions réelles, car d'autres mécanismes que la plasticité du tallage en réponse à la compétition pour la lumière sont alors impliqués. Cependant, étant donné le rôle clef de la plasticité du tallage en réponse à la compétition pour la lumière sur la performance des peuplements de blé (Xie *et al.*, 2016), ces résultats semblent suffisants pour suggérer qu'une attention particulière doit être portée aux dimensions foliaires et à la capacité de tallage des plantes pour la conception de peuplements hétérogènes. Ces traits ont déjà fait l'objet d'attentions pour la sélection de variétés de blé pour les cultures pures (Donald, 1968). Si, pour les cultures pures, sélection et agronomie se focalisent sur la valeur optimale de ces traits pour maximiser la performance, il est possible de considérer également la variance des valeurs des traits lors de la conception de peuplements hétérogènes. Les résultats du Chapitre 5 ont ainsi montré qu'une diversité importante de dimensions foliaires et de capacité de tallage est favorable à la performance des mélanges binaires simulés. La bonne performance des mélanges présentant ces caractéristiques semble liée à un effet de sélection, puisque le génotype le plus compétitif présente un gain de productivité important, qui dépasse la perte de productivité subie par le 2^{ème} génotype du mélange. Si cet effet de sélection peut s'avérer avantageux pour la performance à l'échelle d'une saison de culture, il pourrait s'avérer problématique pour la stabilité et la performance du mélange dans le temps. L'étude présentée dans le Chapitre 6 a ainsi montré que l'effet de sélection qui prend place dans une population composite présentant une grande diversité de hauteur conduit à une homogénéisation et à une baisse de productivité de la population au cours des générations de resemis, ce qui a des implications pour la conception et la gestion des peuplements hétérogènes de blé.

Il est important de souligner que les éléments de compréhension apportés par cette thèse, sur les interactions prenant place dans les peuplements hétérogènes de blé, doivent être interprétés en prenant en considération les limites de l'étude, qui ont été discutées dans les différents chapitres. Une limite importante est ainsi que WALTer ne simule aucune autre ressource que la lumière. Les simulations sont donc faites sous l'hypothèse que les autres ressources, en particulier l'eau et l'azote, ne sont pas limitantes, et la calibration du modèle et sa validation ont été faites à partir de données reflétant ces conditions. Cependant, dans le cadre de la transition agroécologique, ce sont les systèmes à bas niveau d'intrants qui sont favorisés, ce qui correspond à des conditions plus limitantes (Hatt *et al.*, 2016). De plus, plusieurs études suggèrent que les bénéfices associés à la culture de peuplements hétérogènes seraient plus importants dans des conditions stressantes (Borg *et al.*, 2018). D'autre part, les travaux présentés dans cette thèse sont centrés sur le critère de productivité de parcelles de blé hétérogènes. Il semble toutefois important de rappeler que l'agroécologie promeut une diversification des cultures dans l'optique de favoriser un ensemble de services écosystémiques et pas seulement le rendement (Hatt *et al.*, 2016). L'augmentation du rendement n'est ainsi pas la première motivation avancée par les agriculteurs pour la culture de

peuplements hétérogènes de blé (van Frank, 2018). Cependant, nos résultats suggèrent que certaines combinaisons de traits pourraient être défavorables à la performance à l'échelle de la parcelle. Cela confirme l'importance de mieux comprendre les mécanismes à l'œuvre dans les peuplements hétérogènes, car une bonne productivité est capitale pour les agriculteurs.

Perspectives

En plus des travaux présentés dans cette thèse, un travail est en cours sur l'utilisation de la nouvelle fonctionnalité de WALTer, permettant de simuler la mort d'une partie des plantes de la parcelle. Ce travail vise à étudier la capacité de compensation offerte par le tallage lorsque des stress touchent la parcelle et entraînent la mort d'une partie des plantes. En particulier, les impacts du moment du stress et de la proportion de plantes touchées dans la parcelle seront étudiés. Cette étude apportera ainsi des informations sur les conditions dans lesquelles la plasticité du tallage permet une compensation d'une perte de peuplement. Cette meilleure compréhension de la compensation par le tallage aidera à la réflexion sur les règles d'assemblage pour les mélanges de variétés.

L'utilisation de WALTer, et du couplage avec GenoWALT, pour la simulation de peuplements hétérogènes ouvrent de nombreuses perspectives, comme cela a été évoqué dans les différents chapitres de cette thèse. Il sera notamment intéressant de réaliser des simulations visant à identifier des combinaisons de traits conduisant à la fois à une bonne performance du couvert et à une stabilité du peuplement au cours du temps. Une autre perspective intéressante est l'intégration d'un module de précocité de floraison à WALTer, pour déterminer le nombre final de feuilles des plantes en fonction de la séquence climatique et de paramètres génotypiques. Ce développement serait un complément intéressant aux fonctionnalités déjà existantes en modélisation structure-fonction des peuplements hétérogènes de blé. L'intégration d'un tel module ouvrirait un vaste champ d'analyses dans lequel la séquence climatique ajouterait une forte plasticité de phénologie, mais aussi d'architecture. En effet, vernalisation et photopériode conditionnent la transition florale en fonction des conditions climatiques, et plus cette transition est tardive, plus la plante peut émettre de feuilles et d'axes (Allard *et al.*, 2012). Ainsi, la phénologie est connue pour avoir un impact important sur la performance des peuplements hétérogènes de blé, en interaction avec l'architecture aérienne (Borg *et al.*, 2018). Des différences de phénologie au sein d'un couvert hétérogène de blé pourraient également avoir des conséquences sur l'évolution du peuplement, comme cela a été montré dans la population PA menée en gestion dynamique (Goldringer *et al.*, 2006 ; Rhone *et al.*, 2008 et 2010). La simulation de peuplements hétérogènes avec une diversité de phénologie et d'architecture aérienne serait donc intéressante, d'autant que le nombre final de feuilles sur le brin maître a été identifié comme étant le paramètre le plus influent de WALTer lors de l'analyse de sensibilité sur des peuplements purs.

Plus largement, les résultats des simulations de WALTer peuvent guider la conception d'essais expérimentaux sur les peuplements hétérogènes de blé. Il est en effet essentiel de produire des données expérimentales pour la calibration des modèles (Gaudio *et al.*, 2019 ; Louarn et Song, 2020) et pour tester certaines combinaisons de traits architecturaux. Les simulations de WALTer peuvent également être utiles à la planification d'expérimentations *in silico* utilisant des FSPM plus complexes et plus lourds à utiliser, ou servir de base pour la conception de modèles plus simples et plus rapides, comme c'est le cas avec le modèle WHEAMM (Gawinowski *et al.*, 2020). La

complémentarité entre les différents modèles existants pourra notamment être mise à profit dans le cadre du projet ANR MoBiDiv. Ce projet, qui vise à produire des connaissances pour la diversification intra-parcelle dans l'optique d'une transition vers une agriculture sans pesticides, intégrera ainsi l'utilisation de plusieurs modèles, dont WALTer, CN-Wheat et WHEAMM.

À plus long terme, une compréhension des interactions prenant place dans les peuplements hétérogènes de blé permettrait d'établir des règles pour la conception de mélanges de variétés. De telles règles d'assemblage peuvent être intégrées dans des outils d'évaluation multicritères visant à évaluer le potentiel de certaines associations de génotypes. L'outil OptiMix ([Flutre *et al.*, 2019](#)) a ainsi été développé dans le cadre du projet ANR Wheatamix pour évaluer l'aptitude au mélange de variétés sur la base de leurs résistances à différentes maladies. À terme, des règles d'assemblage, basées sur l'architecture aérienne des variétés, pourraient être intégrées à cet outil en accès libre pour les agriculteurs et les conseillers.

RÉFÉRENCES

- Allard, V., Veisz, O., Kőszegi, B., Rousset, M., Le Gouis, J., & Martre, P. (2012). The quantitative response of wheat vernalization to environmental variables indicates that vernalization is not a response to cold temperature. *Journal of Experimental Botany*, 63(2), 847-857.
- Barillot, R., Chambon, C., Fournier, C., Combes, D., Pradal, C., & Andrieu, B. (2019). Investigation of complex canopies with a functional–structural plant model as exemplified by leaf inclination effect on the functioning of pure and mixed stands of wheat during grain filling. *Annals of botany*, 123(4), 727-742.
- Borg, J., Kiær, L. P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., ... & Enjalbert, J. (2018). Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*, 221, 298-313.
- Bornhofen, S., Barot, S., & Lattaud, C. (2011). The evolution of CSR life-history strategies in a plant model with explicit physiology and architecture. *Ecological Modelling*, 222(1), 1-10.
- Carmassi, M., Barbillon, P., Keller, M., Parent, E., & Chiodetti, M. (2018). Bayesian calibration of a numerical code for prediction. *arXiv preprint arXiv:1801.01810*.
- Combes, D., Litrico, I., Barre, P., Gutiérrez, A. E., Rodriguez, L. S., & Louarn, G. (2015, November). Effets de l'intensité de la défoliation sur la composition génétique et phénotypique d'une population virtuelle de ray-grass. In *Colloque présentant les méthodes et résultats du projet Climagie (méta-programme ACCAF)* (pp. 223-p). INRA.
- Combes, D., Litrico, I., Grenier, S., Barre, P., Escobar-Gutiérrez, A. J., & Louarn, G. (2013). Effects of defoliation intensity on the genetic and phenotypic composition of virtual ray-grass populations. *Publishers: Finnish Society of Forest Science, Vantaa, Finland Finnish Forest Research Institute, Vantaa, Finland Department of Forest Sciences, University of Helsinki, Helsinki, Finland*, 325.
- Da Silva, D., Han, L., Faivre, R., & Costes, E. (2014). Influence of the variation of geometrical and topological traits on light interception efficiency of apple trees: sensitivity analysis and metamodelling for ideotype definition. *Annals of botany*, 114(4), 739-752.
- de Vries, J., Evers, J. B., Poelman, E. H., & Anten, N. P. (2020). Optimal plant defence under competition for light and nutrients: an evolutionary modelling approach. *in silico Plants*.
- Donald, C. T. (1968). The breeding of crop ideotypes. *Euphytica*, 17(3), 385-403.
- Döring, T. F., Annicchiarico, P., Clarke, S., Haigh, Z., Jones, H. E., Pearce, H., ... & Wolfe, M. S. (2015). Comparative analysis of performance and stability among composite cross populations, variety mixtures and pure lines of winter wheat in organic and conventional cropping systems. *Field Crops Research*, 183, 235-245.
- Evers, J. B., Van Der Werf, W., Stomph, T. J., Bastiaans, L., & Anten, N. P. (2019). Understanding and optimizing species mixtures using functional–structural plant modelling. *Journal of experimental botany*, 70(9), 2381-2388.

- Evers, J. B., Letort, V., Renton, M., & Kang, M. (2018).** Computational botany: Advancing plant science through functional-structural plant modelling. *Annals of Botany*, 121(5), 767-772.
- Evers, J. B., Vos, J., Yin, X., Romero, P., van der Putten, P. E. L., and Struik, P. C. (2010).** Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *Journal of Experimental Botany*, 61(8):2203–2216.
- Flutre, T., Forst, E., Borg, J., Enjalbert, J., Gauffreteau, A., & Goldringer, I. (2019).** OptiMix: Online, Multivariate Evaluation Analysis of Varietal Mixtures of Bread Wheat (version 0.1.2). R. INRA. <http://moulon.inra.fr/optimix/>.
- Gaudio, N., Escobar-Gutiérrez, A. J., Casadebaig, P., Evers, J. B., Gerard, F., Louarn, G., ... & Barillot, R. (2019).** Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agronomy for Sustainable Development*, 39(2), 20.
- Gauthier, M., Barillot, R., Schneider, A., Chambon, C., Fournier, C., Pradal, C., ... & Andrieu, B. (2020).** A functional structural model of grass development based on metabolic regulations and coordination rules. *Journal of Experimental Botany*, 71(18), 5454–5468.
- Gawinowski, M., Enjalbert, J., Cournède, P. H., & Flutre, T. (2020, October).** WHEAMM: a functional-structural model to study plant growth and interactions with neighbors in wheat variety mixtures. In *FSPM 2020: 9th International Conference on FUNCTIONAL-STRUCTURAL PLANT MODELS*.
- Goldringer, I., Enjalbert, J., Raquin, A. L., & Brabant, P. (2001).** Strong selection in wheat populations during ten generations of dynamic management. *Genetics Selection Evolution*, 33(1), S441-S463.
- Goldringer, I., Prouin, C., Rousset, M., Galic, N., & Bonnin, I. (2006).** Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. *Annals of Botany*, 98(4), 805-817.
- Hatt, S., Artru, S., Brédart, D., Lassois, L., Francis, F., Haubruge, E., ... & Boeraeve, F. (2016).** Towards sustainable food systems: the concept of agroecology and how it questions current research practices. A review. *Biotechnologie, Agronomie, Société et Environnement*, 20(Special issue 1), 215-224.
- Knapp, S., Döring, T. F., Jones, H. E., Snape, J., Wingen, L. U., Wolfe, M. S., ... & Griffiths, S. (2020).** Natural selection towards wild-type in composite cross populations of winter wheat. *Frontiers in Plant Science*, 10, 1757.
- Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019).** WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.
- Louarn, G., & Song, Y. (2020).** Two decades of functional–structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany*, 126(4), 501-509.
- Montazeaud, G., Violle, C., Roumet, P., Rocher, A., Ecartot, M., Compan, F., ... & Fréville, H. (2020).** Multifaceted functional diversity for multifaceted crop yield: Towards ecological assembly rules for varietal mixtures. *Journal of Applied Ecology*, 57(11), 2285-2295.

- Perez, R. P., Dauzat, J., Pallas, B., Lamour, J., Verley, P., Caliman, J. P., ... & Faivre, R. (2018).** Designing oil palm architectural ideotypes for optimal light interception and carbon assimilation through a sensitivity analysis of leaf traits. *Annals of botany*, 121(5), 909-926.
- Pradal, C., Dufour-Kowalski, S., Boudon, F., Fournier, C., & Godin, C. (2008).** OpenAlea: a visual programming and component-based software platform for plant modelling. *Functional plant biology*, 35(10), 751-760.
- Raquin, A. L., Brabant, P., Rhoné, B., Balfourier, F., Leroy, P., & Goldringer, I. (2008).** Soft selective sweep near a gene that increases plant height in wheat. *Molecular Ecology*, 17(3), 741-756.
- Renton, M. (2013).** Aristotle and adding an evolutionary perspective to models of plant architecture in changing environments. *Frontiers in plant science*, 4, 284.
- Renton, M., & Poot, P. (2014).** Simulation of the evolution of root water foraging strategies in dry and shallow soils. *Annals of botany*, 114(4), 763-778.
- Rhone, B., Remoue, C., Galic, N., Goldringer, I., & Bonnin, I. (2008).** Insight into the genetic bases of climatic adaptation in experimentally evolving wheat populations. *Molecular Ecology*, 17(3), 930-943.
- Rhoné, B., Vitalis, R., Goldringer, I., & Bonnin, I. (2010).** Evolution of flowering time in experimental wheat populations: a comprehensive approach to detect genetic signatures of natural selection. *Evolution: International Journal of Organic Evolution*, 64(7), 2110-2125.
- Struik, P. C. (2016).** Bridging the genotype–phenotype gap in 3D. *Journal of experimental botany*, 67(15), 4427-4430.
- Utama, D. N. (2015).** The Optimization of the 3-d Structure of Plants, Using Functional-Structural Plant Models. Case Study of Rice (*Oryza sativa* L.) in Indonesia. *Georg-August University*.
- Van Frank, G. (2018).** Gestion participative de la diversité cultivée et création de mélanges diversifiés de blé tendre à la ferme. Doctoral dissertation, Université Paris-Saclay (ComUE).
- Vidal, T., Gigot, C., de Vallavieille-Pope, C., Huber, L., & Saint-Jean, S. (2018).** Contrasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. *Annals of botany*, 121(7), 1299-1308.
- White, J. W., & Hoogenboom, G. (2003).** Gene-based approaches to crop simulation: Past experiences and future opportunities. *Agronomy Journal*, 95(1), 52-64.
- Xie, Q., Mayes, S., & Sparkes, D. L. (2016).** Optimizing tiller production and survival for grain yield improvement in a bread wheat × spelt mapping population. *Annals of botany*, 117(1), 51-66.

ANNEXE 1

**Ressources complémentaires du
chapitre 1**

WALTer et adaptation du modèle

Tableau 1. Liste des entrées de WALTer, définition, valeur par défaut et unité.

Paramètre	Définition	Valeur	Unité
Paramètres génotypiques			
Dse_{mean}	Délai moyen entre le semis et l'émergence de la première feuille	81	Degrés-jours
Dse_{sd}	Ecart-type du délai entre le semis et l'émergence de la première feuille	30	Degrés-jours
Phl	Intervalle entre l'émergence de 2 feuilles successives	99	Degrés-jours
a_w^B	Coefficient utilisé pour le calcul de la largeur maximale du limbe	0.099	-
b_w^B	Coefficient utilisé pour le calcul de la largeur maximale du limbe	-0.3	cm
L_{max}^B	Longueur finale du plus grand limbe de l'axe	17.4	cm
φ_{zen}^B	Angle entre le limbe et la tige à la base du limbe	10	Degrés
d_I	Diamètre des entrenœuds	0.38	cm
d_P	Diamètre du pédoncule	0.25	cm
d_S	Diamètre des gaines	0.4	cm
d_E	Diamètre de l'épi	0.4	cm
L_E	Longueur finale de l'épi	7.93	cm
N_{MS}^B	Nombre final de feuilles sur le brin maître	11.3 ^a	-
Δ_{prot}	Délai minimum entre la mort de 2 talles successives sur une même plante	25	Degrés-jours
PAR_t	Quantité seuil de PAR interceptée par un axe, en dessous de laquelle l'axe meurt	300000	$\mu\text{mol}\cdot\text{cm}^{-2}\cdot\text{degrés-jours}^{-1}$
t_{beg}^{reg}	Âge physiologique de la plante à partir duquel la régression des talles devient possible	4.2	- ^b
H_{MS}	Hauteur finale du brin maître	55.6	cm
shp_I	Formalisme utilisé pour calculer la longueur finale des entrenœuds	ExpIncrease	-
N_I^{el}	Nombre d'entrenœuds allongés	5	-
a_L^S	Coefficient utilisé pour le calcul de la longueur finale des gaines	0.6919	-
b_L^S	Coefficient utilisé pour le calcul de la longueur finale des gaines	-2.6953	cm
L_1^B	Longueur finale de la première feuille de chaque axe	7	cm

a La partie décimale du nombre final de feuilles correspond à la proportion de plantes présentant une feuille supplémentaire. Le reste des plantes présentant un nombre final de feuilles égal à la partie entière de N_{MS}^B . Ce formalisme permet de reproduire l'hétérogénéité observée dans les cultures.

b Le paramètre t_{beg}^{reg} est exprimé en nombre de feuilles non ligulées, *i.e.* en nombre de feuilles dont l'élongation n'est pas terminée, sur le brin maître. Une valeur de 4.2 indique ainsi que 4 feuilles doivent encore finir leur croissance, ainsi qu'une 5^{ème}, qui doit encore effectuer 20% de son élongation.

S_1^B	Incrément de longueur entre 2 feuilles successives (uniquement pour les feuilles qui émergent avant la transition florale)	0.8	cm
N_r^B	Nombre de feuilles de rang supérieur à la plus longue feuille et qui sont moins longues que la plus longue feuille de l'axe	2	-
S_f^B	Facteur de réduction de la longueur du limbe entre 2 feuilles successives (uniquement pour les feuilles de rang supérieur à la plus longue feuille de l'axe)	0.8	cm
$n0_{sen}$	Age physiologique pour le début de la première phase de sénescence des feuilles	4.75	Haun stage
$n1_{sen}$	Nombre de feuilles vertes à la fin de la première phase de sénescence des feuilles	3.31	-
$n2_{sen}$	Nombre de feuilles vertes à la fin de la deuxième phase de sénescence des feuilles (talles uniquement)	4.5	-
$n3_{sen}$	Nombre de feuilles vertes à la fin de la sénescence des feuilles	0	-
$t0_{sen}$	Date de début de la première phase de sénescence des feuilles	468	Degrés-jours
$t1_{sen}$	Date de fin de la première phase de sénescence des feuilles	691	Degrés-jours
$t2_{sen}$	Date de fin de la deuxième phase de sénescence des feuilles (talles uniquement)	1131	Degrés-jours
$t3_{sen}$	Date de la fin de la sénescence des feuilles	2000	Degrés-jours
P_T	Probabilité d'émergence d'une talle (sauf talle du coléoptile)	0.88	-
P_{CT}	Probabilité d'émergence de la talle du coléoptile	0	-
GAI_c	Valeur de GAI des plantes du voisinage d'un individu, au-dessus de laquelle l'individu concerné arrête sa phase d'émission de talles	0.75	-
curvature	Degré de courbure du limbe (angle entre la tige et la pointe du limbe – angle d'insertion du limbe)	30	Degrés
S_{curv}	Position relative sur la nervure centrale du limbe où les 2/3 de la courbure du limbe sont atteints	0.75	-
Paramètres non génotypiques			
densite	Densité de semis	200	Plants/m ²
d_{GAIp}	Distance de prise en compte du GAI de proximité pour l'arrêt du tallage	1	m
nbj	Nombre de jours à simuler	365	-
rep	Paramètre pour fixer la graine aléatoire	1	-
sky _{type}	Type de rayonnement solaire simulé	soc	-
turtle _{sectors}	Nombre de secteurs pour la discrétisation du ciel TURTLE	16	-
add _{sun}	Option de simulation de la course du soleil par le module de rayonnement	0 (inactif)	-
location	Ville dont les coordonnées géographiques sont utilisées pour simuler la position du soleil	Grignon	-
meteo	Séquences météorologiques à utiliser (PAR et température)	Climat_moy_AS	-

infinity _{CARIBU}	Option d'infinetisation de la parcelle pour le module de rayonnement	1 (infinetisation)	-
CARIBU _{state}	Option d'activation du module de rayonnement	enabled	-
crop_ccptn	Type de plan de semis utilisé	Mesh_for_nplants	-
nb_plt_utiles	Nombre de plantes à simuler	56	-
dist_border_x	Largeur des bordures dans lesquelles des plantes supplémentaires seront simulées, à l'est et à l'ouest de la parcelle	0	m
dist_border_y	Largeur des bordures dans lesquelles des plantes supplémentaires seront simulées, au nord et au sud de la parcelle	0	m
leaf_nb_segments	Nombre de segments utilisés pour la représentation de la courbure des limbes	7	-

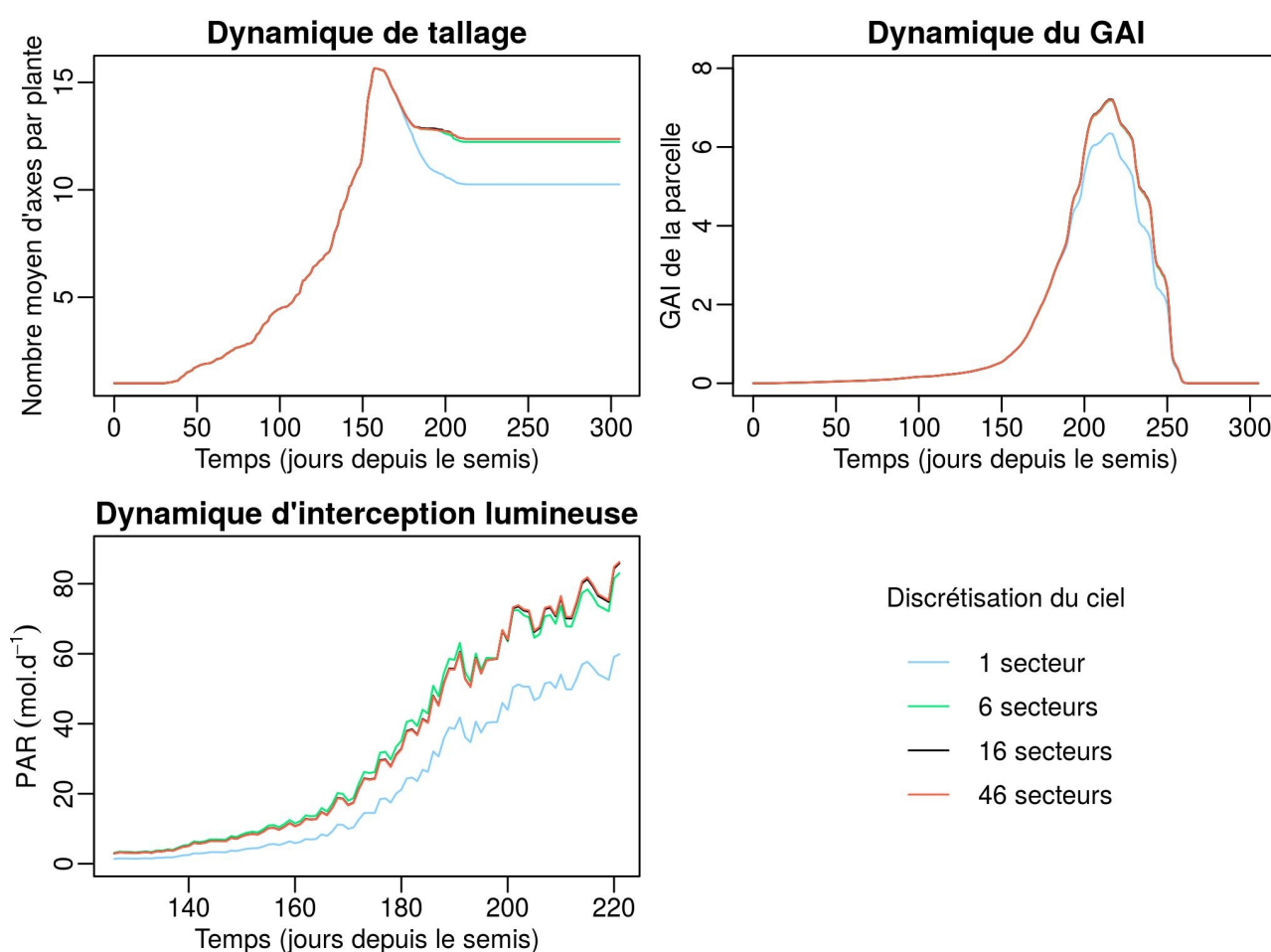


Figure 1. Sorties moyennes à l'échelle de la parcelle en fonction du nombre de secteurs utilisés pour la discrétisation du ciel. Résultats d'une simulation d'un mélange binaire de 110 plantes (55 plantes de chaque génotype) semé à une densité de 50 plantes/m². Les 2 génotypes du mélange diffèrent par leur hauteur finale, leur surface foliaire et leur port foliaire.

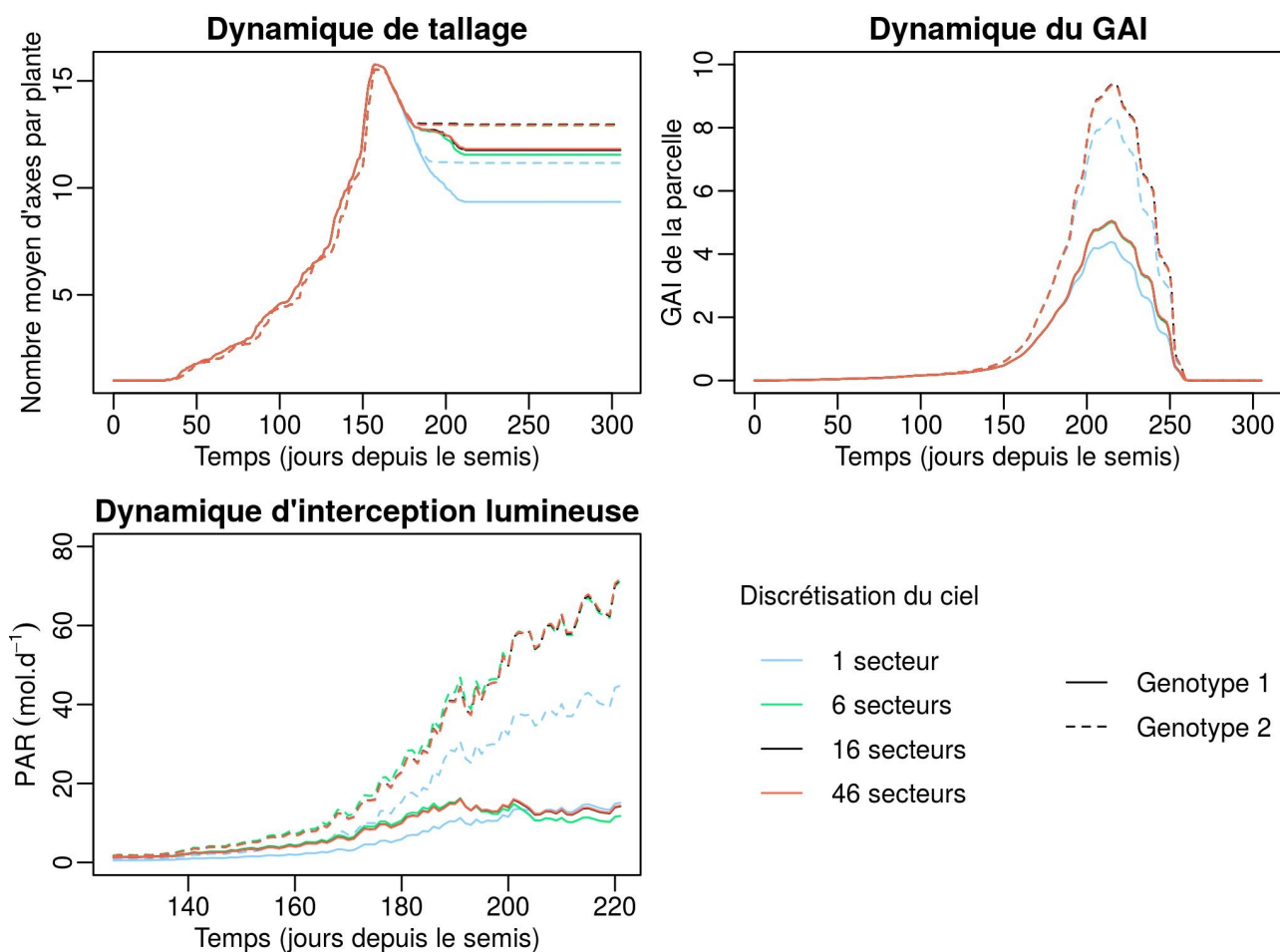


Figure 2. Sorties moyennes par génotype en fonction du nombre de secteurs utilisés pour la discrétisation du ciel. Résultats d'une simulation d'un mélange binaire de 110 plantes (55 plantes de chaque génotype) semé à une densité de 50 plantes/m². Les 2 génotypes du mélange diffèrent par leur hauteur finale, leur surface foliaire et leur port foliaire.

Le nombre de secteurs utilisés pour la discrétisation du ciel a un impact sur les sorties de WALTER. Dans un mélange équilibré de 2 génotype, cet effet est comparable que l'on considère les sorties moyennes à l'échelle de la parcelle (Figure 1) ou les sorties moyennes par génotype (Figure 2). En particulier, les simulations pour lesquelles une unique source lumineuse est utilisée (ciel à 1 secteur) donnent des résultats notablement différents des simulations réalisées avec un ciel plus complexe. Pour la dynamique de GAI, il y a très peu de différences entre les résultats obtenus avec un ciel à 6, à 16 ou à 46 secteurs. En revanche, des différences plus importantes existent entre les résultats obtenus avec une discrétisation à 6 secteurs et ceux obtenus avec 16 ou 46 secteurs pour les dynamiques de tallage et d'interception lumineuse. L'écart entre les résultats de la simulation utilisant 16 secteurs et ceux de la simulation à 46 secteurs est particulièrement faible pour toutes les sorties, si bien que les courbes noire (16 secteurs) et rouge (46 secteurs) sont confondues sur les figures 1 et 2.

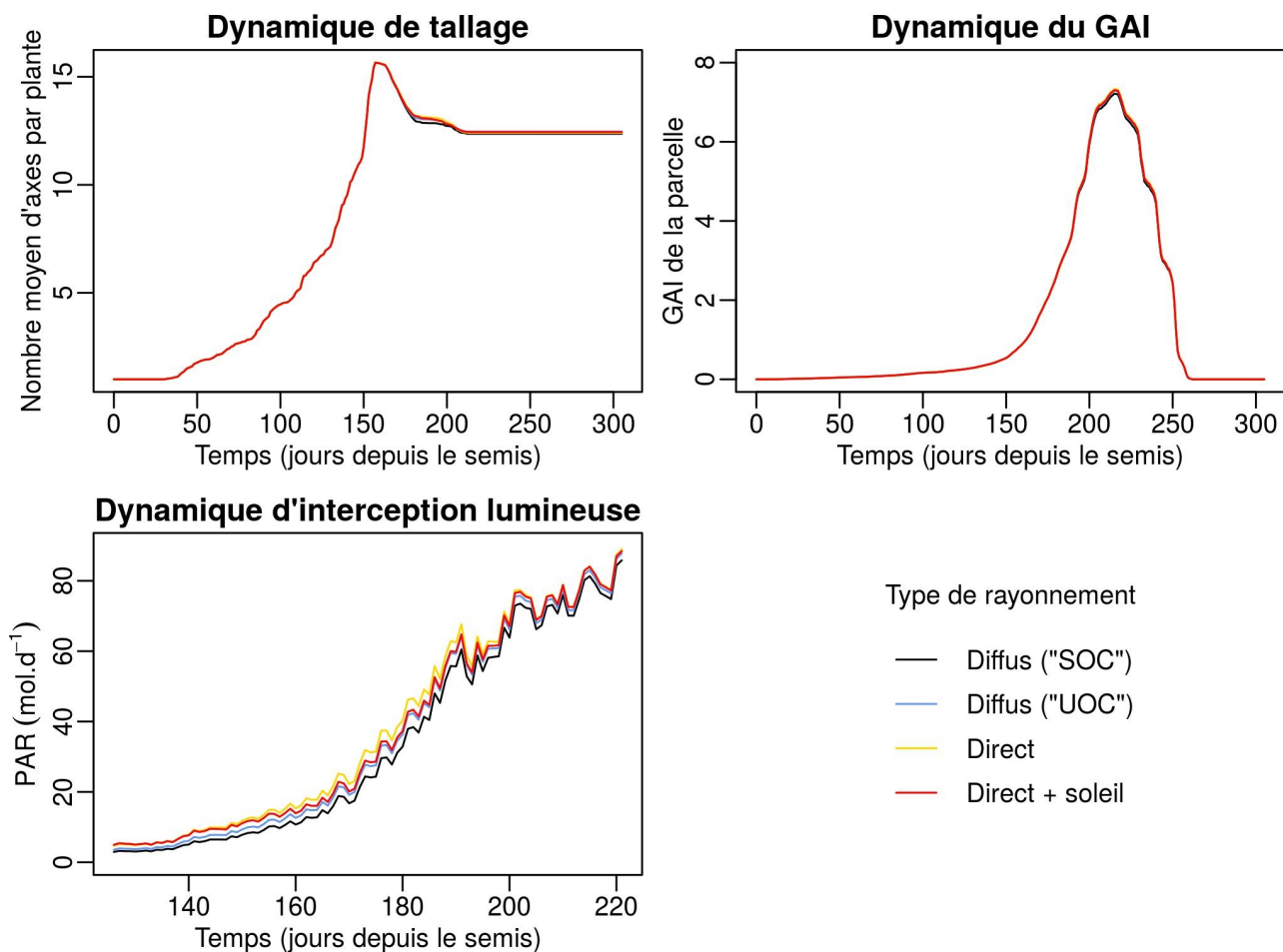


Figure 3. Sorties moyennes à l'échelle de la parcelle en fonction du type de rayonnement utilisé. Résultats d'une simulation d'un mélange binaire de 110 plantes (55 plantes de chaque génotype) semé à une densité de 50 plantes/m². Les 2 génotypes du mélange diffèrent par leur hauteur finale, leur surface foliaire et leur port foliaire.

Le type de rayonnement utilisé pour les simulations a un impact sur la dynamique d'interception lumineuse de la parcelle, avec une quantité de rayonnement intercepté plus faible pour les parcelles soumises à un rayonnement diffus que pour celles soumises à un rayonnement direct. Les différences d'interception de PAR entre simulations avec différents types de rayonnement sont cependant faibles et ont peu d'impact sur les dynamiques de tallage et de GAI, que ce soit à l'échelle de la parcelle (Figure 3) ou à l'échelle du génotype (Figure 4).

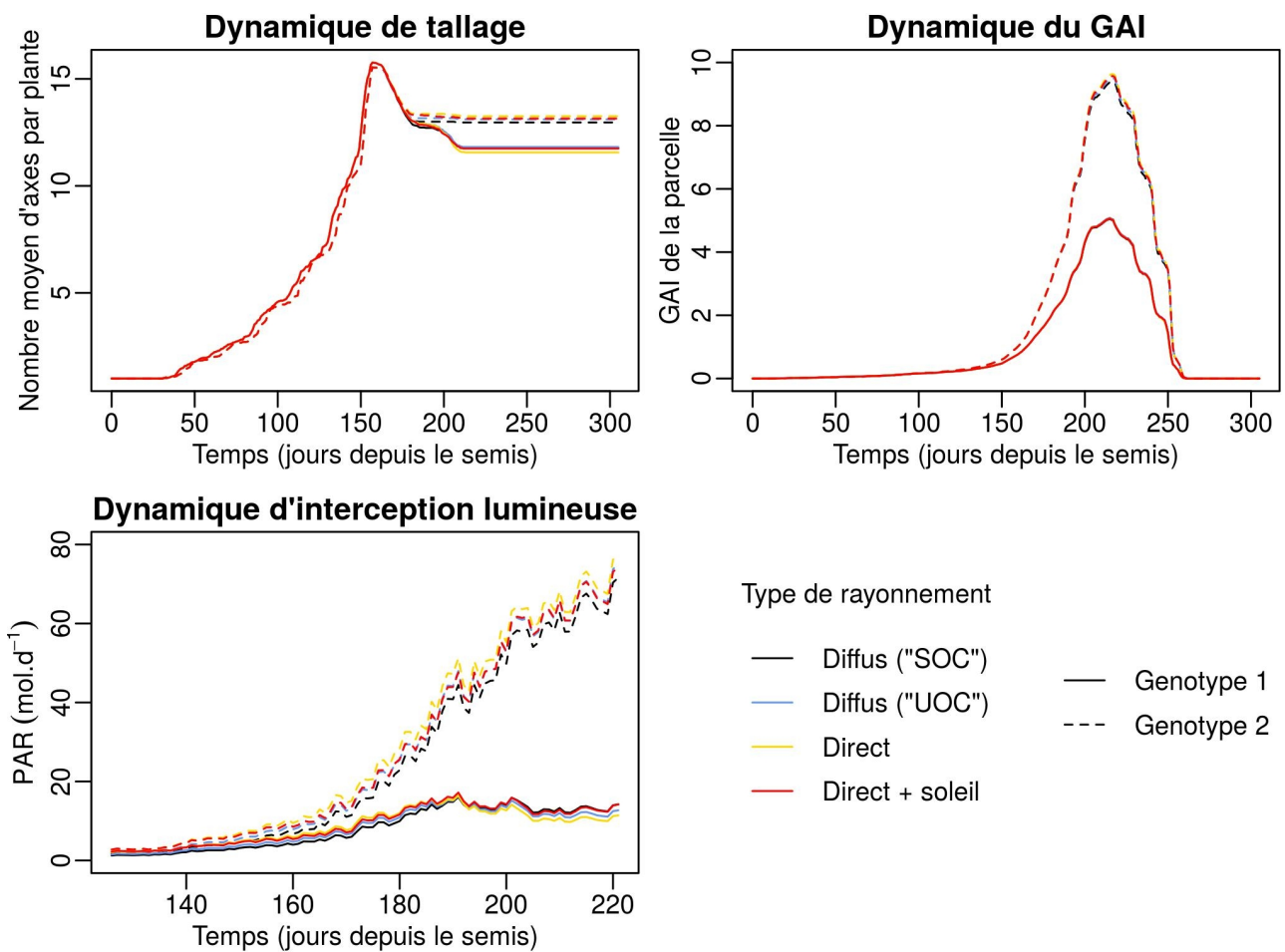


Figure 4. Sorties moyennes par génotype en fonction du type de rayonnement utilisé. Résultats d'une simulation d'un mélange binaire de 110 plantes (55 plantes de chaque génotype) semé à une densité de 50 plantes/m². Les 2 génotypes du mélange diffèrent par leur hauteur finale, leur surface foliaire et leur port foliaire.

Analyse de sensibilité du module de fitness

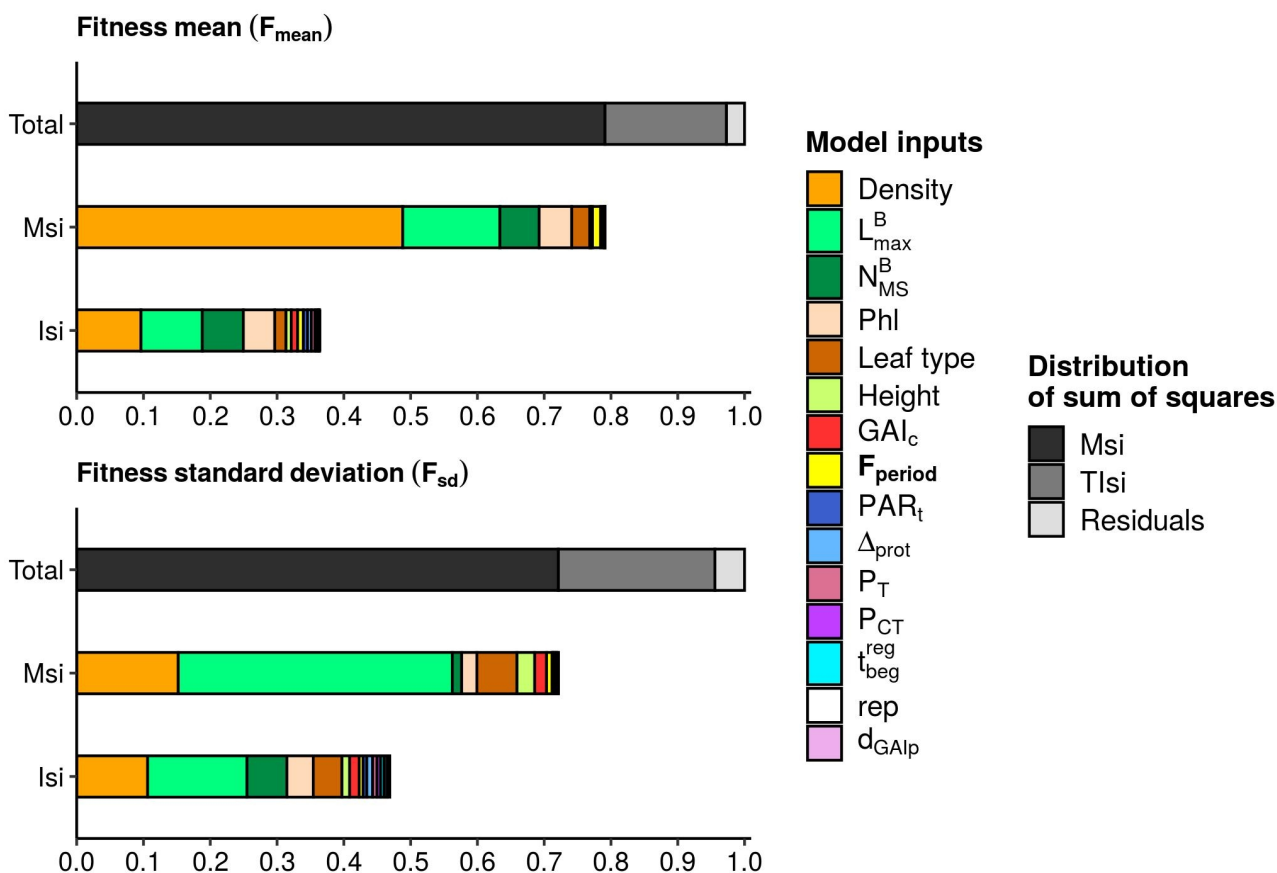


Figure 5. Indices de sensibilité pour la fitness moyenne et l'écart-type de la fitness. Les barres supérieures montrent la distribution des effets principaux, des effets d'interactions de second ordre et de la résiduelle. Les deux barres inférieures montrent l'impact des 15 facteurs d'entrée.

L'impact, sur les valeurs de fitness calculée, d'une modification de F_{period} est faible relativement à l'impact de la variation d'autres entrées du modèle. Dans le plan de simulation décrit au chapitre I.3, des variations de F_{period} n'expliquent que 2,1% de la variance totale de F_{mean} et 1,4% de la variance de F_{sd} .

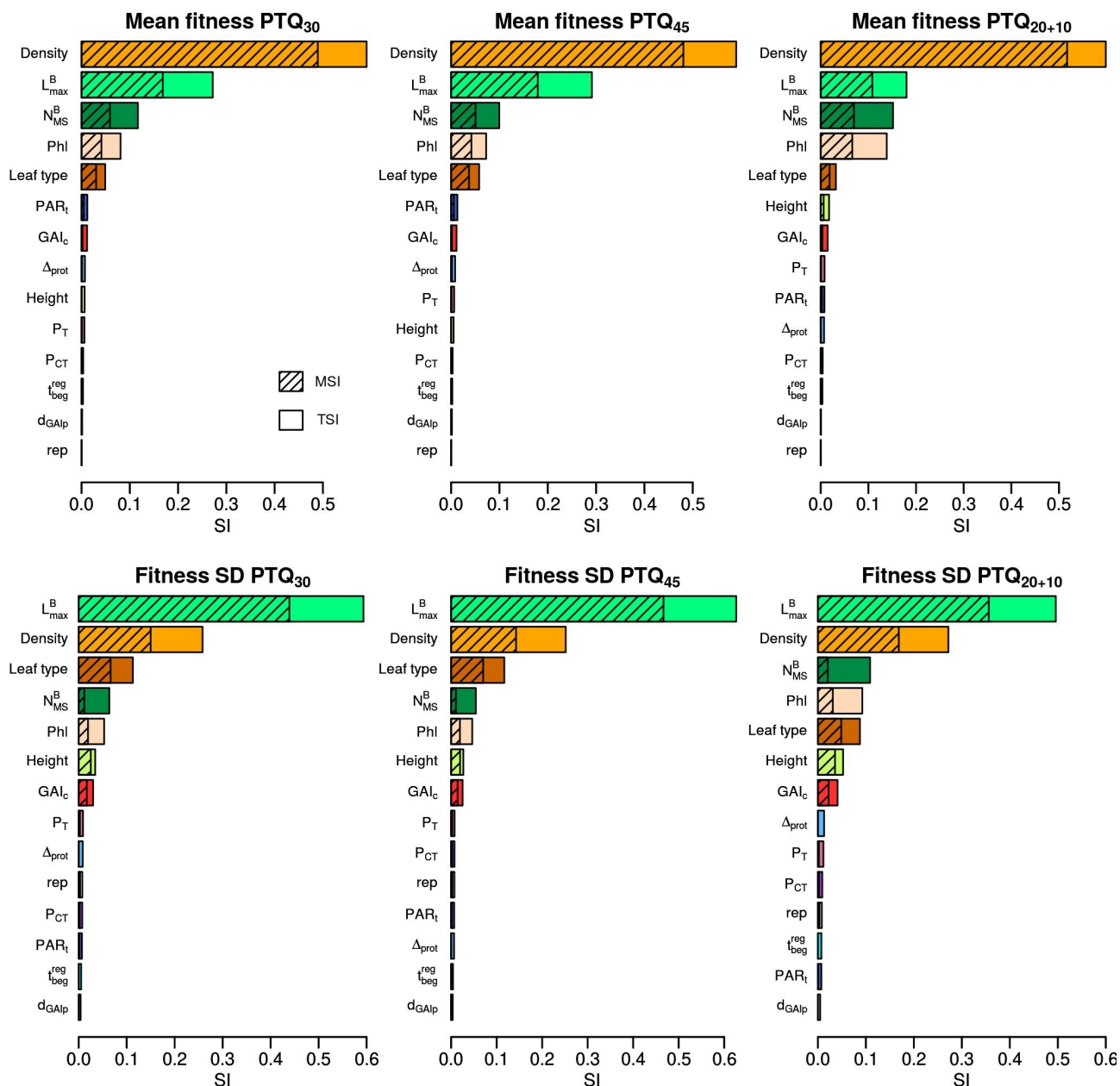


Figure 6. Indices de sensibilité principaux (barres hachurées) et totaux (barres hachurées + barres unies) pour la fitness moyenne (ligne supérieure) et l'écart-type de la fitness (ligne inférieure) en fonction de la période utilisée pour le calcul de la fitness (colonnes).

Le calcul de la fitness moyenne de la parcelle semble peu impacté par les interactions entre F_{period} et les autres entrées du modèle. En effet, quelle que soit la F_{period} utilisée, les 5 entrées ayant le plus d'influence sur F_{mean} sont toujours les mêmes, et leur ordre d'importance reste inchangé. On note toutefois quelques différences dans l'ordre d'importance des 9 autres facteurs d'entrées de l'analyse, mais chacune de ces entrées explique moins de 2% de la variance totale de F_{mean} .

Le choix de F_{period} a un impact sur l'influence des autres entrées sur F_{sd} . Le calcul des indices de sensibilité donne des résultats très proches pour PTQ₃₀ et PTQ₄₅. En effet, pour ces deux périodes, les 8 facteurs d'entrées ayant le plus d'influence sur F_{sd} sont les mêmes, avec le même ordre d'importance. En revanche, des différences plus importantes existent entre les résultats pour PTQ₄₅ et ceux pour PTQ₂₀₊₁₀. Pour ces deux périodes, la longueur finale du plus grand limbe (L_{max}^B) et la densité de semis sont les 2 facteurs d'entrées les plus influents. Cependant, si L_{max}^B explique plus

de 62% de la variance totale de F_{sd} lorsque PTQ_{45} est utilisé, cette même entrée explique un peu de moins de 50% de la variance de F_{sd} pour PTQ_{20+10} . À l'inverse, l'influence du nombre final de feuilles sur le brin maître (N_{MS}^B), du phyllochrone (Phl) et de la hauteur finale du brin maître (Height) sur F_{sd} est plus grande pour PTQ_{20+10} que pour PTQ_{45} .

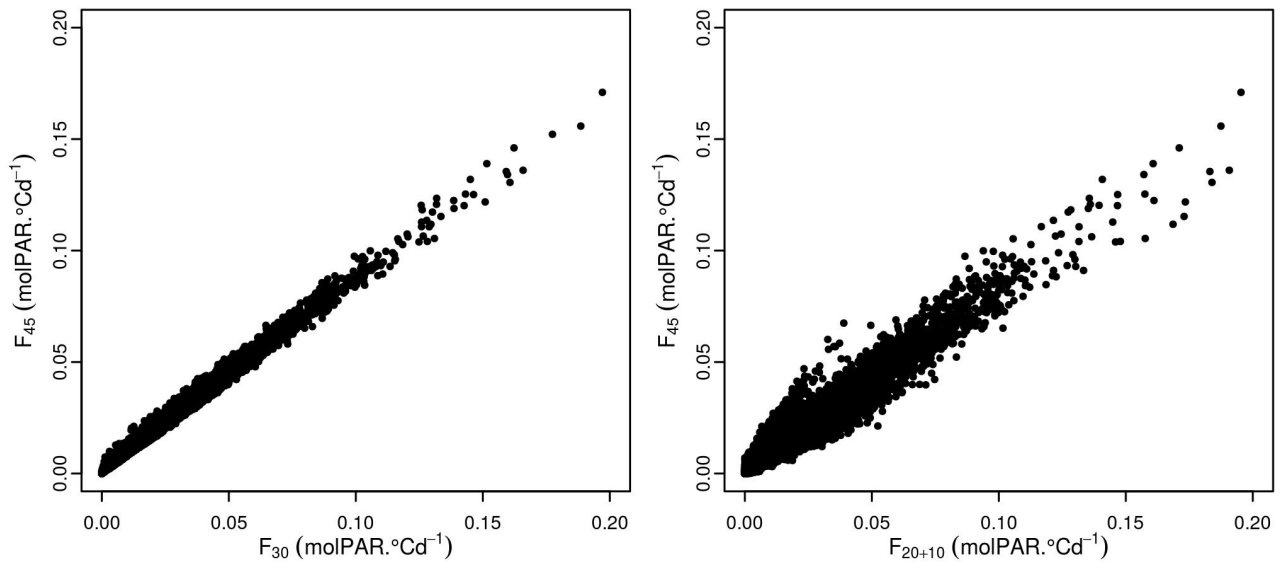


Figure 7. Fitness individuelles (*i.e.* fitness de chaque plante) calculées avec PTQ_{45} , en fonction des fitness individuelles calculées pour PTQ_{30} (graphique de gauche) ou PTQ_{20+10} (graphique de droite). Les plantes considérées sont celles des simulations de l'analyse de sensibilité décrite dans le Chapitre 3 et n'appartiennent donc pas toutes à une même parcelle.

ANNEXE 2

Ressources complémentaires du chapitre 3

Séquence climatique

La séquence climatique utilisée pour les simulations du Chapitre 3 (ainsi que pour les simulations des Chapitres 5 et 6) est une séquence moyenne obtenue à partir de données collectées sur 10 ans (de 2007 à 2017) dans cinq localités de France, sélectionnées pour représenter des conditions climatiques contrastées. Les 5 localités sont : Mons-en-Chaussée (49.9°N, 3.0°E) ; Marseillan-Plage (43.3°N, 3.5°E), Plomelin (47.9°N, 4.2°O), Entzheim (48.5°N, 7.6°E) et Clermont-Ferrand (45.8°N, 3.1°E).

Tillering dynamics

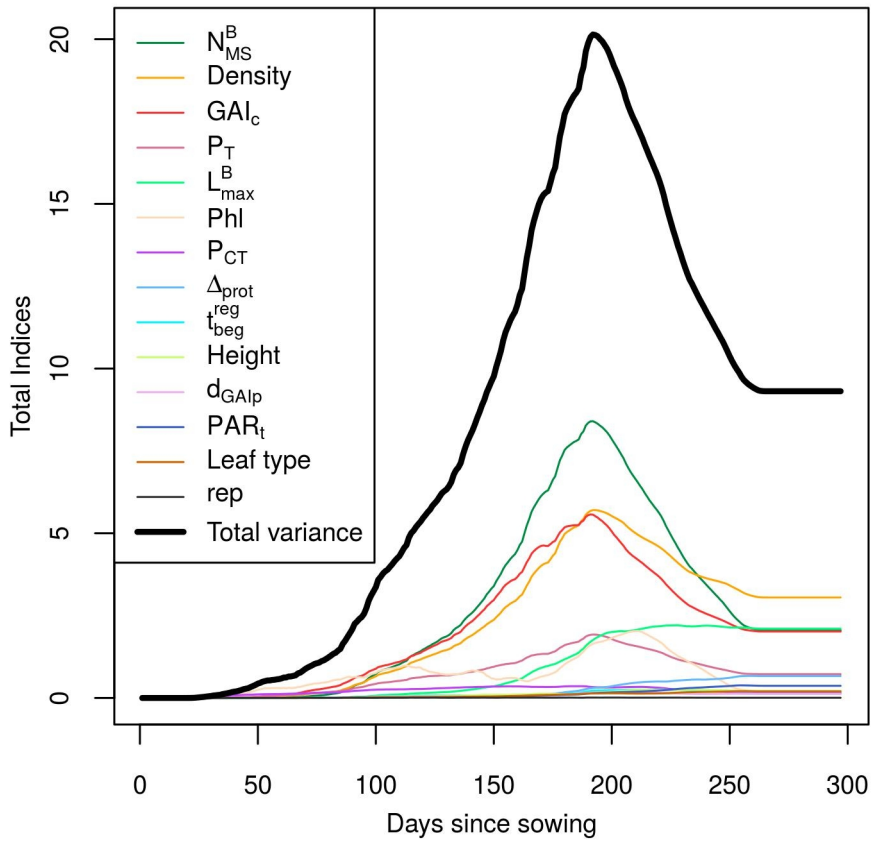


Figure 1. Indices de sensibilités totaux calculés pour chaque date de la dynamique de tallage.

GAI dynamics

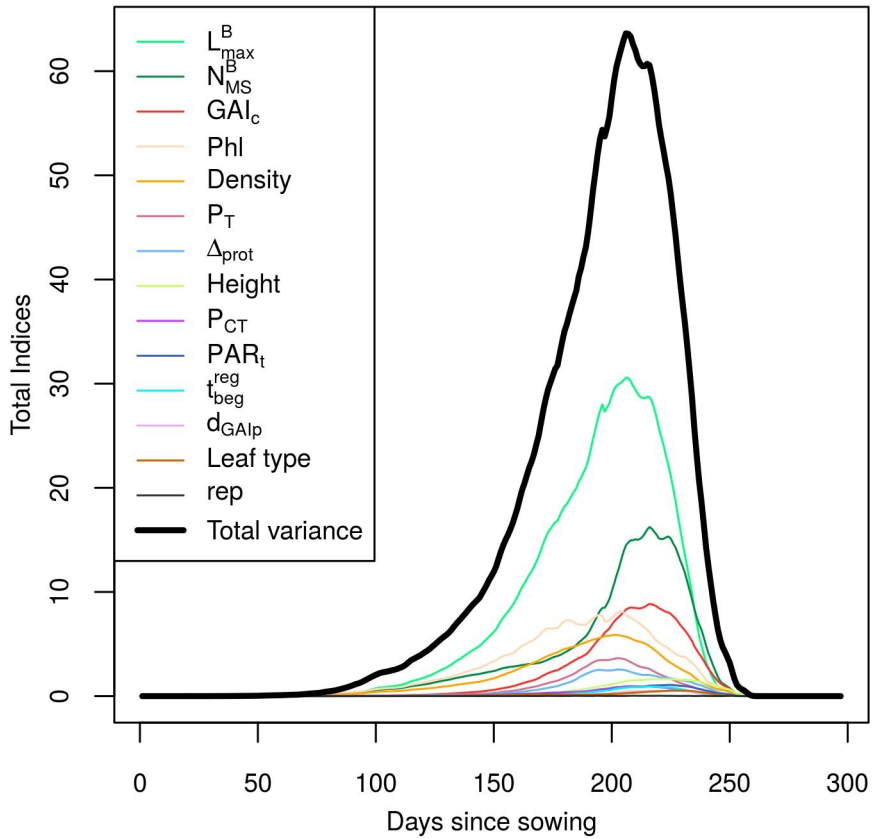


Figure 2. Indices de sensibilités totaux calculés pour chaque date de la dynamique de GAI.

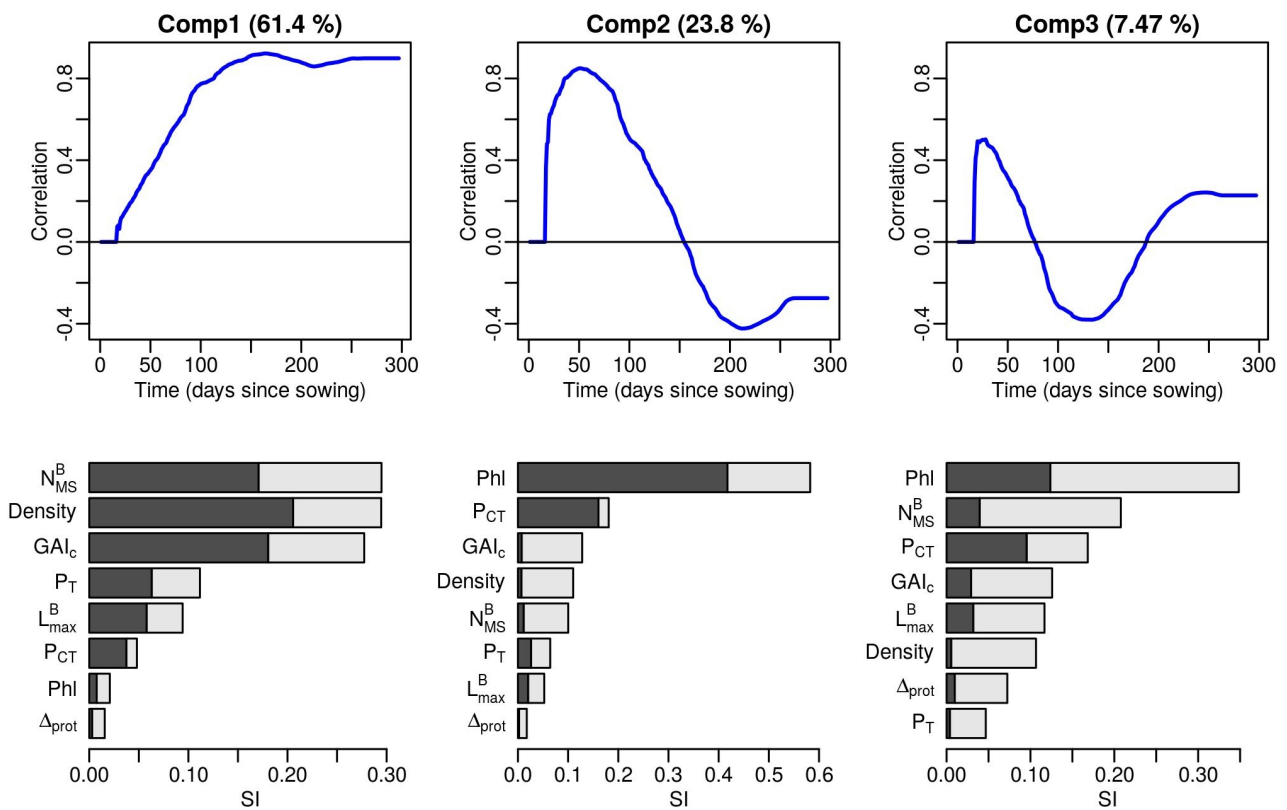


Figure 3. Analyse de sensibilité de la dynamique de tallage basée sur une ACP. Colonnes : composantes principales 1 à 3. Ligne supérieure : coefficient de corrélation (ordonnée) entre la composante principale et les différentes dates de la dynamique de tallage (abscisse). Ligne inférieure : indices de sensibilité principaux (gris foncé) et totaux (gris foncé + gris clair) pour les entrées les plus influentes.

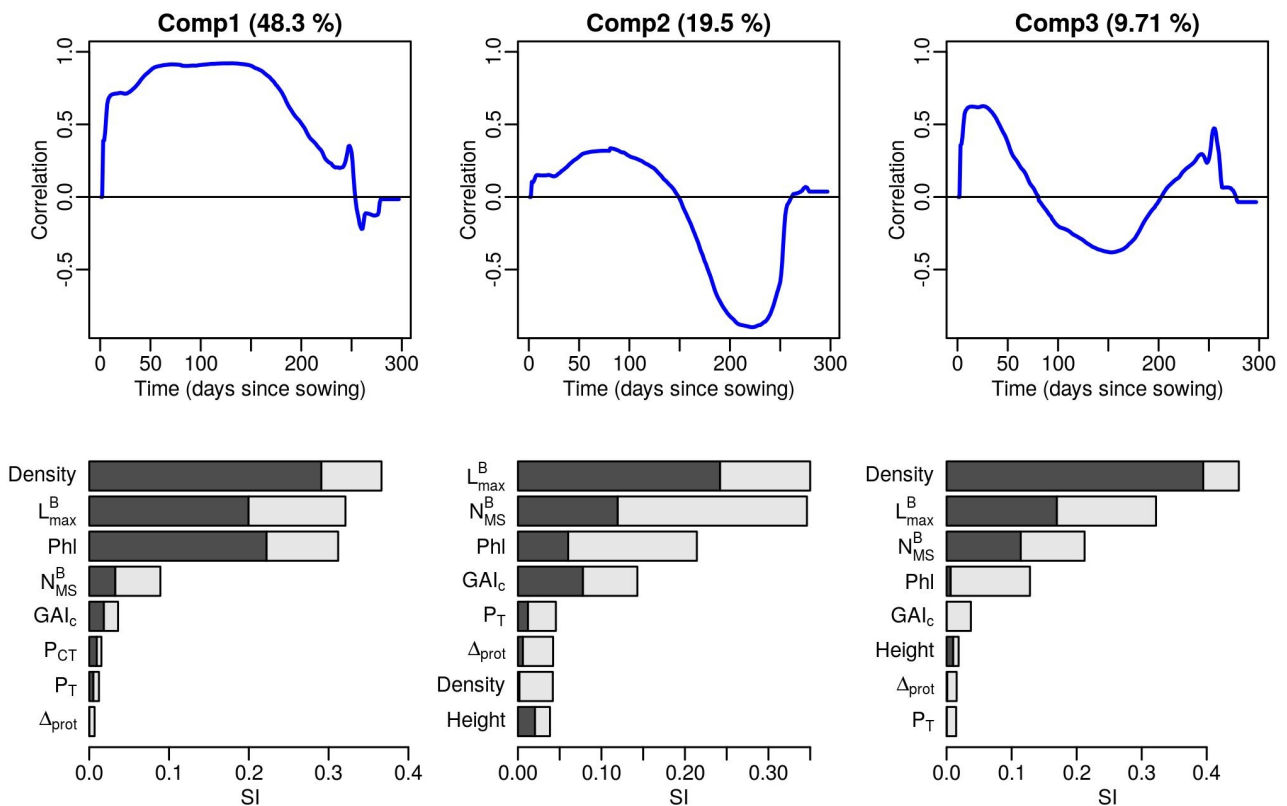


Figure 4. Analyse de sensibilité de la dynamique de GAI basée sur une ACP. Colonnes : composantes principales 1 à 3. Ligne supérieure : coefficient de corrélation (ordonnée) entre la composante principale et les différentes dates de la dynamique de GAI (abscisse). Ligne inférieure : indices de sensibilité principaux (gris foncé) et totaux (gris foncé + gris clair) pour les entrées les plus influentes.

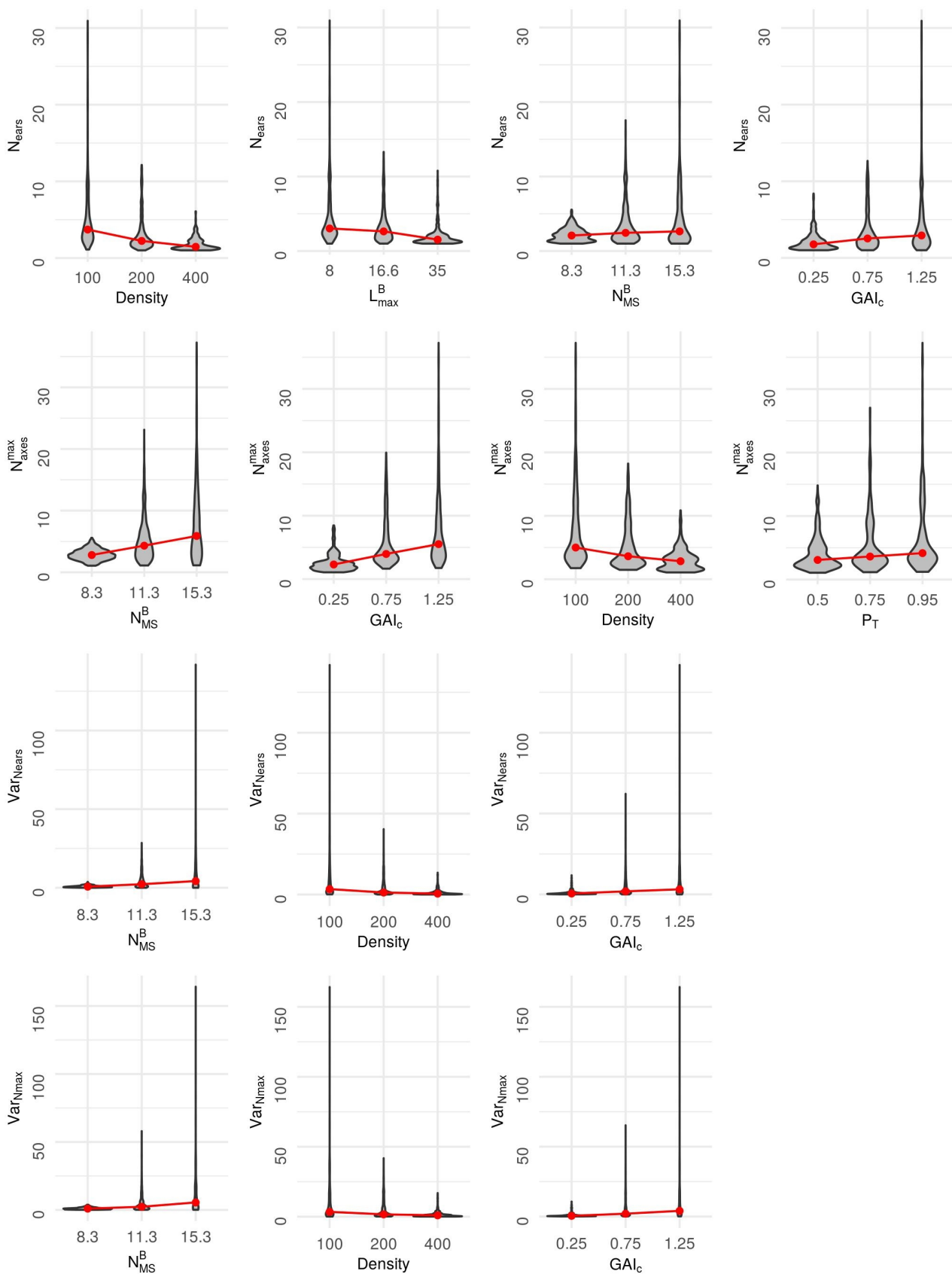


Figure 5. Distributions des différentes sorties descriptives de la dynamique de tallage (N_{ears} , $N_{\text{axes}}^{\text{max}}$, $\text{Var}_{N_{\text{ears}}}$, $\text{Var}_{N_{\text{max}}}$) en fonction des valeurs de leurs facteurs d'entrée les plus influents (facteurs d'entrée expliquant plus de 10% de la variance de la sortie considérée). Les points rouges représentent la médiane pour chaque groupe.

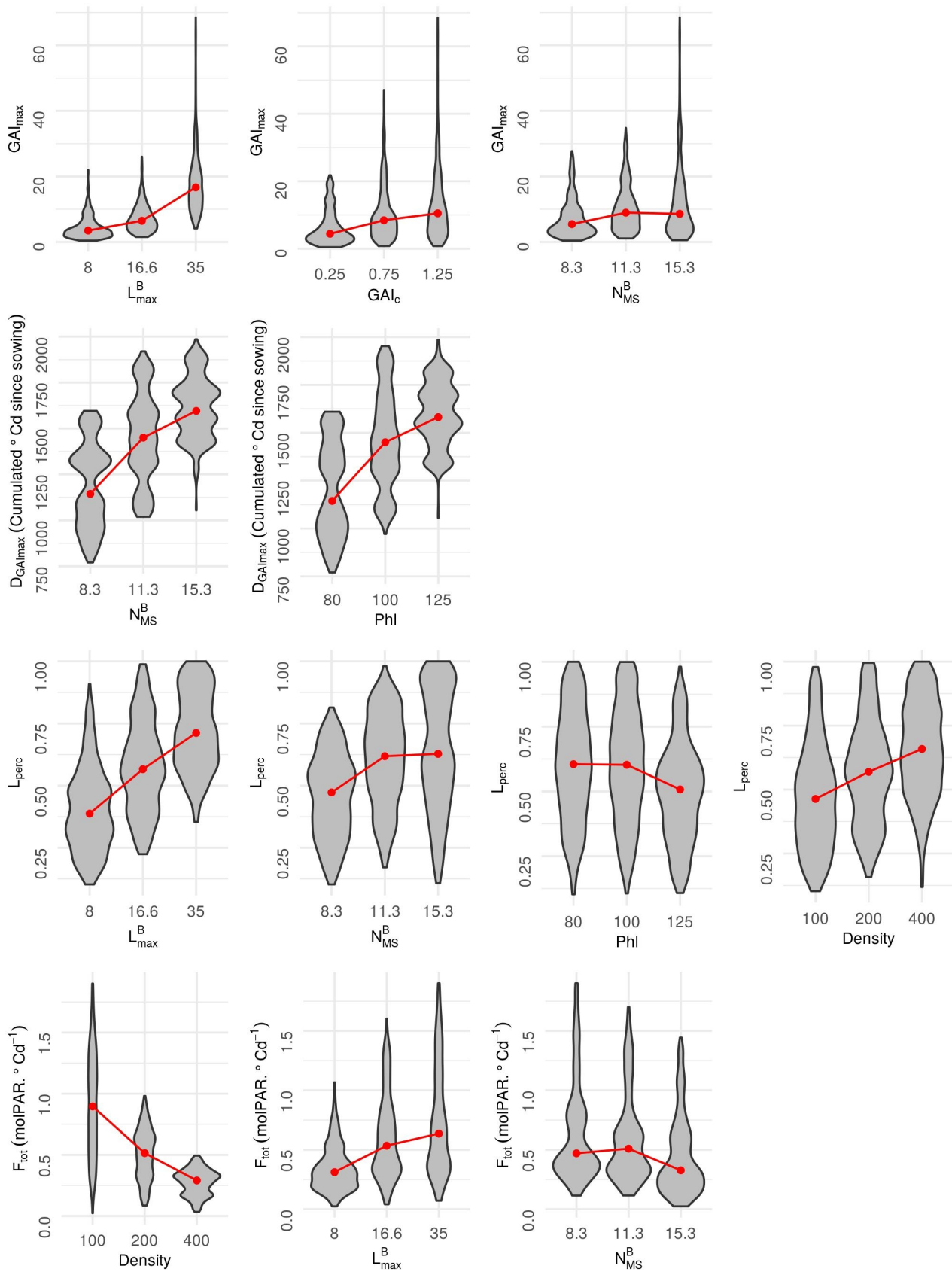


Figure 6. Distributions des différentes sorties descriptives des dynamiques de GAI (GAI_{max} et $D_{GAI_{max}}$) et d'interception lumineuse (L_{perc} et F_{tot}) en fonction des valeurs de leurs facteurs d'entrée les plus influents (facteurs d'entrée expliquant plus de 10% de la variance de la sortie considérée). Les points rouges représentent la médiane pour chaque groupe.

ANNEXE 3

Ressources complémentaires du chapitre 5

Formalism for blade size

As described in Lecarpentier *et al.*, 2019, the final lengths of the blades are computed from a multi-linear function proposed in Fournier *et al.* (2003). This function depends partly on the final length of the longest blade of the axis (L_{max}^B). Consequently, a change in L_{max}^B will result in a change in the final length of all the blades of the plant emerging after the floral transition. Moreover, the maximal width of the blade is computed as a linear function of the blade's final length. Therefore, changes in the value of L_{max}^B also result in changes to the maximal width of the blades emerging after the floral transition. Thus, blades emerging after the floral transition are both longer and wider for plants with high L_{max}^B values.

Additional figure

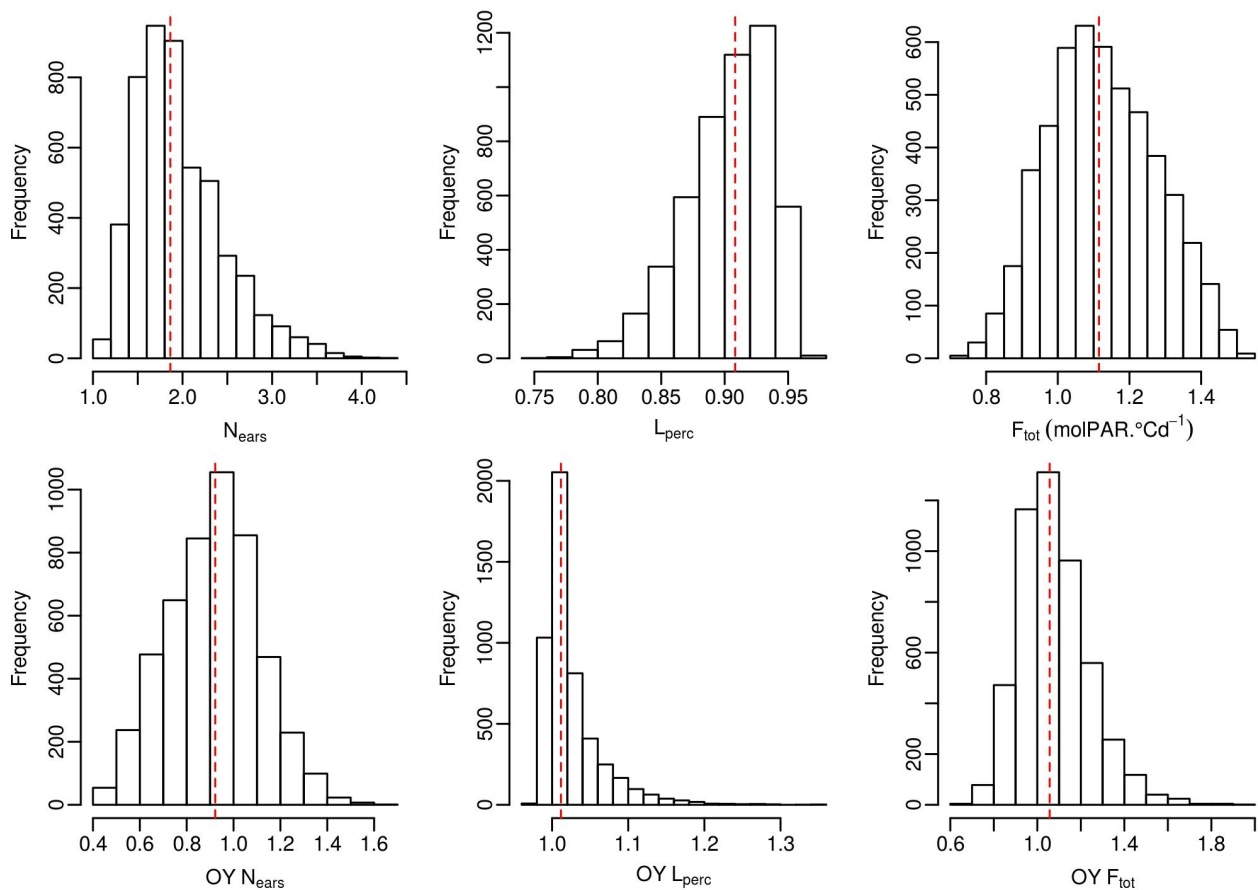


Fig. 1. Distribution of the 3 outputs (N_{ears} , L_{perc} and F_{tot}) and their OY indicators for the 5000 simulations of the initial design. The red dotted lines represent the corresponding median values.

REFERENCES

Fournier, C., Andrieu, B., Ljutovac, S., & Saint-Jean, S. (2003). ADEL-wheat: a 3D architectural model of wheat development. *Plant Growth Modeling and Applications* (eds B.-G. Hu & M. Jaeger), 54-66.

Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., ... & Andrieu, B. (2019). WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Annals of botany*, 123(6), 961-975.

ANNEXE 4

Ressources complémentaires du chapitre 6

Le déséquilibre de liaison

Le déséquilibre de liaison (DL) est une mesure de l'association entre 2 marqueurs dans une population (Feingold, 1991) et peut être estimé par différents indices. En particulier, le DL peut être exprimé en termes de corrélation entre les états alléliques de 2 marqueurs dans une population. Il est alors possible d'utiliser la valeur absolue du coefficient de corrélation (r), qui est comprise entre 0 et 1. Un déséquilibre de liaison non nul entre 2 marqueurs indique ainsi que les allèles des 2 marqueurs ne sont pas associés au hasard dans la population.

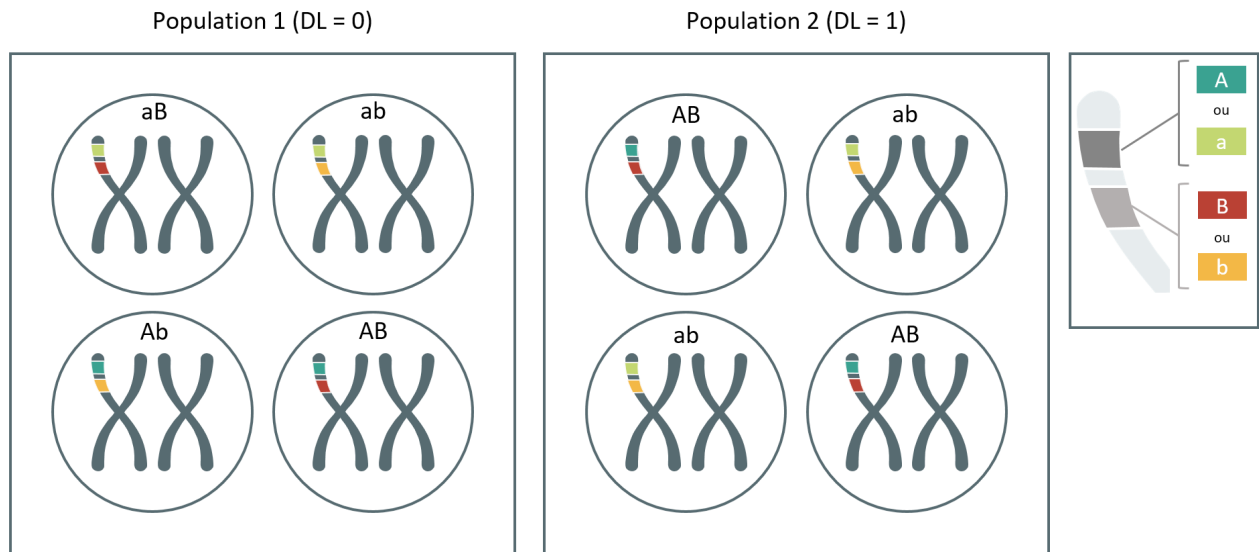


Figure 1. Représentation schématique des génotypes des individus appartenant à 2 populations fictives différentes. 2 gènes bi-alléliques sont considérés. Dans la population 1, le DL entre les 2 gènes est nul : les allèles des 2 gènes sont associés au hasard. Dans la population 2, l'allèle A du gène 1 est systématiquement associé à l'allèle B du gène 2 et l'allèle a du gène 1 est toujours associé à l'allèle b du gène 2. Le DL est donc maximum (DL de 1) dans la population 2.

Lorsqu'un déséquilibre de liaison existe entre un marqueur et un gène sous sélection, il est possible que le marqueur augmente en fréquence dans la population en même temps que le gène sous sélection. On parle alors d'effet auto-stop (ou "*genetic hitchhiking*" en anglais).

Expérience de gestion dynamique in silico

Méthode

Les effets des marqueurs (gènes majeurs et QTL) sont intégrés à un modèle additif simple permettant de calculer la hauteur finale des individus à partir de leur génotype. La hauteur finale du brin maître d'un individu (H_{MS}) est ainsi calculée selon l'équation suivante :

$$H_{MS} = m_H + \sum_{i=1}^n \alpha_i G_i$$

dans laquelle n est le nombre total de marqueurs considérés ; α_i est l'effet additif associé au marqueur i (cf. Tableau 1 du Chapitre 6) et G_i est un coefficient indicateur du génotype de l'individu

au marqueur i . m_H correspond à la hauteur minimale d'une plante de blé à maturité. Cette valeur est ici fixée à 40 cm.

Pour les marqueurs dont les allèles sont co-dominants, le coefficient G s'apparente au nombre d'allèles 'haut' de l'individu à ce locus, tel que :

$G = 2$ pour un individu homozygote pour l'allèle 'haut' (2 allèles 'haut')

$G = 1$ pour un individu hétérozygote (1 allèle 'court' et 1 allèle 'haut')

$G = 0$ pour un individu homozygote pour l'allèle 'court' (2 allèles 'court')

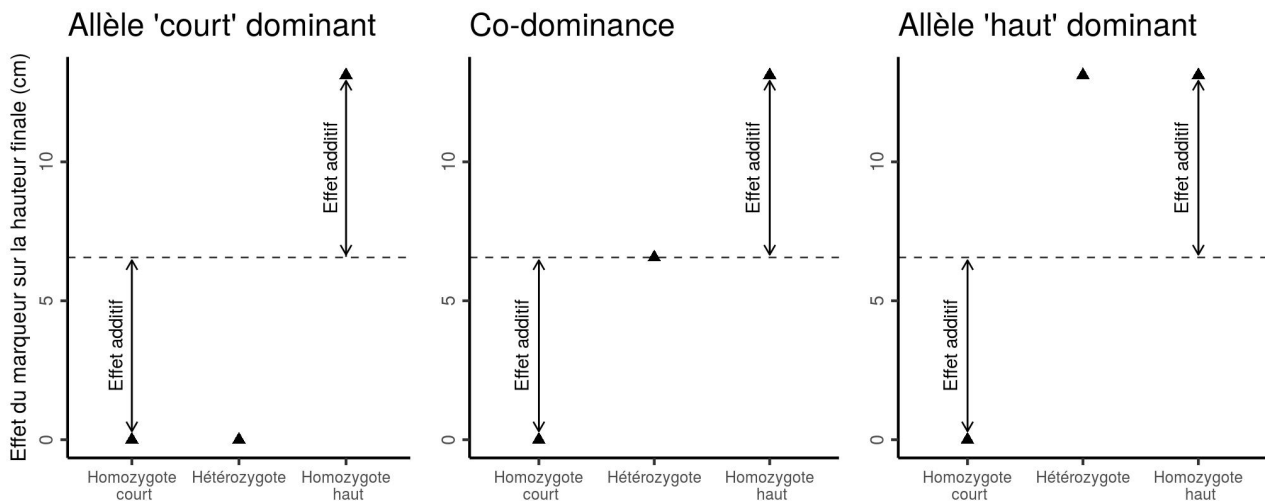


Figure 2. Représentation schématique de l'effet d'un gène bi-allélique sur la hauteur finale d'un individu en fonction des différents génotypes (homozygote court, hétérozygote et homozygote haut) et pour différents régimes de dominance (dominance de l'allèle 'court', co-dominance ou dominance de l'allèle 'haut'). Les individus homozygotes présentent 2 exemplaires du même allèle (2 allèles 'court' ou 2 allèles 'haut') et les hétérozygotes présentent un exemplaire de chaque allèle (1 allèle 'court' et 1 allèle 'haut'). La ligne pointillée représente la demi différence entre les valeurs d'effet du gène pour les 2 homozygotes.

Pour les marqueurs dont l'allèle 'haut' est dominant, le coefficient G a la même valeur pour les individus homozygotes 'haut' et pour les hétérozygotes. En effet, les effets du marqueurs sur la hauteur de la plante sont les même pour ces 2 types d'individus. Ainsi :

$G = 2$ pour un individu homozygote pour l'allèle 'haut' (2 allèles 'haut')

$G = 2$ pour un individu hétérozygote (1 allèle 'court' et 1 allèle 'haut')

$G = 0$ pour un individu homozygote pour l'allèle 'court' (2 allèles 'court')

Résultats

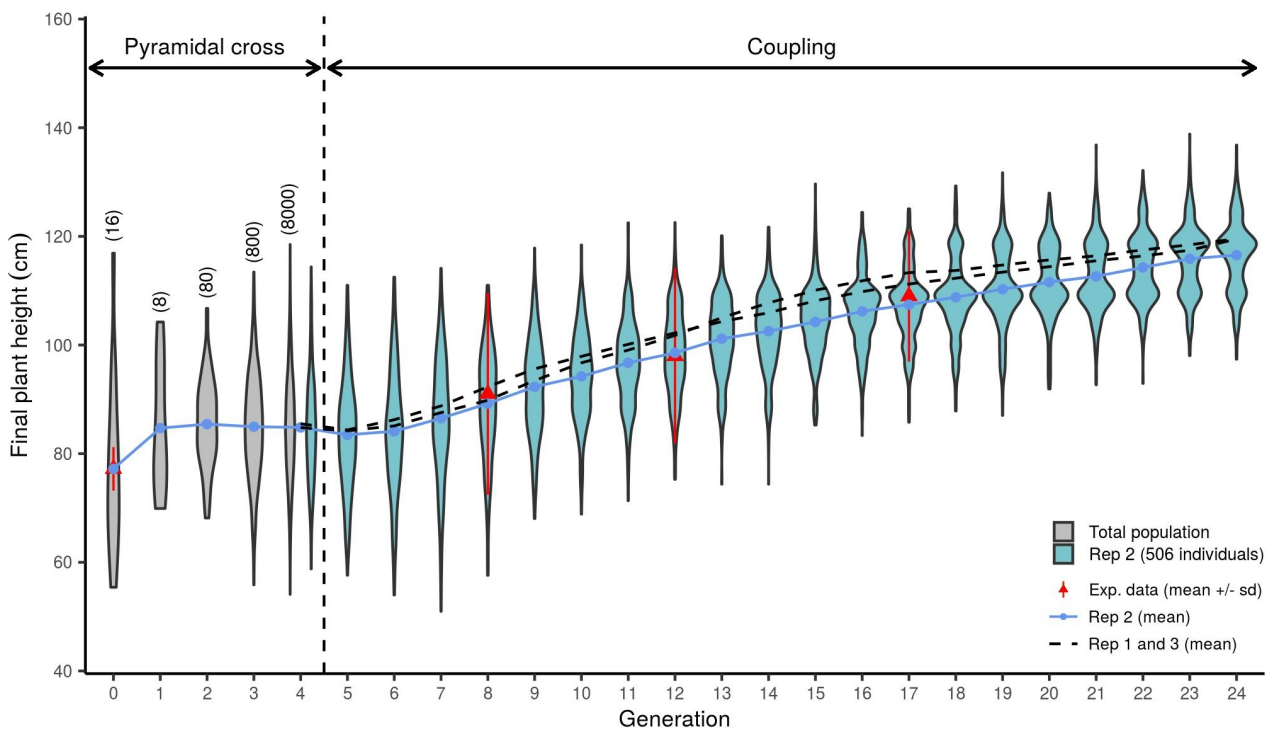


Figure 3. Evolution de la hauteur finale des plantes au cours des générations. La hauteur moyenne au cours des générations pour un des 3 réplicats simulés (rep 2) est représentée par une courbe bleue. La hauteur moyenne au cours des générations pour les 2 autres réplicats simulés est représentée par des courbes noires pointillées. Pour le croisement pyramidal, les distributions de hauteur pour chaque génération sont représentées en gris et le nombre d'individus pour chaque génération est indiqué entre parenthèses. Les distributions de hauteur pour chaque génération du réplicat 2 sont représentées en bleu. Les données expérimentales disponibles dans l'étude de [Raquin et al. \(2008\)](#) pour la hauteur moyenne et l'écart-type de hauteur sont représentés en rouge par des triangles et des segments respectivement.

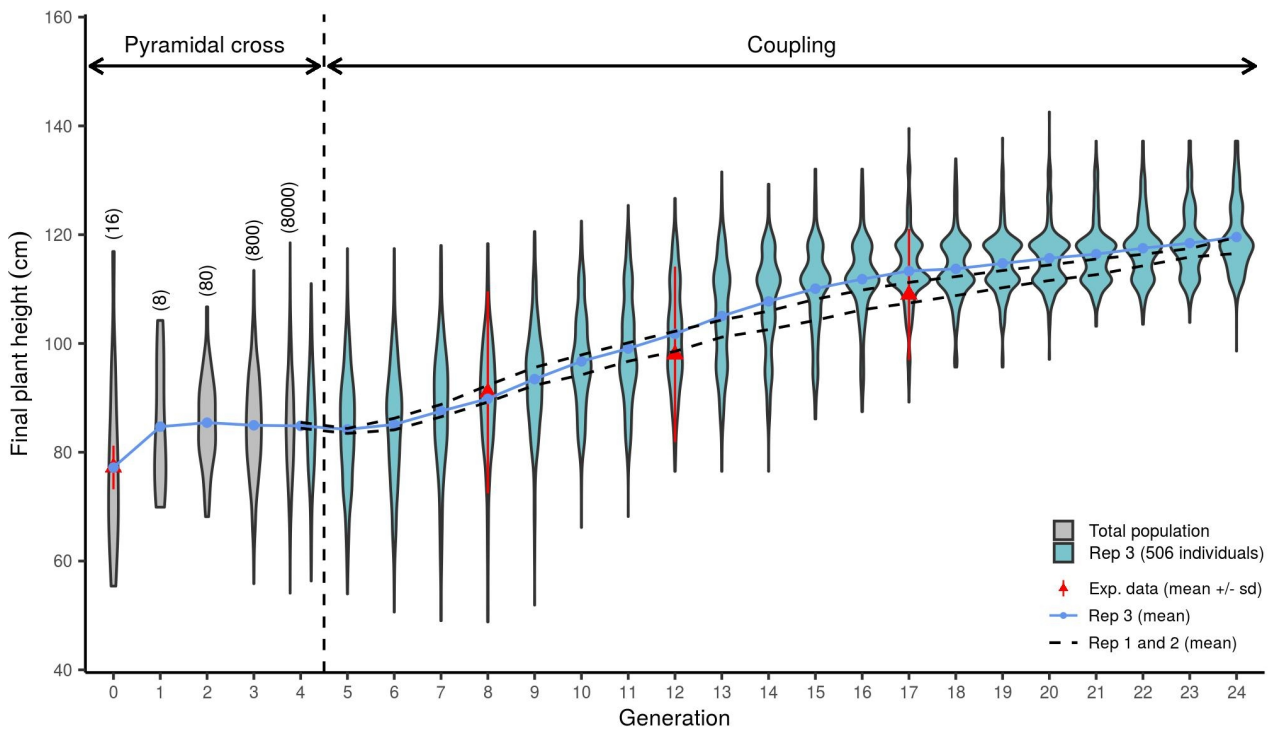


Figure 4. Evolution de la hauteur finale des plantes au cours des générations. La hauteur moyenne au cours des générations pour un des 3 réplicats simulés (rep 3) est représentée par une courbe bleue. La hauteur moyenne au cours des générations pour les 2 autres réplicats simulés est représentée par des courbes noires pointillées. Pour le croisement pyramidal, les distributions de hauteur pour chaque génération sont représentées en gris et le nombre d'individus pour chaque génération est indiqué entre parenthèses. Les distributions de hauteur pour chaque génération du réplicat 3 sont représentées en bleu. Les données expérimentales disponibles dans l'étude de [Raquin et al. \(2008\)](#) pour la hauteur moyenne et l'écart-type de hauteur sont représentés en rouge par des triangles et des segments respectivement.

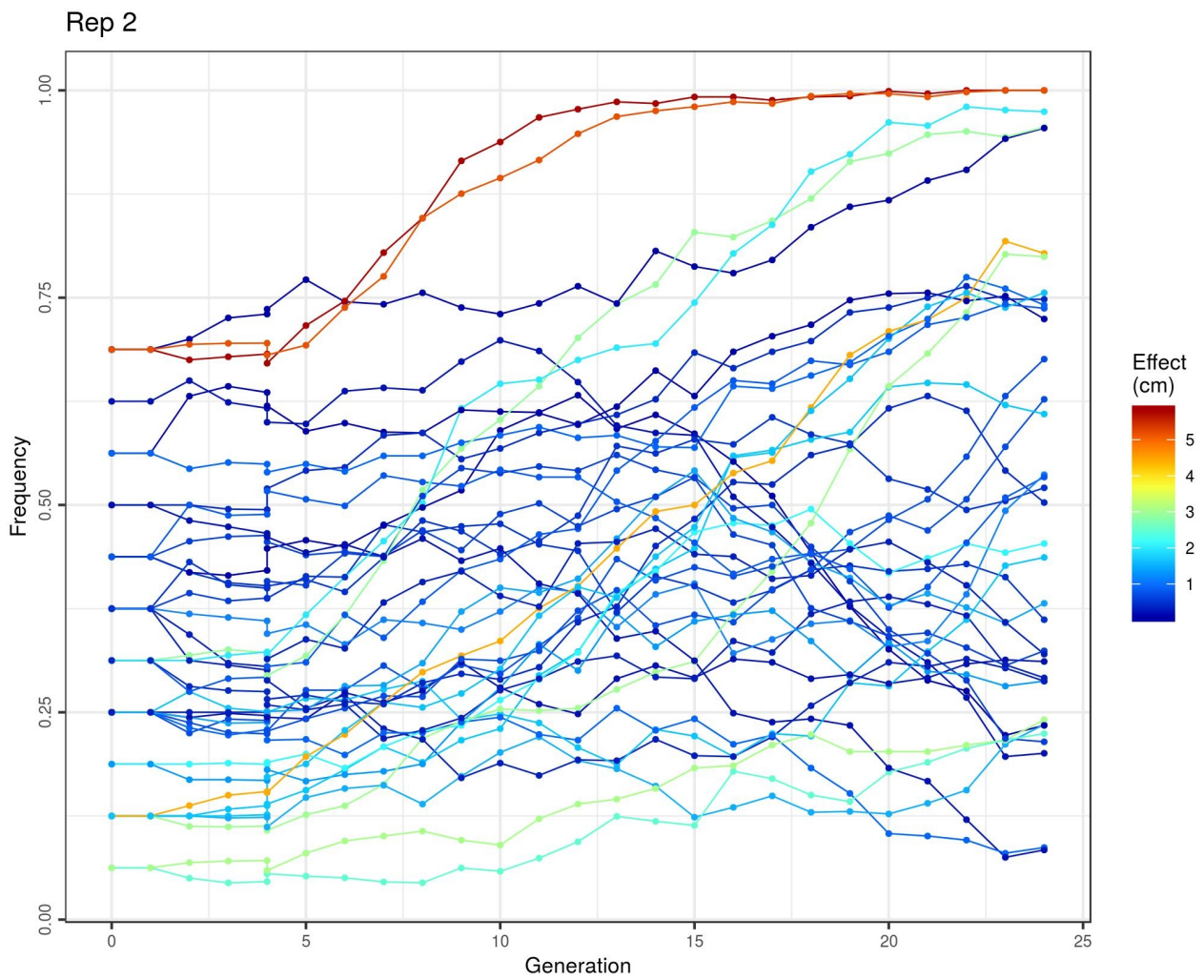


Figure 5. Evolution de la fréquence des allèles 'haut' de tous les gènes et QTL simulés au cours des générations pour un des 3 réplicats (Rep 2). La couleur des courbes indique l'effet de chaque marqueur sur la hauteur finale des plantes.

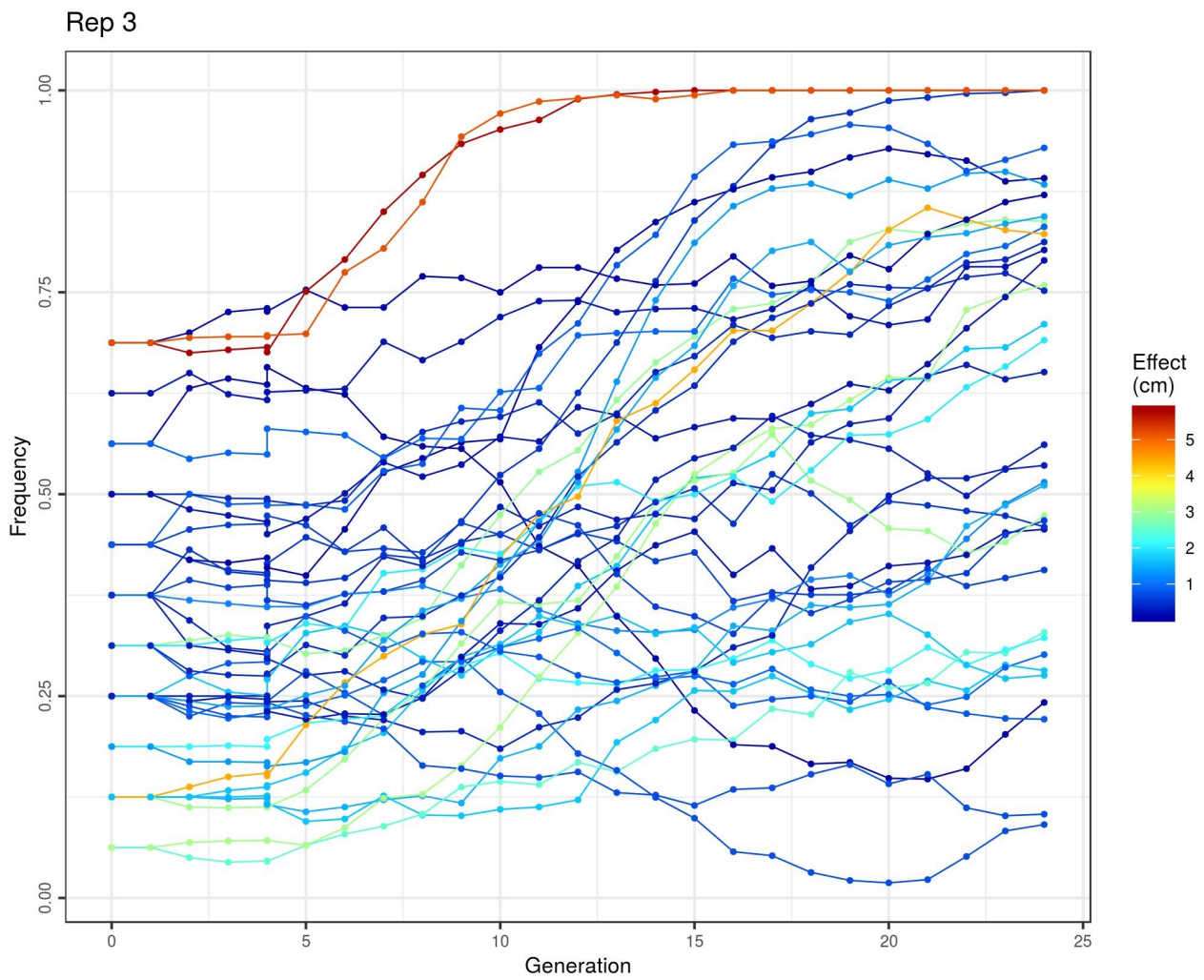


Figure 6. Evolution de la fréquence des allèles '*haut*' de tous les gènes et QTL simulés au cours des générations pour un des 3 réplicats (Rep 3). La couleur des courbes indique l'effet de chaque marqueur sur la hauteur finale des plantes.

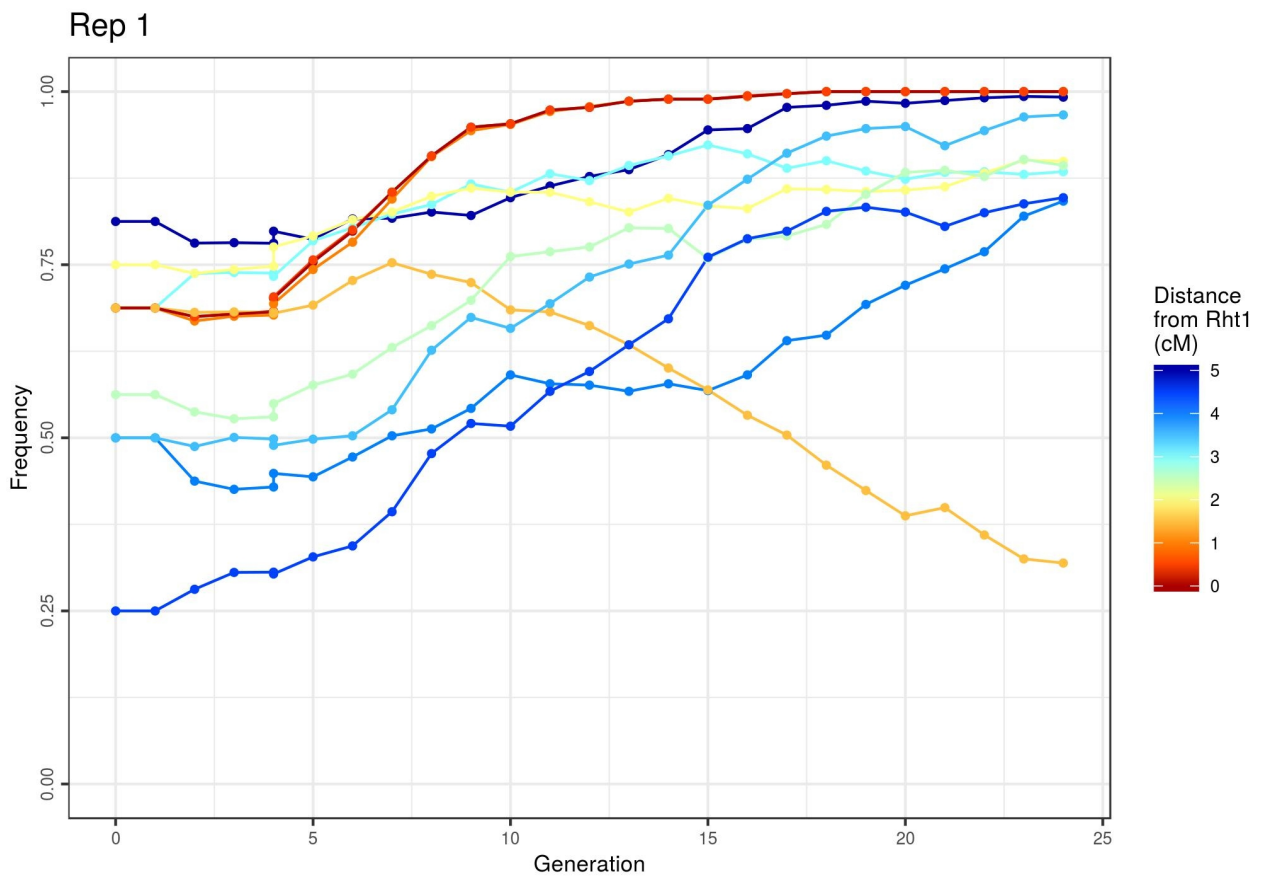


Figure 7. Evolution de la fréquence de l'allèle '*haut*' du gène Rht1 et des marqueurs neutres adjacents au cours des générations pour un des 3 réplicats (rep 1). La couleur des courbes renseigne sur la distance entre le marqueur et le gène Rht1.

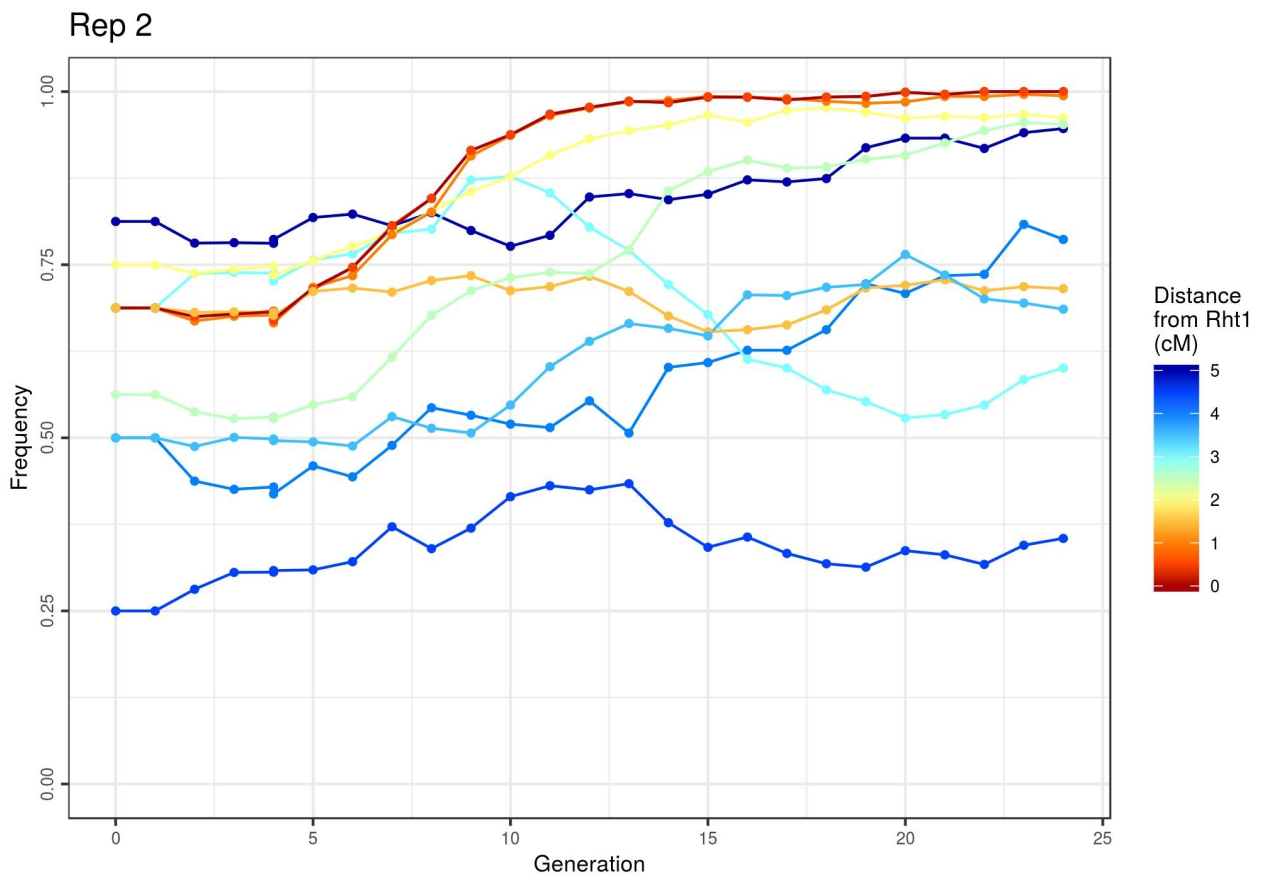


Figure 8. Evolution de la fréquence de l'allèle '*haut*' du gène Rht1 et des marqueurs neutres adjacents au cours des générations pour un des 3 réplicats (rep 1). La couleur des courbes renseigne sur la distance entre le marqueur et le gène Rht1.

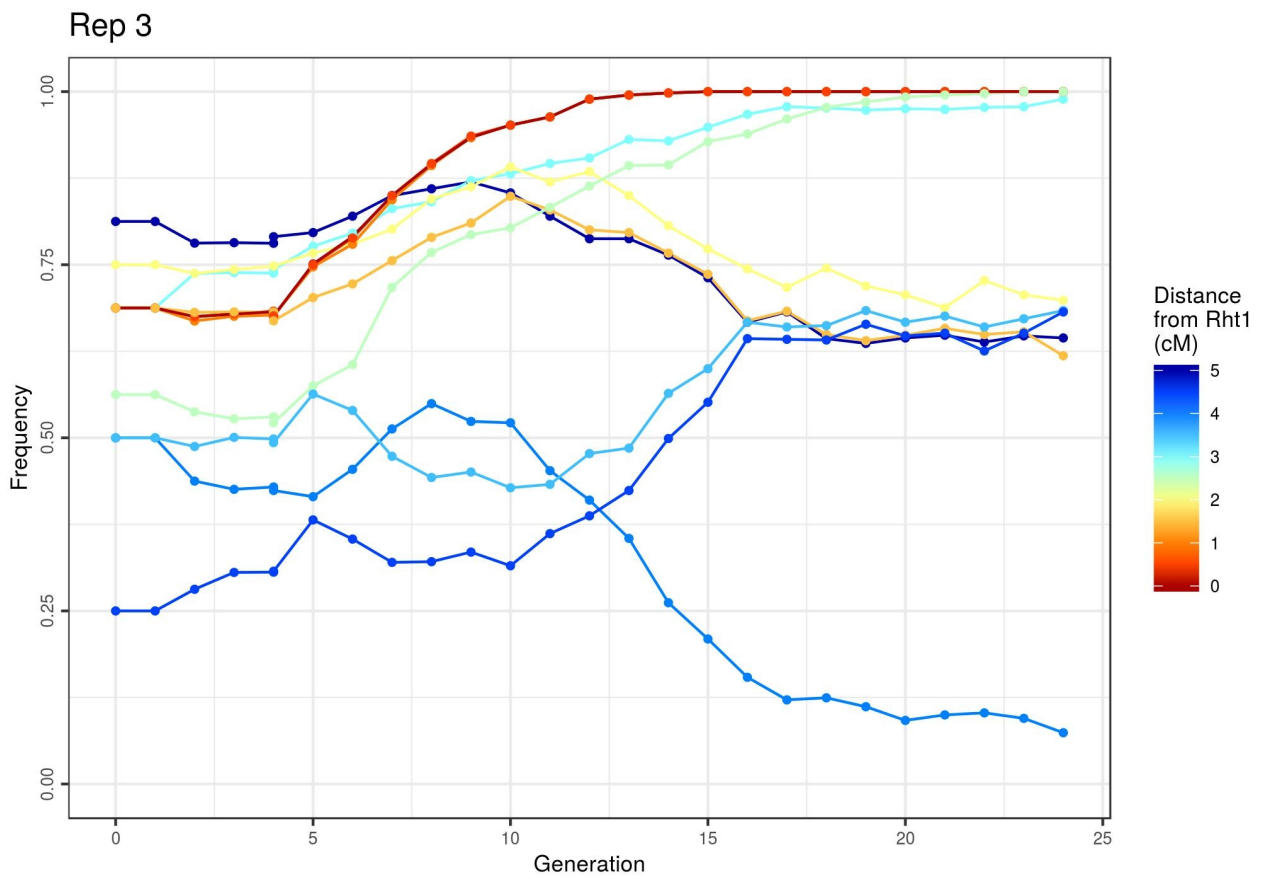


Figure 9. Evolution de la fréquence de l'allèle '*haut*' du gène Rht1 et des marqueurs neutres adjacents au cours des générations pour un des 3 réplicats (rep 1). La couleur des courbes renseigne sur la distance entre le marqueur et le gène Rht1.

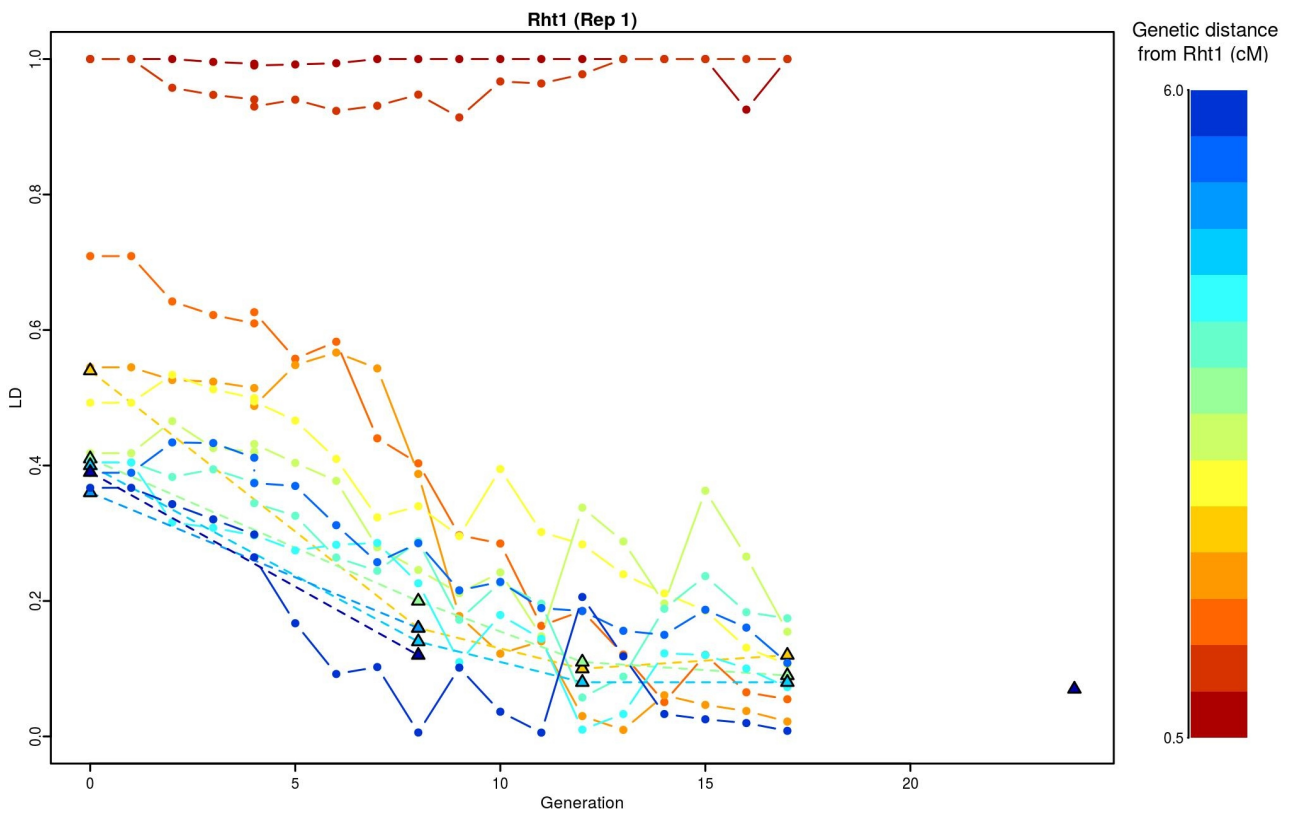


Figure 10. Déséquilibre de liaison entre Rht1 et les marqueurs neutres adjacents en fonction de la génération pour la rep 1. La distance par rapport à Rht1 est indiquée par la couleur. Les points reliés par des lignes continues correspondent au DL pour les données simulées du réplicat 1. Les triangles reliés par des lignes pointillées correspondent à des données expérimentales de [Raquin et al. \(2008\)](#).

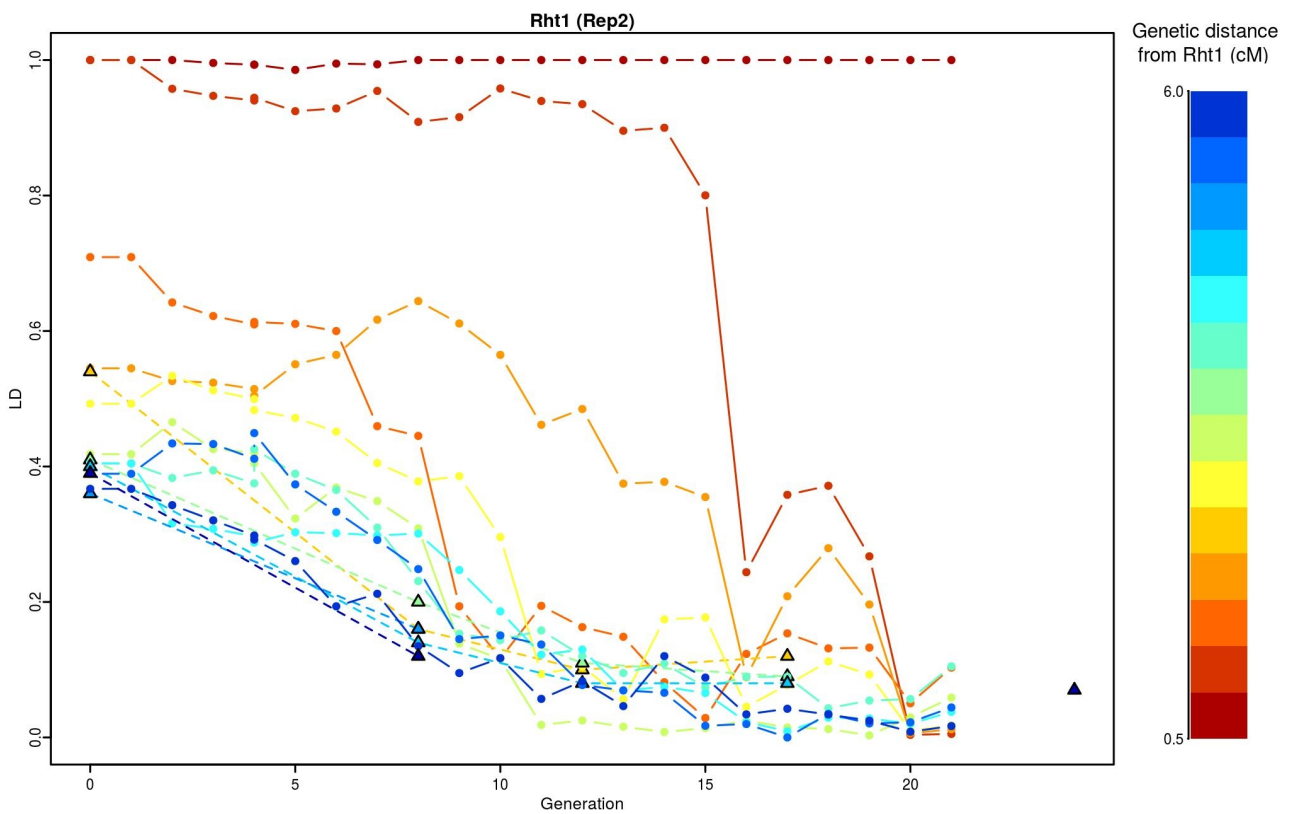


Figure 11. Déséquilibre de liaison entre Rht1 et les marqueurs neutres adjacents en fonction de la génération pour la rep 2. La distance par rapport à Rht1 est indiquée par la couleur. Les points reliés par des lignes continues correspondent au DL pour les données simulées du réplicat 2. Les triangles reliés par des lignes pointillées correspondent à des données expérimentales de [Raquin et al. \(2008\)](#).

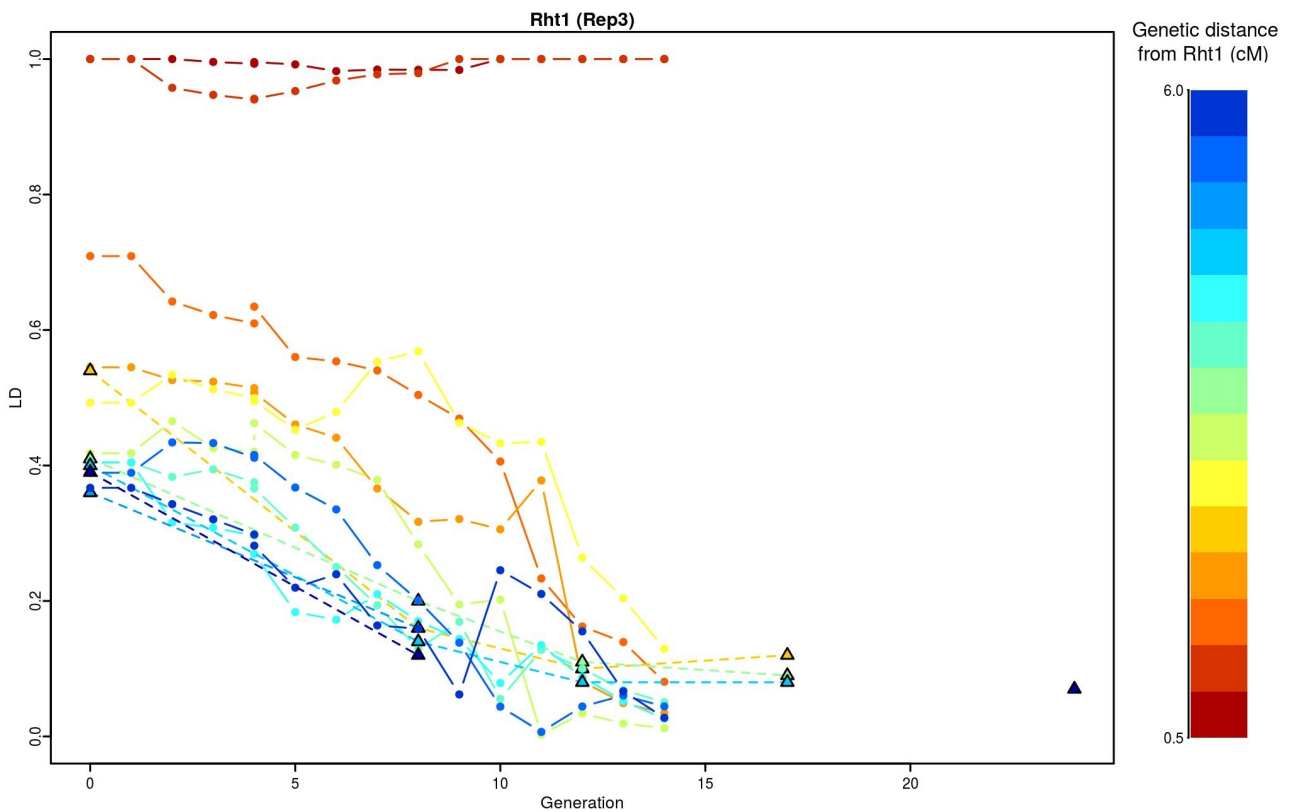


Figure 12. Déséquilibre de liaison entre Rht1 et les marqueurs neutres adjacents en fonction de la génération pour la rep 3. La distance par rapport à Rht1 est indiquée par la couleur. Les points reliés par des lignes continues correspondent au DL pour les données simulées du réplicat 3. Les triangles reliés par des lignes pointillées correspondent à des données expérimentales de Raquin *et al.* (2008).

Resemis d'un mélange binaire

Méthode

Afin d'illustrer les possibilités offertes par l'association de GenoWALT et de WALTer, et pour s'assurer du bon fonctionnement du couplage, l'évolution de la hauteur lors du resemis d'un mélange simple a été étudiée.

Le resemis d'un mélange binaire équilibré (2 variétés en proportions 50/50) a ainsi été simulé pour 20 générations successives. Le mélange était composé d'une variété courte, d'une hauteur finale de 80 cm^a, et d'une variété haute, d'une hauteur finale de 140 cm. Les deux génotypes du mélange simulé ont été générés à partir de la carte génétique décrite dans le Chapitre 6, en considérant que chaque variété était homozygote à tous les loci considérés. Hormis leur différence de hauteur finale, les deux variétés composant le mélange étaient parfaitement identiques et leurs valeurs de paramètres génotypiques étaient celles estimées pour le génotype Maxwell (Annexe 1). Les simulations ont été réalisées avec des parcelles de 110 plantes semées en maillage (distance inter-rang = distance entre plantes dans le rang) à une densité de 200 plantes/m². Le nombre de plantes était constant au cours des générations. Afin de tenir compte de la part d'aléatoire de GenoWALT, 6 réplicats ont été réalisés.

a Les valeurs indiquées correspondent à la hauteur du brin maître de la plante, à maturité.

Résultats et discussion

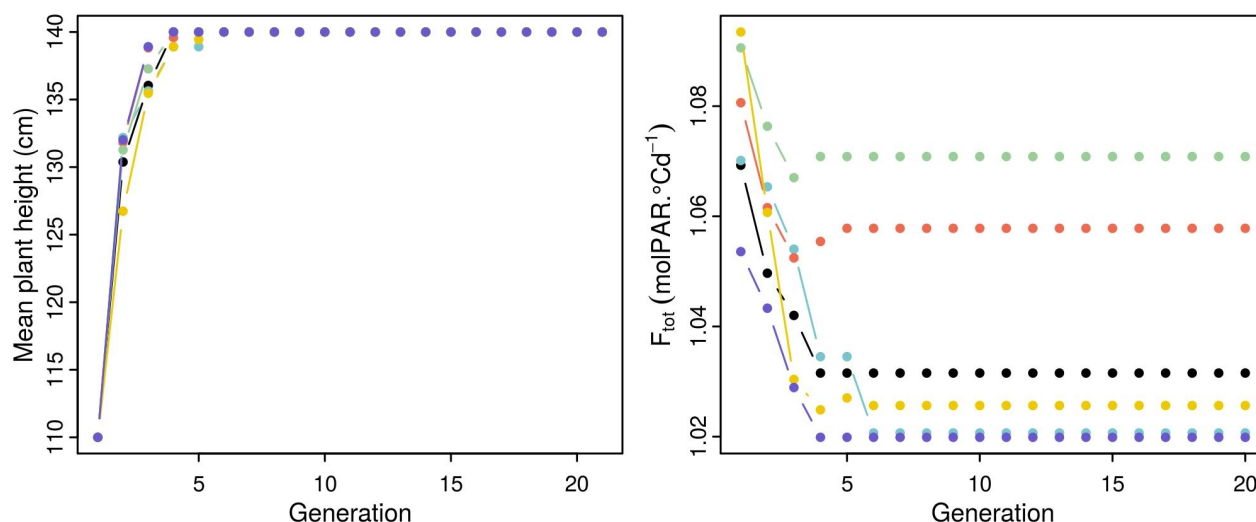


Figure 13. Évolution de la hauteur moyenne dans le couvert (gauche) et de F_{tot} (droite) au cours des générations de resemis. Les différentes couleurs de courbes représentent différents réplicats.

Pendant les 3 premières générations de resemis, la hauteur moyenne des plantes de la parcelle augmente rapidement jusqu'à atteindre la hauteur de la variété haute (140 cm). Dans le même temps, la fitness totale de la parcelle (F_{tot}) diminue. À partir de la 4^{ème} génération, il y a peu de variations de F_{tot} et de hauteur, et à partir de la 6^{ème} génération, tous les réplicats correspondent à des parcelles de culture pure avec seulement des individus de la variété haute. La variété haute, plus compétitive, a donc rapidement envahi la population. Ce processus s'est fait au détriment de la performance de la population, comme cela est prévu par la théorie de la tragédie des biens communs (TOC pour "tragedy of the commons" en anglais) (Anten et Vermeulen, 2016). Ce phénomène a été observé expérimentalement dans plusieurs études, dont celle de Goldringer *et al.* (2001) sur la population PA, pour laquelle l'augmentation de hauteur au cours des générations de resemis s'est accompagnée d'une baisse de rendement. La diminution de F_{tot} au cours des générations, bien que limitée (baisse de 3,6% en moyenne), est également en accord avec les simulations des Chapitres 5 et 6 qui ont montré une corrélation négative entre la hauteur moyenne de la parcelle et sa fitness totale.

Les résultats de cette expérience *in silico* de resemis d'un mélange binaire simple sont donc cohérents et indiquent un bon fonctionnement du couplage entre WALTER et GenoWALT. De prochaines simulations sur le resemis de mélanges plus complexes, avec une diversité sur plusieurs traits de l'architecture aérienne, permettraient de profiter pleinement des possibilités offertes par cette association entre FSPM et modèle de génétique des populations.

RÉFÉRENCES

Anten, N. P., & Vermeulen, P. J. (2016). Tragedies and crops: understanding natural selection to improve cropping systems. *Trends in Ecology & Evolution*, 31(6), 429-439.

Feingold, J. (1991). Le déséquilibre de liaison. *Médecine/Sciences*, 7, 161–168.

Goldringer, I., Enjalbert, J., Raquin, A. L., & Brabant, P. (2001). Strong selection in wheat populations during ten generations of dynamic management. *Genetics Selection Evolution*, 33(1), S441-S463.

Titre : Modélisation des conséquences évolutives de la compétition pour la lumière dans des populations composites de blé

Mots clés : Couplage de modèles, compétition inter-génotypique, génétique des populations, FSPM - modèle structure-fonction, analyse de sensibilité

Résumé : La diversité intra-spécifique, mobilisée à l'échelle de la parcelle (populations et associations variétales) devrait jouer un rôle important dans la transition agroécologique, en contribuant à l'adaptabilité et à la résilience aux stress des systèmes agricoles à bas niveaux d'intrants. Toutefois, la complexité des interactions entre plantes dans ces peuplements soulève de nombreuses questions sur leur conception et leur gestion, pour lesquelles l'approche par modélisation est particulièrement pertinente, mais peu développée à ce jour. Cette étude porte ainsi sur la modélisation structure-fonction et son utilisation pour étudier l'impact de l'architecture aérienne des plantes sur la performance et l'évolution de peuplements hétérogènes de blé. Ce projet de thèse est basé sur l'utilisation de WALTer : un modèle de plante structure-fonction préexistant qui représente le développement d'une parcelle de blé du semis à maturité et simule la plasticité du tallage (propriété de ramification) en réponse à la compétition pour la lumière. Dans un premier temps, des adaptations ont été effectuées sur WALTer pour réduire le coût computationnel des simulations, augmenter le réalisme du modèle et permettre la simulation de peuplements hétérogènes complexes. Le fonctionnement de cette nouvelle version de WALTer a ensuite été caractérisé en analysant l'impact de l'aléatoire sur les simulations et en réalisant une analyse de sensibilité sur des parcelles homogènes. D'autre part, une méthode de calibration automatique, peu développée pour ce type de modèle, a été appliquée à WALTer sur la base de données expérimentales afin d'estimer les valeurs des paramètres précédemment identifiés comme étant importants par l'analyse de sensibilité. WALTer a ensuite été utilisé pour identifier, via une analyse de sensibilité sur des mélanges binaires équilibrés, les traits

de l'architecture aérienne (décrits par leurs valeurs moyennes dans le mélange et les valeurs de différences entre les 2 variétés) qui ont le plus d'impact sur la performance des mélanges. L'analyse a ainsi révélé le rôle prépondérant de la capacité de tallage et des dimensions foliaires comme déterminants de la performance des mélanges simulés. Grâce à une phase d'optimisation, les combinaisons de traits les plus favorables pour la performance des mélanges simulés ont été identifiées. Les résultats des simulations ont ainsi souligné l'importance de sélectionner soigneusement les combinaisons de traits associés en mélange, en montrant l'intérêt potentiel d'une diversité d'architecture dans les mélanges, mais également des effets défavorables possibles d'une telle diversité. Enfin, WALTer a été couplé à un modèle de génétique des populations afin de simuler l'évolution de peuplements hétérogènes de blé au cours de générations de resemis. Le couplage a été utilisé pour simuler une expérience de gestion dynamique d'une population présentant une grande diversité de hauteur. L'évolution génétique et phénotypique de la population sur 20 générations a ainsi pu être comparée à des données expérimentales pour confirmer le bon fonctionnement du couplage. Cette thèse a permis de développer des méthodes et outils pour la modélisation structure-fonction des peuplements hétérogènes de blé. Les simulations réalisées ont également apporté des éléments de compréhension sur l'impact de l'architecture aérienne sur la compétition pour la lumière dans ces peuplements. À terme, ces travaux seront utiles pour la conception et la gestion des peuplements hétérogènes de blé en champ.

Title : Modelling evolutionary outcomes of competition for light in wheat composite populations

Keywords : Model coupling, inter-genotypic competition, population genetics, FSPM – functional-structural plant model, sensitivity analysis

Abstract : Intra-specific diversity, mobilized at the scale of the field (cultivar mixtures and composite populations) should play an important role in the agroecological transition, contributing to the adaptation and resilience of low input cropping systems, more prone to stresses. However, many questions are raised by the complex plant-to-plant interactions occurring in such heterogeneous canopies, and there is a growing need to develop modelling approaches to improve their design and management. Thus, our study focuses on the use of functional-structural plant modelling to study the impact of the aerial architecture of plants on the performance and the evolution of heterogeneous wheat populations. The thesis project is based on WALTer: a pre-existing functional-structural plant model that simulates the development of a wheat field from sowing to maturity and the plasticity of tillering (i.e. branching) in response to competition for light. In a first step, modifications were made to WALTer to reduce the computational cost of simulations, enhance its realism and improve its ability to simulate complex heterogeneous stands. The functioning of this new version of WALTer was then characterized by analysing the impact of stochasticity on the simulations and by performing a sensitivity analysis on homogeneous plots. Furthermore, on the basis of field observations, an automatic calibration method, which is rarely developed for this type of model, was applied to WALTer in order to estimate the values of the parameters previously identified as important by the sensitivity analysis. Thanks to a sensitivity analysis of balanced binary mixtures, WALTer was then used to identify the traits of the aerial architecture (described by their mean values in the mixture

and the values of the differences between the 2 varieties) that have the most impact on the performance of the plots. The analysis thus revealed the preponderant role of tillering capability and dimensions of the leaves as determinants of the performance of the simulated mixtures. Thanks to an optimization phase, the combinations of traits that are the most favourable for the performance of the simulated mixtures were identified. The results of the simulations highlighted the importance of carefully selecting combinations of traits to associate in mixtures, showing the potential interest of a diversity of architecture in heterogeneous stands, but also the possible adverse effects of such diversity. Finally, WALTer was coupled to a population genetics model to simulate the evolution of heterogeneous wheat stands along generations of re-sowing. The coupling was used to simulate the dynamic management of a population presenting a large diversity of plant height. The genetic and phenotypic evolution of the population over 20 generations of re-sowing could thus be compared with experimental data to confirm the good functioning of the coupling. This thesis allowed the development of methods and tools for the functional-structural modelling of heterogeneous wheat stands. The simulations carried out have also improved our understanding of the impact of the aerial architecture on the competition for light in these stands. Ultimately, this work will be useful for the design and management of heterogeneous wheat stands adapted to farmers' needs.