



HAL
open science

Spatialization methods of crop models and metrics for evaluating spatialized crop model performances in a precision agriculture context

Daniel Pasquel

► **To cite this version:**

Daniel Pasquel. Spatialization methods of crop models and metrics for evaluating spatialized crop model performances in a precision agriculture context. Environmental Engineering. Montpellier SupAgro, 2023. English. NNT : 2023NSAM0027 . tel-04448152

HAL Id: tel-04448152

<https://theses.hal.science/tel-04448152>

Submitted on 9 Feb 2024

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'INSTITUT AGRO MONTPELLIER ET DE L'UNIVERSITE DE MONTPELLIER

En Sciences Agronomiques

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par

Les unités de recherche

ITAP – Technologie et méthodes pour les agricultures de demain
MISTEA – Mathématiques, informatique et statistique pour l'environnement et l'agronomie

Spatialization methods of crop models and metrics for evaluating spatialized crop model performances in a precision agriculture context

Présentée par Daniel PASQUEL
Le 05 octobre 2023

Sous la direction de James A. TAYLOR
Et le co-encadrement de Sébastien ROUX et Bruno TISSEYRE

Devant le jury composé de

André CHANZY, Directeur de recherche, INRAE, France

Raffaele CASA, Professeur, University of Tuscia, Italie

Taru PALOSUO, Principal Scientist, Luke, Finlande

Alessia PEREGO, Associate Professor, University of Milan, Italie

Julie CONSTANTIN, Chargée de recherche, INRAE, France

James A. TAYLOR, Directeur de recherche, INRAE, France

Davide CAMMARANO, Professeur, Aarhus University, Danemark

Sébastien ROUX, Ingénieur de recherche, INRAE, France

Rapporteur (Président)

Rapporteur

Examinatrice

Examinatrice

Examinatrice

Directeur de thèse

Invité

Invité



UNIVERSITÉ
DE MONTPELLIER

L'INSTITUT
agro Montpellier

Acknowledgements

Déjà trois années que mon projet de thèse a commencé et voilà qu'il arrive à son terme, bien que ce ne soit que le début du chemin. Mais c'est une occasion particulière pour remercier l'ensemble des personnes qui m'ont montré la voix et aidé dans cet accomplissement.

Tout d'abord à mon encadrement de thèse.

Un immense merci à James, tu as su être un remarquable directeur de thèse. Trouver l'équilibre entre autonomie et assistance n'est pas toujours simple et tu as su parfaitement le faire. Tu m'as beaucoup apporté pour mener à bien ce projet, je te remercie grandement. Un très grand merci à Sébastien pour ces nombreuses heures passées devant des tableaux à trouver des solutions à des problèmes qui semblaient à la base insolubles, merci pour toute ton aide ! Et aussi merci beaucoup à Bruno pour ton esprit critique, ton soutien infaillible et toutes tes relectures assidues.

Also I would like to give a special thanks to Davide, I learn so much about the crop modeling with you, thank you for welcomed me in Aarhus University, it was an awesome experience.

Un grand merci à vous, vous avez été hors-pair sur cet encadrement de thèse. Vous m'avez énormément appris ! Je suis fier d'avoir pu travailler avec vous et j'espère pouvoir continuer nos collaborations par la suite.

A mes très chers collègues.

Une thèse c'est éprouvant mais heureusement il y a les collègues pour décompresser ! Un grand merci à l'équipe AgroTIC pour toutes les pauses café, repas à la cantine aux sujets de conversations toujours très ... instructifs ! Merci à Simon, Yoann, Sarah, Léo, Jean-Philippe, Thomas, Arnaud, Laure, Pauline, Romain, Aurélien pour toutes ces heures passées ensemble. Heureux d'avoir pu prendre part aux compétitions de GeoGuesser, Pédantix, EuroVision, et tant d'autres ... je me souviendrai toujours de ce repas aux Grands Buffets organisé (presque) exprès pour mon anniversaire ! Un merci aussi pour mes premières heures de TD passées avec Nicolas et Guilhem ! Je n'oublie bien entendu pas les collègues de Lavalette avec qui j'ai passé de très bons moments à l'autre bout de Montpellier (ce qui a permis de maintenir ma forme olympique à vélo). Un merci particulier à Héloïse ma très chère collègue de bureau 2 années durant qui m'a évité des crises d'hypoglycémie à de nombreuses reprises ! Merci aussi à Anice pour tes relectures minutieuses en parallèle pour gérer les coquilles et pour la gestion post pot !

Un grand merci à mes collègues doctorants, votre soutien aura été d'une grande aide à tous moments (mais on va pas se mentir particulièrement à la fin, merci de m'avoir supporté et écouté !). Merci à vous Cécile et Baptiste passés avant moi pour vos nombreux conseils avisés et merci aussi à Maxime, Yulin, Nicolas, Chloé, Clémence, Shérif, Belal et Leire, à qui je souhaite bon courage pour leur fin de thèse et bien évidemment nous continuerons à aller boire des bières c'est certain !

Un merci très chaleureux et tout spécifique à Eva, merci pour tout ton soutien pendant ces 3 années, tu es incroyable et je ne suis pas prêt d'oublier notre escapade new-yorkaise ; et à Victoria, merci pour ton soutien en toutes circonstances, ta joie, ta bonne humeur et ton rire si communicatif ! Je suis heureux d'avoir pu vous rencontrer au cours de cette expérience !

Mais aussi à mon lieu d'origine en recherche agronomique.

Je voudrais adresser un merci tout particulier à nos voisins de bâtiment, Aurélie et Léo, qui m'ont transmis autrefois leurs connaissances agronomiques. Sans vous aujourd'hui, je ne serai pas dans le milieu de la recherche et je vous remercie pour cela. Je vous suis reconnaissant d'avoir eu votre soutien dans ce projet et espère pouvoir retravailler avec vous un jour sur des thématiques d'agroécologie !

C'est l'occasion aussi de remercier mes autres collègues de l'UMR ABSys, Johanna, Elsa et Julie. Merci pour ces repas et cafés (et apéros) en votre compagnie, ils ont toujours été une vraie bouffée d'oxygène.

A l'indéfectible équipe des Ministres.

Un merci tout particulier à ma super coloc Gazzot, toujours prête à me ramasser à la petite cuillère quand il le faut et me soutenir dans toutes sortes d'épreuves ! A quand la SCI ?

Et merci à tout le reste de l'équipe sans qui mes moments en thèse à Montpellier auraient eu un goût plus fade notamment avec Max et ses découvertes musicales (et aussi aux Oudotte pour mes lieux de début et de fin de thèse forts appréciables), Iapteff et ses décoctions, Pitzouille et notre résidence secondaire, Kelvin et son potager qui m'a nourri quelques mois, Alex et sa détermination aux jeux de société (tu pensais à autre chose ?), Gabin et ses toasts incroyables, Lulu et sa bonne compagnie sur le dancefloor, Mimi et son soutien sans faille pour aller en soirée, Hugo et ses apéros très proche du bureau, Pierre et son retour tant attendu, Remy et Juliette pour votre point de chute très bien placé et aux repas très fournis. Un grand merci à vous tous, votre soutien aura été plus que bénéfique !

A tous les autres, je ne vous oublie bien évidemment pas !

Vous avez été nombreuses et nombreux à me permettre de près ou de loin, sur le long ou court terme de mener à bien ce projet. Que ce soit en direct sur le campus de la Gaillarde à la cantine ou autour d'un café noisette : merci Lolo, Fafa et Guillaume pour ces petits interludes fort appréciables ! Ou alors pour des escapades fugaces pendant les périodes estivales avec la Team de Hoche : Hugues, Laura, Anna, Gaëlle, Palisson, Carla et Michel, merci pour tous nos rendez-vous annuels que j'apprécie tant ; à Paname où tout a commencé : Antoine, Asma, Guillaume et Doriane, merci d'être toujours présents après toutes ces années ; en pays arlésien : merci Sirine pour ces moments privilégiés en ta compagnie ; perdu dans le Massif Central : merci Nini pour la découverte de la région ou en Italie : merci Clairon pour tes chaleureux accueils et ta gentillesse sans égale. Je n'aurai évidemment pas assez de place pour remercier toutes les personnes qui le méritent, mais soyez en sûr je ne vous oublie pas pour autant même si vous n'êtes pas personnellement cité ici !

A ma famille.

Merci à ma famille qui a toujours su croire en moi. Ce n'aura pas été toujours de tout repos de trouver une voix qui me corresponde, mais on s'en approche petit à petit (et oui le chemin n'est pas encore réellement terminé ...). Merci à mes (nombreux) parents pour leurs conseils avisés et à tout le reste de ma famille sur qui j'ai toujours pu compter. Même si finalement tout le monde n'aura malheureusement pas pu voir l'aboutissement de mon projet de thèse, je sais au fond de moi que vous auriez été fiers si vous aviez encore été là pour voir cela. Merci pour tout ce que vous avez fait pour moi et merci de votre soutien toutes ces années durant !

Table of contents

Table of contents	4
List of Figures	7
List of Tables.....	10
Publication and international conferences.....	11
Short abstract.....	12
Résumé court.....	12
Abstract	13
Résumé.....	14
Foreword	15
Chapter 1 State of the art, research positioning and scientific questions of the PhD project.....	17
1.1. Intention note.....	17
1.2. A review of methods to evaluate crop model performance at multiple and changing scales	18
1.2.1. Details about the paper	18
1.2.2. Introduction	20
1.2.3. Issues and methods of model spatialization process	21
1.2.4. Evaluation of model performances.....	28
1.2.5. Conclusion.....	34
1.2.6. Acknowledgements	35
1.3. Research positioning and scientific questions of the PhD project.....	35
1.3.1. Research objectives of the PhD project.....	35
1.3.2. Research approach implemented for the PhD project	37
1.3.3. Manuscript sectioning	39
Chapter 2 Achieving a relevant assessment of spatialized crop model performance when simulating agronomic variables at changing scales	40
2.1. Intention note.....	40
2.2. Comparison of different aspatial and spatial indicators to assess performance of spatialized crop models at different within-field scales	41
2.2.1. Details about the paper	41
2.2.2. Introduction	42
2.2.3. Material and methods	43
2.2.4. Results and Discussion.....	47
2.2.5. Conclusion.....	51
2.2.6. Acknowledgements	51
2.3. A new metric to evaluate spatial crop model performances.....	52
2.3.1. Details about the paper	52
2.3.2. Introduction	52
2.3.3. Material and Methods.....	53
2.3.4. Results and discussion.....	57
2.3.5. Conclusion.....	58
2.3.6. Acknowledgements	58
2.4. General conclusion of the chapter	58

Chapter 3 Focus on spatial calibration approach as a downscaling process to spatialize crop models .	60
3.1. Intention note.....	60
3.2. What drives the performance of a crop model spatialized at the within field scale?	62
3.2.1. Details about the paper	62
3.2.2. Introduction	63
3.2.3. Material and Methods.....	65
3.2.4. Results	72
3.2.5. Discussion	77
3.2.6. Conclusion.....	80
3.2.7. Acknowledgements	80
3.3. General conclusion of the chapter	80
 Chapter 4 Simulation scale impact on the relevance of utilizing downscaled spatialized crop models at the within-field scale	 83
4.1. Intention note.....	83
4.2. Downscaling the APSIM crop model for simulation at the within-field scale.....	84
4.2.1. Details about the paper	84
4.2.2. Introduction	85
4.2.3. Material and Methods.....	88
4.2.4. Discussion	104
4.2.5. Conclusion.....	107
4.2.6. Acknowledgements	108
4.3. General conclusion of the chapter	108
 Chapter 5 General discussion and perspectives.....	 110
5.1. Retrospective on the results obtained during the PhD project.....	110
5.2. Discussion and contribution for a relevant evaluation and comparison of spatialized crop model performances	112
5.2.1. The needs for assessing spatialized crop models with a metric based on spatial pattern instead of spatial structure	112
5.2.2. The spatial balanced accuracy: expected limitations and benefits	112
5.3. Discussion and contribution of the results for downscaling existing crop models at within-field scale.....	113
5.3.1. Consistency of the spatial calibration approach for the spatialization of crop models at the within-field scale (compared to existing methods of downscaling)	113
5.3.2. Dependencies of the spatial calibration performances	116
5.4. Perspectives for further research	118
5.4.1. How to improve spatialized crop models evaluation?.....	118
5.4.2. How the spatial calibration approach could be improved to increase the performances of the spatialized crop model?	118
5.4.3. How the spatialized crop models could be involved in more operational decision-making in a precision agriculture context?.....	120

Chapter 6 Résumé étendu de la thèse	121
6.1. Introduction, contexte, positionnement et questions scientifiques du projet de thèse.....	121
6.1.1. Etat de l’art sur la spatialisation des modèles de culture et sur l’évaluation de leur performance prédictive	121
6.1.2. Positionnement et questions scientifiques du projet de thèse	122
6.2. Parvenir à une évaluation pertinente de la performance des modèles de culture spatialisés lors de la simulation de variables agronomiques à des échelles variables	124
6.2.1. Objectif du chapitre	124
6.2.2. Comparaison de différentes métriques pour évaluer la performance des modèles de culture spatialisés à différentes échelles intra-parcellaires.....	124
6.2.3. Présentation de la métrique Spatial Balanced Accuracy (SBA).....	125
6.2.4. Validation de la métrique SBA.....	125
6.3. Focus une approche de calibration spatiale comme processus de réduction d’échelle pour spatialiser les modèles de culture	126
6.3.1. Objectif du chapitre	126
6.3.2. Présentation de la méthode de calibration spatiale comme processus de réduction d’échelle des modèles de culture spatialisés.....	126
6.3.3. Quels sont les déterminants de la performance d’un modèle de culture spatialisé à l’échelle intra-parcellaire par processus de réduction d’échelle ?.....	127
6.4. Effet de l’échelle de simulation sur la pertinence d’utilisation des modèles de culture spatialisés par réduction d’échelle à un niveau intra-parcellaire	128
6.4.1. Objectif du chapitre	128
6.4.2. Utilisation du SBA et de l’approche de réduction d’échelle par calibration spatiale appliquées à une cas d’étude sur un modèle de culture complexe	128
6.5. Discussion et perspectives.....	130
Appendix A: Supplementary data of the Chapter 3.....	132
Appendix B: Supplementary data of the Chapter 4.....	135
References	141

List of Figures

Figure 1.1 Notion of crop model spatial footprint.....	22
Figure 1.2 Schematic illustrating the main reasons why users of point-based crop and environment models have decided to spatialize their models.	22
Figure 1.3 Illustration of scaling methods used for model spatialization.....	24
Figure 1.4 Schematic illustration of pathways, from data collection to final output, to apply spatialization processes to classical point-based crop models to obtain spatial model outputs.	25
Figure 1.5 Attribution of noise depending on observed PLWP values.	32
Figure 2.1 Experimental field with locations of observed predawn leaf water potential ($\Psi_{PD,obs}$) within-field measurements in a 1.2 ha Syrah vineyard at INRAE Pech Rouge (Gruissan, Aude, France) and the interpolated (kriged) maps of ancillary data used to define within-field zones.	43
Figure 2.2 Weather conditions for 2003 and 2004 for the period of measurement dates used for respectively model calibration (2003) and performance evaluation (2004).	44
Figure 2.3 Maps of different within-field zones defined with ancillary data using a segmentation algorithm for the Syrah vineyard.....	44
Figure 2.4 Different modeling cases of predawn leaf water potential (Ψ_{PD}) by WaLIS according to different spatial scales defined at the observation scale, the whole field scale or intermediate zonal scales (2 to 5 zones) based on ancillary data.	45
Figure 2.5 Ψ_{PD} values below which irrigation is required over the season.	47
Figure 2.6 Comparison between spatial structures of observed predawn leaf water potential ($\Psi_{PD,obs}$) and modeled predawn leaf water potential ($\Psi_{PD,mod}$) using the Cambardella index and Z-score.	48
Figure 2.7 Evaluation of the WaLIS simulations of the predawn leaf water potential (Ψ_{PD}) according to the different spatial scales along the measurement period.	49
Figure 2.8 Maps of within-field irrigation according to the date and the value of predawn leaf water potential for the observed point data and the modeled data at each considered scale of simulation.....	51
Figure 2.9 Difference between spatial structure and spatial pattern, the same spatial structure can result in different spatial patterns.	53
Figure 2.10 Representation of spatial balanced accuracy (SBA) computation.	56
Figure 2.11 Illustration of the simulated spatial agronomic variable maps used to evaluate theoretical spatialized crop models.	57
Figure 3.1 The method used to compare the classical calibration approach and the spatial calibration approach and to assess the relevance of this latter.	66
Figure 3.2 Experimental field of Syrah vineyard at INRAE Pech Rouge (Gruissan, Aude, France)...	67
Figure 3.3 Weather conditions for production seasons 2003 and 2004.....	688
Figure 3.4 Predawn leaf water potential (Ψ_{PD}) measured on the experimental vineyard on July 8 th 2004 and August 18 th 2004.....	70
Figure 3.5 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) regarding each considered spatial structure of Ψ_{PD} (strong, moderate and weak).	73
Figure 3.6 Change in the SBA distributions using a spatial calibration approach compared to a classical calibration approach depending on the spatial structure of the agronomic variable, the variability of the agronomic variable and the correlation between the agronomic variable and the ancillary data.....	75
Figure 3.7 Relevance of using a spatial calibration compared to a classical calibration regarding the difference of SBA score for different spatial scales of modelling.	76
Figure 3.8. Relevance of using a spatial calibration compared to a classical calibration regarding the difference of RRMSE for different spatial scales of modeling for a date of high σ_{Ψ}	77

Figure 4.1 Daily precipitation, air temperatures and solar radiation for 2005-2006 and 2007-2008 growing seasons for the experimental field.....	89
Figure 4.2 Durum wheat yield characterization.	90
Figure 4.3 Experimental field with the location of 100 measurement sites at within-field scale and the interpolated maps of apparent soil electrical conductivity in horizontal mode (EC _{aH}) and in vertical mode (EC _{aV}) using inverse distance weighting.	90
Figure 4.4 Example of a general soil profile shape for soil water properties used to model durum wheat yield with APSIM for one of the considered measurement sites.	91
Figure 4.5 Maps of different within-field zones defined using a segmentation algorithm based on soil apparent electrical conductivity data on horizontal and vertical mode together for the experimental field.	Erreur ! Signet non défini.
Figure 4.6 Different modeling scales of durum wheat yield using APSIM from site measurement observation scale, intermediate within-field scales (2 to 20 zones) defined with ancillary data related to soil characteristics up to the whole field scale.	94
Figure 4.7 Calibration steps followed for the spatial calibration of APSIM.	95
Figure 4.8 Maps of spatially calibrated fraction of plant available water able to be extracted (KL) profiles for each simulation scale on 2005-2006 data.	Erreur ! Signet non défini.
Figure 4.9 Maps of spatially calibrated drained upper limit (DUL) profiles for each simulation scale on 2005-2006 data.	Erreur ! Signet non défini.
Figure 4.10 Spatial patterns of durum wheat yield for observed data from 2007-2008 and modeled data from APSIM calibrated on 2005-2006 data for different modeling scales.	102
Figure 4.11 Spatial patterns of durum wheat yield for observed data from 2005-2006 and modeled data from APSIM calibrated on 2007-2008 data for different modeling scales.	103
Figure 6.1 Evaluation en contexte opérationnel de cartes d'irrigation intra-parcellaire en fonction de la date et de la valeur du potentiel hydrique foliaire (Ψ_{PD}) résultant des données observées et des données modélisées à différentes échelles spatiales.	124
Figure 6.2 Représentation du calcul de le Spatial Balanced Accuracy (SBA).	125
Figure 6.3 Approche de calibration spatiale pour la réduction d'échelle des modèles de culture pour la prédiction intra-parcellaire d'une variable agronomique.	127
Figure 6.4 Pertinence de l'utilisation d'une calibration spatiale par rapport à une calibration classique par rapport à la différence de SBA pour différentes échelles spatiales de modélisation avec WaLIS.	128
Figure 6.5 Exemple de résultats obtenus avec la version spatialisées d'APSIM. Motifs spatiaux du rendement du blé dur pour les données observées de 2007-2008 et les données modélisées par APSIM calibré sur les données de 2005-2006 pour différentes échelles de modélisation.	129

Figure S.A1 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a strong spatial structure of Ψ_{PD} .	135
Figure S.A2 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a moderate spatial structure of Ψ_{PD} .	135
Figure S.A3 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a weak spatial structure of Ψ_{PD} .	135
Figure S.B1 Different possibilities of fractions of plant available water able to be extracted (KL) from soil profiles for the durum wheat biomass calibration.	135
Figure S.B2 Different possibilities of soil hydraulic limits profiles for the durum wheat yield calibration.	135
Figure S.B3 Maps of spatially calibrated fraction of extractable plant available water (KL) profiles for each simulation scale on the 2007-2008 data.	136
Figure S.B4 Maps of spatially calibrated drained upper limit (DUL) profiles for each simulation scale on the 2007-2008 data.	136
Figure S.B5 Maps of soil hydraulic limits (SHL) profiles represented by difference between crop lower limit (CLL) and drained upper limit (DUL) on the first layer for each simulation scale.	137
Figure S.B6 Comparison between observed and calibrated durum wheat biomass from APSIM for prediction in 2007-2008 (calibrated on 2005-2006 data) for different modeling scales.	137
Figure S.B7 Comparison between observed and calibrated durum wheat yield from APSIM for prediction in 2007-2008 (calibrated on 2005-2006 data) for different modeling scales.	138
Figure S.B8 Comparison between observed and calibrated durum wheat biomass from APSIM prediction in 2005-2006 (calibrated on 2007-2008 data) for different modeling scales.	138
Figure S.B9 Comparison between observed and calibrated durum wheat yield from APSIM prediction in 2005-2006 (calibrated on 2007-2008 data) for different modeling scales.	139
Figure S.B10 Comparison between observed and modeled durum wheat yield from APSIM prediction in 2005-2006 data (calibrated on 2007-2008 data) for different modeling scales.	139
Figure S.B11 Comparison between observed and modeled durum wheat yield from APSIM predictions in 2007-2008 data (calibrated on 2005-2006 data) for different modeling scales.	140

List of Tables

Table 1.1 Indicators used to measure the distance between observed and simulated data.	30
Table 1.2 Observed field data and three simulated Predawn Leaf Water Potential (PLWP) models classified into Management Zones (MZs) based on a 3-class classification of the observed data (and the same thresholds used for the simulated data). Metrics of model fit are shown as the RMSE between observed and simulated data ($n = 49$) and Cohen’s Kappa value associated with the similarity of the simulated MZ models to the observed MZ model.....	32
Table 2.1 Calibration values for each scale simulation (except site-scale) and for each WaLIS calibrated parameters: maximum crop coefficient of the vine (K_c) and total transpirable soil water (TTSW).....	47
Table 2.2 Comparison of different metrics to identify the theoretical SCMs with the best performance. Each indicated value was computed between observed and modelled data.....	57
Table 3.1 Characteristics of the predawn leaf water potential (Ψ_{PD}) measured for both evaluation dates related to its variability and spatial structure. σ_Ψ = standard deviation of Ψ_{PD} (characterizing the variability), $C_{i,\Psi}$ = Cambardella index of Ψ_{PD} (characterizing the spatial structure).....	69
Table 4.1 Crop practices of the experimental field used as management information for the modeling. Adapted from Wallor et al., 2018.....	89
Table 4.2 Summary of calibrated and estimated parameters necessary as APSIM inputs to simulate durum wheat yield. Further details on the values of the aspatially calibrated parameters are available in Supplementary Table S.1 and S.2.....	97
Table 4.3 Spatial balanced accuracy (SBA) to assess calibration of APSIM spatialized version for simulating durum wheat biomass and yield at the field level and at different within-field spatial scales for the two different years.	101
Table 4.4 Root mean square error of prediction ($RMSE_p$) to assess the ability of the spatialized APSIM version to simulate durum wheat yield at field level and different within-field spatial scales for two different years of calibration and evaluation. Values indicated are in t/ha.	102
Table 4.5 Spatial balanced accuracy (SBA) scores to assess the ability of the spatialized APSIM version to simulate durum wheat yield at field scale and at different within-field spatial scales for two different years of calibration and evaluation.....	104
Tableau 6.1 Comparaison de différentes métriques pour identifier la modélisation théorique la plus performante. Modification par rapport au données observées : (A) ajout d’un bruit aléatoire, (B) ajout du même bruit de manière positive par rapport aux données observées, (C) ajout du même bruit de manière positive par rapport aux données observées avec ajout d’un biais numérique, (D) translation du motif spatial de 10 pixels et (E) données modélisées ayant la même structure spatiale mais des motifs spatiaux différents. Chaque valeur indiquée a été calculée entre les données observées et modélisées. Résultats extraits de Pasquel et al (2023).....	126
Table S.1 APSIM cultivar parameters aspatially calibrated for the 2005-2006 and 2007-2008 production years.	140
Table S.2 APSIM soil parameters aspatially calibrated for the both production years 2005-2006 and 2007-2008.....	140

Publication and international conferences

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B. and Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture*, 23(4), 1489–1513. <https://doi.org/10.1007/s11119-022-09885-4>.

Pasquel, D., Cammarano, D., Roux, S., Castrignanò, A., Tisseyre, B., Rinaldi, M., Troccoli, A. and Taylor, J.A. (2023). Downscaling the APSIM crop model for simulation at the within-field scale. *Agricultural System*, 212, 103773. <https://doi.org/10.1016/j.agsy.2023.103773>.

Pasquel, D., Roux, S., Tisseyre, B. and Taylor J. A. (2022). Comparison of different aspatial and spatial indicators to assess performance of spatialized crop models at different within-field scales. In: Proceedings of the 15th International Conference on Precision Agriculture, Minneapolis, Minnesota, United-States, June 26-29.

Pasquel, D., Roux, S., Tisseyre, B. and Taylor, J. A. (2023). A new metric to evaluate spatial crop model performances. In: Proceedings of the 14th European Conference on Precision Agriculture, Bologna, Italy, July 02-06.

Short abstract

Spatialization enables crop models to be applied at different spatial scales than their original, designed scale. Models can be upscaled or downscaled. Downscaling allows well-accepted, existing crop models to be used at finer modeling scales. This PhD project investigated how downscaling of crop models could work so that they can be used for precision agriculture. It realized that new metrics for evaluating the spatialized crop model performance were needed, as existing crop model metrics were limiting. To do this, the Spatial Balance Accuracy metric was proposed to assess both spatial and non-spatial errors in the crop model predictions. The project also proposed that spatial calibration can be used to spatialize crop models. Spatial data was used to define zones in fields where key model parameters were calibrated differently without changing the model equations. This proved to be a viable approach when the calibrated target variable itself exhibited a spatial structure.

Keywords Downscaling, Spatial patterns, Spatial calibration, Evaluation metrics

Résumé court

La spatialisation permet aux modèles de culture d'être appliqués à des échelles spatiales différentes de leur échelle d'origine. Les prédictions de ces modèles peuvent être sujettes à une réduction ou un agrandissement. La réduction d'échelle permet d'utiliser les modèles de culture existants bien acceptés à des échelles de modélisation plus fines. Cette thèse a étudié comment la réduction d'échelle de ces modèles pourrait fonctionner afin qu'ils puissent être utilisés pour l'agriculture de précision. Il a été montré que de nouvelles métriques pour évaluer les performances de ces modèles spatialisés étaient nécessaires, car les métriques existantes étaient limitantes. Pour ce faire, la métrique Spatial Balanced Accuracy a été proposée pour évaluer les erreurs spatiales et non spatiales dans les prédictions des modèles de culture. La thèse a également proposé que la calibration spatiale puisse être utilisée pour spatialiser ces modèles. Les données spatiales ont été utilisées pour définir les zones dans les parcelles où les paramètres clés du modèle ont été calibrés différemment sans changer les équations du modèle. Cela s'est avéré être une approche viable lorsque la variable cible calibrée elle-même présentait une structure spatiale.

Mots clefs Réduction d'échelle, Motifs spatiaux, Calibration spatiale, Métriques d'évaluation

Abstract

Crop models play a key role in simplifying and understanding complex agronomic systems. However, not all practitioners are interested in modeling agronomic variables at the same spatial scale. Changing the spatial scale at which such variables are modeled is therefore a necessary process to meet environmental and societal expectations. Spatialization enables a crop model to be applied at a different spatial scale from its native spatial footprint. More specifically, downscaling spatialization processes are identified as an opportunity to use existing crop models, initially designed at field scale use, at finer modeling scales (within-field scale) without modifying the internal structure of the model. This will permit a more tactical use of crop models for management, compared to their current, mainly strategic use. Particular attention was paid to mechanistic crop models, as they provide a better understanding of the biological, physiological and physical processes associated with the agronomic variables modeled. However, these biophysical crop process equations are generally designed at the field scale, and it is still unclear how a change of spatial resolution will particularly affect mechanistic crop models. This PhD project is based on the assumption that existing crop models are efficient and well recognized by the agronomic community. Thus, using them at finer spatial scales, by rethinking their use, would make it possible to employ these models into precision agriculture without having to use 'true' spatial crop models, which are more complicated to design. This led to the general research issue: is the spatialization of existing crop models, by using downscaling processes, conceivable and relevant for their use at within-field scales? However, the evaluation of the performance of these spatialized crop models at different scales needed to be rethought to take into account both aspatial and spatial pattern errors. These statements have led to the following specific scientific questions: how to perform a relevant evaluation and comparison of spatialized crop model performances across different spatial scales? And, is the spatial calibration of selected crop model parameters an effective method of downscaling existing crop models to permit modeling at within-field scales? Evaluation of the spatialized crop model performances at different spatial scales should be possible with the right metric. However, the metrics currently used are not the most relevant for assessing the performance of such models. A new metric has therefore been proposed: Spatial Balanced Accuracy (SBA). The SBA enables a relevant evaluation of spatialized crop models, taking into account the aspatial and the spatial pattern-based error of the considered variable(s). A spatial calibration approach was also implemented to downscale the spatial scale of two crop models, a simple and a complex model, to the within-field scale. This method proved successful, for both model types, when the modeled variable was strongly spatially structured and when ancillary data correlated with this variable were available. The intention was not to draw general conclusions on the spatialization of crop models, but to formalize this concept in a precision agriculture context and to build a basis for future research on the tactical use of these models at the within-field scale.

Résumé

Les modèles de culture jouent un rôle clé dans la simplification et la compréhension des systèmes agronomiques complexes. Cependant, tous les utilisateurs ne sont pas intéressés par la modélisation de variables agronomiques à la même échelle spatiale. Changer l'échelle spatiale à laquelle ces variables sont modélisées est donc un processus nécessaire pour répondre aux attentes environnementales et sociétales. La spatialisation permet d'appliquer un modèle de culture à une échelle spatiale différente de son empreinte spatiale native. Plus précisément, les processus de spatialisation par réduction d'échelle sont identifiés comme une opportunité d'utiliser les modèles de culture existants, initialement conçus à l'échelle de la parcelle, à des échelles de modélisation plus fines (échelle intra-parcellaire) sans modifier la structure interne du modèle. Cela permettra une utilisation plus tactique des modèles de culture pour la gestion, par rapport à leur utilisation actuelle, principalement stratégique. Une attention particulière a été portée aux modèles de culture mécanistes, car ils permettent de mieux comprendre les processus biologiques, physiologiques et physiques associés aux variables agronomiques modélisées. Cependant, ces équations biophysiques des processus des cultures sont généralement conçues à l'échelle de la parcelle, et l'impact d'un changement de résolution spatiale est encore mal connu sur les prédictions des modèles de culture mécanistes. Ce projet de thèse est basé sur l'hypothèse que les modèles de culture existants sont efficaces et bien reconnus par la communauté agronomique. Ainsi, les utiliser à des échelles spatiales plus fines, en repensant leur utilisation, permettrait d'employer ces modèles en agriculture de précision sans avoir à recourir à de « vrais » modèles spatiaux de culture, plus compliqués à concevoir. Cela a conduit à la question générale de recherche : la spatialisation des modèles de culture existants, en utilisant des processus de descente d'échelle, est-elle envisageable et pertinente pour leur utilisation à des échelles intra-parcellaires ? Cependant, l'évaluation des performances de ces modèles de culture spatialisés à différentes échelles a dû être repensée pour prendre en compte les erreurs de modèle aspatiales et spatiales. Ces constats ont conduit aux questions scientifiques spécifiques suivantes : comment effectuer une évaluation et une comparaison pertinentes des performances des modèles de culture spatialisés à différentes échelles spatiales ? Et, est-ce que la calibration spatiale des paramètres du modèle de culture sélectionné est une méthode efficace de réduction d'échelle des modèles de culture existants pour permettre la modélisation à l'échelle intra-parcellaire ? L'évaluation des performances des modèles de culture spatialisés à différentes échelles spatiales devrait être possible avec la bonne métrique. Cependant, les métriques actuellement utilisées ne sont pas les plus pertinentes pour évaluer les performances de tels modèles. Une nouvelle métrique a donc été proposée : Spatial Balanced Accuracy (SBA). Le SBA permet une évaluation pertinente des modèles de culture spatialisés, en tenant compte de l'erreur aspatiale et spatiale de la (ou des) variable(s) considérée(s). Une approche de calibration spatiale a également été mise en œuvre pour réduire l'échelle spatiale de deux modèles de culture, un modèle simple et un complexe, à l'échelle intra-parcellaire. Cette méthode s'est avérée efficace, pour les deux types de modèles, lorsque la variable modélisée était fortement structurée spatialement et que des données auxiliaires corrélées à cette variable étaient disponibles. L'intention n'était pas de tirer des conclusions générales sur la spatialisation des modèles de culture, mais de formaliser ce concept dans un contexte d'agriculture de précision et de construire une base pour de futures recherches sur l'utilisation tactique de ces modèles à l'échelle intra-parcellaire.

Foreword

Agrosystems are complex systems in which several biotic and abiotic interactions combine to affect the crop production component. Moreover, agroecosystems tend to further increase the number of interactions and complicate the understanding of the production and environmental systems. Crop models have a key role to play in the simplification and understanding of complex agronomic production systems. Such crop models, whatever their underlying processes (e.g. statistical, mechanistic, machine learning based), are used to model agronomic variables by taking into account parameters and interactions that affect the considered agronomic variable. If well developed and correctly deployed, crop models can be used to answer specific agronomic questions, especially in cases where experimental measurements in a field are too time consuming and/or expensive to sufficiently answer a considered agronomic question. Modeling also has the advantage of being applicable across temporal dimensions to assess past or future production scenarios, which is particularly important in the understanding of agricultural systems in a context of climate change or longer-term implications of modifications to cultural practices. Therefore, crop models are useful tools to provide data to facilitate agricultural system understanding.

Precision agriculture aims to increase the sustainability of agrosystems and input efficiency by adapting cultural practices at the within-field scale. Used within decision support tools, crop models can be used by farmers to support their site-specific management strategy. However, crop models are commonly as point-based models and are initially designed and experimented at the field scale, i.e. the field is considered to be a homogeneous spatial modeling unit. Thus, they are designed to reproduce bioprocesses at that specific spatial scale. Using such crop models at spatial scales other than the field scale is made possible by spatialization processes. Many studies, related to climate change impact on crop production, have applied crop models on a larger scale (e.g. county, regional, national scale) through the use of upscaling methods. However, only a few studies have investigated the use of crop models at smaller spatial scales than the field scale using downscaling methods. Such spatialized crop models at the within-field scale should be able to reproduce within-field variability and the spatial pattern of the considered agronomic variable. If this is achieved, this will allow spatialized crop models to be used as another data layer for short-term, in-season spatial management. In this manuscript, spatialization refers to a change in the spatial resolution (footprint) of the crop modeling, i.e. there is no modification of the core internal crop model structure and it remains a point-based process. There is a distinction to be made with spatial models, which do take into account neighboring modeling units to compute the value of each considered modeling unit. This kind of model is often used in hydrology especially in relation to modeling water dynamic fluxes.

In this PhD project, there is an assumption made that spatialized (or spatial) crop models have an important point to play in site-specific crop management. Given this, and the investment spent in developing and improving existing crop models, coupled with recent development in spatializing these crop models; then the spatialization approach to generating site-specific crop predictions seems more appropriate than redeveloping spatial crop models, which would require significant changes to the core crop model equations. A focus on mechanistic (also called process-based) crop models was also made for their ability to understand the underlying biophysical processes within model sub-processes and equations. Therefore, the term 'crop models' is used in this manuscript to refer to process-based crop models. For parts of this work, synthetic data had to be generated to investigate specific subjects. These synthetic data will be designated as 'simulated data', while 'modeled data' will be reserved for use with crop model outputs, i.e. the considered agronomic variable modeled by the crop model.

To address the scientific questions posed in this project (Section 1.3.1), this manuscript is divided into five chapters. The first chapter presents a global introduction emphasizing the context of the crop model spatialization. It highlights scientific questions that arise from using and evaluating spatialized crop model data. The next three chapters deal with research developed during the PhD project. This research is divided into two broad issues. The first one addresses theoretical aspects related to performance evaluation of spatialized crop models, while the second addresses methods for spatializing crop models to within-field scales for precision agriculture purposes. These four chapters are made of one or two scientific papers (in total there are three published and two submitted papers within the manuscript). Each chapter is introduced by an intention note and finished with a conclusion to position the published/submitted papers within the global scientific approach of the PhD project. The last chapter consists of a general discussion on the PhD project outcomes and its contribution to precision agriculture and crop modeling. It concludes with some perspectives linked to future research needs.

Chapter 1

State of the art, research positioning and scientific questions of the PhD project

1.1. Intention note

Crop models are commonly used to advise farmers, policymakers or agronomic researchers in their decision-making thanks to their capacities to facilitate understanding of bio physiological crop processes (Silva and Giller, 2020). However, all of these practitioners are not necessarily interested in modeling at the same spatial scales because their decisions do not have the same spatial reach. In the literature, many studies highlight the spatial dependency of crop models and wonder if these crop models are able to correctly model the spatial variation whatever the considered spatial scale (Challinor et al., 2017; Han et al., 2019; Xu et al., 2020; You et al., 2022). Crop models can be based on various approaches that could be statistical (Lobell et al., 2008; Ray et al., 2015), based on machine learning processes (Maimaitijiang et al., 2020; van Klompenburg et al., 2020) or mechanistic (Jones et al., 2017a). The question about the change in spatial resolution is particularly relevant for mechanistic crop models, because equations describing crop biophysical processes are generally designed at the field scale. Many studies have used these kind of crop models at other spatial scales by using spatialization methods (Acevedo-Opazo et al., 2008; Baralon et al., 2012; Basso et al., 2001; Chen et al., 2017; Huard et al., 2019). Moreover, the goals and methods to achieve the spatialization of such crop models are diverse. To our knowledge, there is no state of the art allowing to identify and classify in a logical way the different approaches proposed in the literature. Thus, a first working assumption is that reviewing the reasons and the methods employed by practitioners and crop modelers will be useful to clarify and understand how and why spatialization processes are used. Particular attention will be paid to mechanistic (also called process-based) crop models. Process-based crop models are designed with equations that describe processes, i.e. agronomic variables are modeled by considering driving parameters and variables (e.g. sunlight, temperature, water, nutrients, etc.) and their interactions at the field scale. The choice to focus on these types of crop models within the PhD project was made because these models have a better potential to understand the biological, physiological and physical processes associated with the modeled agronomic variable compared to others crop model types (e.g. statistical or machine learning-based models). However, the use of mechanistic crop models in a spatialization process does raise questions of whether or not the described processes are scale-independent or if different processes and interactions could occur at different spatial scales.

Changing the spatial resolution (footprint) of a point-based model also raises questions about the evaluation of the modeling performances. Spatialized crop models should maintain a spatial consistency in the modeled agronomic variable. However, classical metrics used to evaluate crop model performances do not explicitly take into account spatial data characteristics. Thus, these classical metrics could lead to bias in the interpretation of the quality of spatialized crop models at different scales. Therefore, another working assumption was that by reviewing how spatialized crop models are assessed could show a relevant assessment of such models and if not, could identify the statistical needs for a relevant evaluation. The assessment of spatialized crop models could be improved by using metrics that account for both aspatial and spatial error in model outputs.

To address these questions, under the assumptions outlined above, a literature review was performed (Section 1.2). Questions highlighted in the literature review were used to inform and to specify the research positioning and scientific questions of the PhD project (Section 1.3). This includes the context of the research, objectives of the project and the specific scientific questions addressed (Section 1.3.1.), the general approach implemented across the project (Section 1.3.2.) and a general description of layout of the thesis (Section 1.3.3).

1.2. A review of methods to evaluate crop model performance at multiple and changing scales

1.2.1. Details about the paper

1.2.1.1. Title and publication information

The following section was published in *Precision Agriculture* as a literature review in February 2022. The full citation is:

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B. and Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. Precision Agriculture, 23(4), 1489–1513. <https://doi.org/10.1007/s11119-022-09885-4>.

1.2.1.2. Authors

Pasquel D.^{1,*}, Roux S.², Richetti J.³, Cammarano D.^{4,5}, Tisseyre B.¹, Taylor J.A.¹

¹ ITAP, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

² MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

³ CSIRO Agriculture & Food, Floreat, WA, Australia

⁴ Department of Agronomy, Purdue University, West Lafayette, IN, USA

⁵ Department of Agroecology, iClimate, CBIO, Aarhus University, Tjele, Denmark

* Corresponding author: daniel.pasquel@inrae.fr

1.2.1.3. Abstract

Crop models are useful tools because they can help understand many complex processes by simulating them. They are mainly designed at a specific spatial scale, the field. But with the new spatial data being made available in modern agriculture, they are being more and more applied at multiple and changing scales. These applications range from typically at broader scales, to perform regional or national studies, or at finer scales to develop modern site-specific management approaches. These new approaches to the application of crop models raise new questions concerning the evaluation of their performance, particularly for downscaled applications. This article first reviews the reasons why practitioners decide to spatialize crop models and the main methods they have used to do this, which questions the best place of the spatialization process in the modelling framework. A strong focus is then given to the evaluation of these spatialized crop models. Evaluation metrics, including the consideration of dedicated sensitivity indices are reviewed from the published studies. Using a simple example of a spatialized crop model being used to define management zones in precision viticulture, it is shown that classical model evaluation involving aspatial indices (e.g. the RMSE) is not sufficient to characterize the model performance in this context. A focus is made at the end of the review on potentialities that a complementary evaluation could bring in a precision agriculture context.

1.2.1.4. Keywords

Spatialization, Scaling methods, Crop model uncertainty, Sensitivity analysis, Spatial pattern

1.2.1.5. Glossary

Crop model: A system-based model that is used to simulate daily dynamic interactions between the “soil-plant-atmosphere-management” (Wallach et al., 2019). In a broader sense, they complement field experiments and can be used to extrapolate/integrate observed data.

Point-based model: A model designed to predict variables on a point (i.e. a unit support) without taking into account neighboring data or effects to compute the variable. This is what is considered a ‘classical’ crop model.

Spatialization (or model spatialization): Means to apply point-based models spatially across an area different from the native model area (unit support) on which it was designed.

Spatialized model: A point-based model that has had a spatialization process applied to it (see above).

Spatial model: A model designed to compute output variables taking into account neighboring data or/and effects.

Spatial footprint: The native model spatial area.

Finer scale: Corresponds to a more accurate resolution (disaggregation method, smaller pixels) or the use of a model on a smaller scale than which the model was initially designed.

Larger scale: Corresponds to a more coarse resolution (aggregation method, larger pixels) or the use of a model on a larger scale than which the model was initially designed.

Processes: Corresponds to an activity ensemble which are correlated or interactive. These activities can be biological, environmental, physical or chemical.

The term ‘*variable*’ is only used to describe a model input or model output.

A ‘*state variable*’ is a variable that is internal (calculated) and not an input and is not necessarily expected to be given as a model output.

The term ‘*parameter*’ is only used to describe a part of a model. Parameters can be calibrated using different methods.

Calibration: A process to find the best values of model parameters by using the observed data, a consequence of the calibration is that simulated data are better fitted to observed data.

Resolution: Refers to the minimum scale of both spatial and/or temporal phenomena (i.e. is non-specific). If a statement only refers to a spatial or a temporal phenomenon, then it will be described as such (i.e. spatial resolution or temporal resolution), otherwise it may mean either or both.

Data assimilation: A suite of methods to combine simulated data from a model and observed data. It aims to find an optimal combination between both to improve model predictions (e.g. by recalibrating or updating a model) (Huang et al., 2019).

Data fusion: Methods to combine data from different sources into an integrated and unified compound with higher quality of information (Bleiholder and Naumann, 2009; Oliveira et al., 2021).

Evaluation: Refers to the question of knowing how well model predictions are relevant with measures collected in real-world situations. The aim is to ascertain the value computed by models.

1.2.2. Introduction

In many scientific domains, including agronomy, environmental sciences and hydrology, models are a way to simplify reality through a series of assumptions and by representing processes (Bouman et al., 1996; Sinclair and Seligman, 1996; van Ittersum and Donatelli, 2003). These models are often necessary to answer a specific question and are designed around this objective. Models come in various forms. Statistical models (also called empirical models) use a mathematical relation between different variables. The principal drawback of statistical models is that they are designed on observed data and are ill suited for use in sites or applications that were not involved in the model parameterization and development. They also cannot predict values in an uncertain context (e.g. impact of climate change on crop growth) (Jones et al., 2017a). In contrast, purely mechanistic models (also called process-based models) rely on the modelling of biophysical processes. They are based on mathematical equations that describe physiological processes (and not on mathematical equations that simply link two variables, as in statistical modelling). They can be derived in the absence of any real data as long as the process has been described. Mechanistic models can be deterministic, i.e. without random variations within the model and equations, so that for a given set of inputs, the result will always be the same. Mechanistic models can also be stochastic, i.e. the models and equations include random effects, so that results will change between simulations even if the inputs remain constant. Most of the crop models are a combination between process-based and empirical models, resulting in mechanistic deterministic models.

Models are useful tools in agro-environmental fields because they can help understand many complex processes by simulating them. Indeed, models can be used as a surrogate to estimate data that are hard, expensive or cumbersome to measure. Models account for relationships between crop growth and environmental, management and genetic factors. Therefore, there is a huge interest in using crop modelling to see how crop growth is impacted by these factors or to quantify ecosystem services. In other words, crop models are system-based models that aim to simulate interactions between the “soil-plant-atmosphere-management” (Hoogenboom, 2000; Wallach et al., 2019). To achieve this, multidisciplinary approaches are needed and crop models can take into account biological, physiological, ecological, physical or economical components. Integrating these approaches in crop modelling has led to the development of large crop models, such as APSIM (Holzworth et al., 2014), DSSAT (Hoogenboom et al., 2021, 2019; Jones et al., 2003), STICS (Brisson et al., 2003, 2002, 1998), WOFOST (de Wit et al., 2019), CropSyst (Stöckle et al., 2003) or AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009). Crop models are explanatory tools that are typically used in scenarios testing. For example, (Asseng et al., 2018) used an ensemble of crop models to understand the climate change impact and adaptation for wheat protein on a global level.

While crop modelling has become common within agricultural research domains for long-term strategic applications, it has traditionally been poorly used in shorter-term (single-season) production contexts (Asseng et al., 2013; Cammarano et al., 2020). This is changing. Modern agriculture has increasing access to data, including spatial data, which is providing increased possibilities to model agricultural systems, particularly using statistical modeling coupled to machine-learning approaches. It also provides an opportunity to integrate these data into conventional crop modelling platforms and to change the way these ‘traditional’ crop models can be used. One of the main ways that this is occurring is via the spatialization of crop models.

Most existing crop models are “point-based models” (Heuvelink et al., 2010). Spatialization is a way to apply these point-based models spatially across an area, by taking advantage of the new data available and applying these models to new scenarios without fundamentally changing the underlying model. Spatialization of crop models is of interest to the agricultural community as predictive crop modelling, particularly short to medium term predictions at field or subfield scales, is becoming an important part of modern site-specific management approaches. This is shifting the use of these

‘traditional’ crop models from long-term strategic applications, such as understanding long-term crop production potential under a changing climate, to short-term tactical applications and spatial applications. Examples of short-term tactical applications would be the determination of local fertilizer requirements given within-season production potential or generating in-season production estimates across local, regional or national scales to inform food security policy and actions.

The concept of model spatialization is not new. Faivre et al. (2004) defined spatializing a crop model as “using a crop model over areas larger than those over which it was developed”. At this stage, the idea was to upscale crop models from field/farm level modelling to regional level modelling. With a push toward precision agriculture, the concept of spatialization has evolved and became less restrictive. Within this review, spatialization is more simply defined as “using a crop model on another scale than which it was initially designed”. Thus, it could be applied at a larger scale or a smaller scale. This review will be focused on crop models but it is noted that the concepts developed could be applicable to any environmental models or other models in general.

Finally, the difference between spatializing a crop model and a spatial crop model is important. Point-based models do not take into account neighboring data or effects to compute a result at a point (or unit support) (Heuvelink et al., 2010). So with spatialized crop models, each point, regardless of its spatial footprint, is an independent simulation. An alternative would be to create crop models that do take into account spatial interactions between the unit supports to compute their results. These would be considered ‘true’ spatial crop models. However, this would require a fundamental change in the underlying crop model equations to achieve this and a considerable effort from the crop modelling community. Given the investment that has been made in current crop modelling platforms, short-term development seems better suited to spatializing crop models rather than redeveloping spatial models. Consequently, this review will focus on model evaluation with an emphasis on spatialized crop models, although it is recognized that some aspects in this review will be equally relevant to spatial crop model evaluation.

In the context of crop model use, shifting from strategic to tactical applications, model spatialization is expected to increase among agro-environmental models. Therefore, the purpose of this article is twofold: (i) to present an overview of different ways to spatialize a crop model and characterize more precisely spatialization methods and (ii) to review current ways that the outputs from these spatialized model are being evaluated and should be evaluated going forward. The article concludes with a comment on how these emerging spatialized crop models can be used in precision agriculture.

1.2.3. Issues and methods of model spatialization process

1.2.3.1. Model spatialization: goals and reasons

1.2.3.1.1. Crop model spatial footprint

Crop models were designed to understand and explain plant biophysical processes and developed with an assumption which considers a homogeneous unit support, i.e. same weather, soil and management in the simulated area; so they are point simulations or point-based models (Heuvelink et al., 2010). Crop models were also initially designed to operate at the field scale (van Ittersum and Donatelli, 2003). A common feature of crop models is that they have often been designed for a specific scale, and this scale refers to the scale of the processes that the models seek to predict. Nevertheless, Sinclair and Seligman (1996) highlight that processes described by models need to be described at a finer scale than the scale at which the models simulate outputs. Some models, such as DSSAT or STICS, were developed as a point represented by a small unit of support, for instance a homogeneous plot of one m², and then scaled to the field scale, and model the crop as a single entity grown within a field with homogeneous production conditions (Faivre et al., 2004). Others, e.g. the MAPP potato model (MacKerron et al., 2004), are based on small pot trials and simulate an individual plant, which is then grown in standard conditions at all points in the field. Note that this is not a spatialization of the model as all model inputs

and parameters are kept spatially constant. Regardless of whether it is modelled at the individual plant or the individual field/plot level, the observation implies that the crop model was designed to the scale of the specific object or process of study (e.g. leaf, plant, plot, field, watershed, region, etc.). For this review, the term spatial footprint of the model is defined as the scale of model outputs and conditions model inputs (Figure 1.1). Typically, the scale is the same for model inputs as for model outputs. Thus, to run a crop model, inputs need to correspond to the model spatial footprint and outputs will be obtained at the model spatial footprint scale. Therefore, if users need to have a different scale in output than the model spatial footprint, they will need to do some modifications, i.e. by spatialization (Ginaldi et al., 2019).

For this review, the spatialization process will not consider changes of scale to the molecular level, although the importance of the intersection of the ‘omics’ and the crop modelling community, especially using advances in high-throughput phenology platforms is noted. However, this intersection is more focused on advancing genetic improvements, rather than for crop management applications.

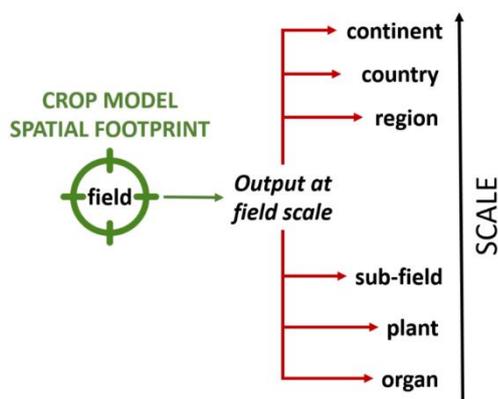


Figure 1.1 Notion of crop model spatial footprint. The spatial footprint is illustrated here with a crop model designed at the field scale. Model outputs in the native form are at the field scale and then spatialization (red arrows) can be used to change the scale of the outputs (color Figure online).

1.2.3.1.2. Reasons for spatialization presented in the agro-environmental literature

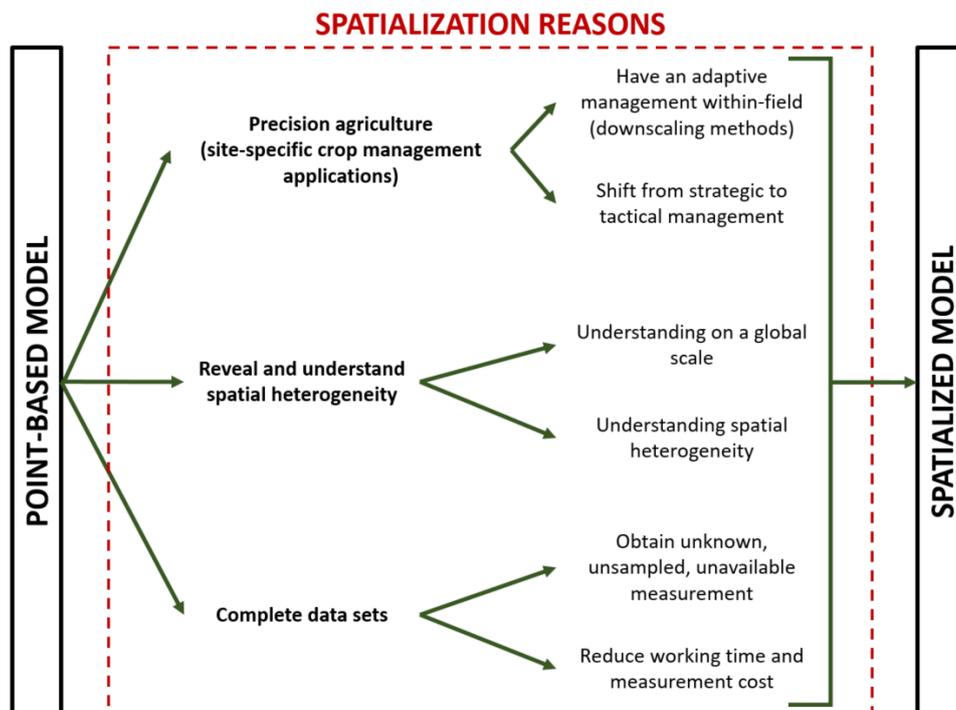


Figure 1.2 Schematic illustrating the main reasons why users of point-based crop and environment models have decided to spatialize their models.

Some users choose to spatialize crop models for specific purposes in order to obtain results that were not possible by using the crop model in its native spatial footprint. These reasons can be diverse but can be grouped together into several classes depending on the intended application (Figure 1.2).

Site-specific crop management applications: Users may aim to shift from strategic to tactical use of crop models with a desire to inform short-term management at finer spatial scales (site-specifically within fields). Thus, this refers to the use of crop models for precision agriculture purposes by aiming to have differential management across the field (Basso et al., 2011, 2001; Cammarano et al., 2021; Chen et al., 2017). For example, sensors are commonly used to provide variable rate applications of nitrogen (Colaço and Bramley, 2018), however, a spatialized crop model could be the main driver for the variable map or to understand spatial variability of soil-plant interactions.

Reveal and understand spatial heterogeneity: Models are often constructed to improve understanding of crop development, however, model phenomena and processes, and thus the understanding, are usually restricted to a specific scale (Balkovič et al., 2013). Nevertheless, users could wish to have an understanding on a finer scale of a phenomenon that is simulated over a relatively large area, e.g. with climate change simulations (Huard et al., 2019) and/or to characterize the local spatial heterogeneity (Li et al., 2020) by downscaling a variable that was originally too coarse. Therefore, model spatialization is proving to be useful for developing an understanding of processes at different spatial scales using both upscaling and downscaling methods (Blanchoud et al., 2020; Domínguez-Álvarez et al., 2021).

Complete data sets: As well as an improved understanding of processes, models can be used to predict unknown or unsampled points within a population or area. Spatialization can provide more accurate model simulations using spatially varying inputs within a known domain. When models have been calibrated and evaluated, their outputs can be used instead of real observations, thus model spatialization may be desirable to reduce the working time and cost of obtaining measurements in the field (Acevedo-Opazo et al., 2010, 2008; Baralon et al., 2012). Models can be used to obtain difficult, infeasible or unavailable measurements (Constantin et al., 2019).

1.2.3.2. Methods used for crop model spatialization

In order to spatialize crop models, different methods have been applied according to the practitioner's objectives. These methods can relate to either or both the inputs and outputs of the models and will lead to a change of scale in the model input or output variables via the use of scaling methods at one particular point in the modelling process or, alternatively, scaling methods can be used in succession within a modified framework of crop model spatialization. An example of this would be the successive variable transformation of model inputs (e.g. to calculate an unknown model variable with a known/measured variable) then a change of scale of the model input/output variables.

1.2.3.2.1. Change of scale of model outputs

This is the simplest method of spatialization, whereby the model is run in its native form, without any changes to the inputs, model equations or the form of the outputs (i.e. there is no change of input/output). Once the output has been computed the scaling is achieved via spatial processing (e.g. geostatistical operators) only. The scaling methods (Figure 1.3) for model spatialization can be classified into different categories depending on whether they increase or decrease the resolution: upscaling and downscaling methods (Blöschl, 2005; Ewert et al., 2011; Faivre et al., 2004). The aim of downscaling methods is to increase the variable resolution over a given area. Upscaling methods have the opposite goal, they generate a coarser resolution of the variables. Different approaches to up/downscaling have different consequences on the data and may lead to a change of extent, change of coverage or change of spatial resolution.

Change of extent: Extrapolation is used for this purpose and aims to give a prediction on a wider area (e.g. farm, regional, national, etc.) than the inputs. Predictions are made into areas outside the spatial coverage of the original observations, i.e. the extent becomes larger (Acevedo-Opazo et al., 2010; Baralon et al., 2012; Roux et al., 2019), but the quality of prediction may be uncertain. The inverse process is termed ‘singling out’ to reduce the extent of the observations. This is a simple extraction process and the data quality is equivalent to the original observation(s).

Change of coverage: Interpolation is used for this purpose and aims to provide estimates at locations where input variables are not available. Interpolation is performed over the entire area between known locations, for example by inverse distance weighting (IDW), kriging, spline functions or modern machine learning techniques. Reducing the coverage, or sub-setting the data, is performed using sampling approaches.

Change of spatial resolution: Aggregation aims to give a coarse prediction scale of an event or a phenomenon, for example, by averaging the finer scale data to the desired coarser scale. Disaggregation is the opposite, obtaining a finer prediction scale of an event or a phenomenon that of the basic model pixel, it can be achieved by simply resampling the coarser data, such that a 10 x 10 m pixel could be disaggregated into 100 pixels of 1 m² with the same value, or by trying to differentially partition values spatially across the finer scale grid using some form of disaggregation model (Malone et al., 2013).

The lack of fine resolution data for some inputs or low computational capacity relative to the large quantity of fine resolution data available for other inputs are reasons that have led to the use of upscaling in many studies (Grosz et al., 2017). Data aggregation can be useful or even necessary in order to simplify the understanding of the processes represented and to be able to draw applicable conclusions (Jankowski et al., 2001). Some issues are related to the spatialization of models when moving from a local scale to a more global scale, in particular when using aggregation. This raises questions of whether or not to use averaged data, in order to try to quantify heterogeneity, or to keep and use very fine resolution data (Allain et al., 2018). For instance, the over-simplification of the considered process is cited as critical to the use of aggregation and aggregated data (Scholes et al., 2013), but over-simplification may not be suitable for the intended model use.

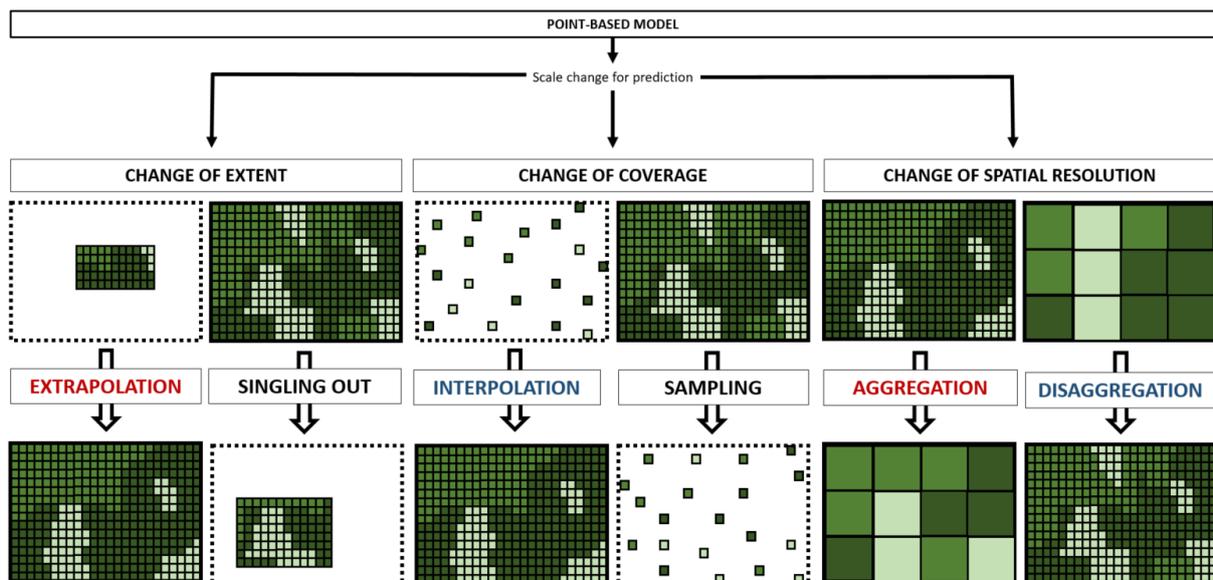


Figure 1.3 Illustration of scaling methods used for model spatialization. Red processes refer to upscaling methods and blue processes refer to downscaling methods that use a spatial process. Black processes refer to a change of scale via direct extraction without using a spatial process (color Figure online). Adapted from Faivre et al. (2004).

Change of scale can be used to spatialize input or/and output data of a crop model. Thus, a crop model can be spatialized by running independent simulations in each unit (or pixel) of the desired area using spatial inputs or based on coarse independent simulations, after which the outputs undergo a change in scale.

1.2.3.2.2. Spatial alterations of the crop modelling framework

The methods outlined in Figure 1.3 aim to manipulate the output data but could equally be applied to scale model inputs before running the model. This generates, as highlighted by Ewert et al. (2011), other potential methods to manipulate the models to achieve a spatialized crop models. These model manipulation methods correspond for instance to the modification of model parameters, the simplification of model structures or the use of nested models. Note that the processes in Figure 1.3 are relevant to scaling inputs or outputs into a spatialized (crop) model. It is also potentially relevant even if the crop model framework has already been spatialized or a spatial crop model is being used.

In the case of input modification, by rescaling the model inputs to take advantage of modern sensing technologies, such as satellite imagery, spatially explicit model input data can be generated. This then generates the question of how and when these spatially explicit data can be incorporated into the model and at what moment the spatialization process takes place in the modelling framework. Given the diversity of crop model types and approaches and the diversity in the type and availability of spatially explicit model input data, it is not surprising that in a very short time there have been various methods of crop model spatialization proposed. The variety of methods include, for instance, studies about vine water status at different scales (Acevedo-Opazo et al., 2010; Baralon et al., 2012), adaptation of wheat in a global warming context on a global scale (Asseng et al., 2018) or yield prediction at differing scales (Battude et al., 2016; Claverie et al., 2012). Figure 1.4 presents different crop model spatialization methodologies that have been synthesized from studies that have aimed to spatialize crop models.

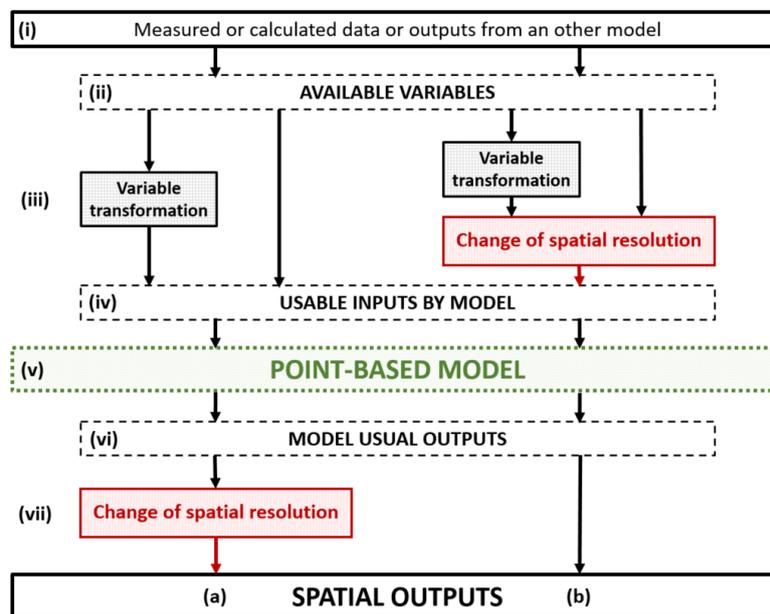


Figure 1.4 Schematic illustration of pathways, from data collection to final output, to apply spatialization processes to classical point-based crop models to obtain spatial model outputs. Common pathways (a and b) in the literature are indicated and represent the main spatialization framework, but other methods can be used inside these pathways. Red boxes and red arrows correspond to the moment where spatialization really occurs in the pathway. Change of spatial resolution refers to methods that change the data resolution by processes described in Figure 1.3 (extrapolation, interpolation, aggregation, disaggregation). Variable transformation refers to ancillary data being converted into model input variables. Black arrows correspond to simple transfers of data without changing data (color Figure online).

In Figure 1.4, the horizontal bars indicate the stages of the crop modelling process, from the collection of available data (top) to final model outputs (bottom), while the vertical arrows show potential pathways for modelling and the red arrows specifically indicate points where spatialization can occur. Figure 1.4 is constructed to indicate typical pathways for model spatialization. These include (a) a change of scale of model outputs and (b) a change of scale of the model inputs. Choosing whether to change the scale of model inputs or outputs or both is important because these methods will not have the same impact (Al-Shammari et al., 2021).

The available spatial variables can be either variables measured in the field, calculated data or output data from an upstream model (Figure 1.4(i)). However, these available spatial variables (Figure 1.4(ii)) may not necessarily be the same as the native variables (data) used as input by the original crop model (Figure 1.4(iv)). In some cases, these spatial data/variables are the same and can be directly used in the model. However, in a majority of cases, sensing and modelling systems do not directly measure the correct model variable at the correct spatial resolution to be usable by the crop model. Therefore, variable transformation may be required to obtain the correct variable to run the model (Figure 1.4(iii)). For example, many crop models use Leaf Area Index (LAI) but canopy sensors usually return a surrogate of LAI, such as a vegetation index (VI), at very high spatial resolutions. These available VI data may be subject to a mathematical relationship, for instance a transfer function, to modify and transform the VI values into LAI values for the considered crop. In the example of pathway (b), available variables may be subject to a change of resolution before being used in the model. Any approaches outlined in Figure 1.3 can be applied to the input data to change the support, coverage or extent of the model, leading to a change of spatial resolution of inputs (Teixeira et al., 2017) (Figure 1.4(iii)). Often this change of scale is possible by using other ancillary data; the available data can be coupled with these ancillary data, such as high-resolution remotely-sensed imagery, to try to reduce the uncertainty in the available spatial input data (Kasampalis et al., 2018).

Once the scale of usable model inputs has been correctly adjusted (Figure 1.4(iv)), the spatialized crop model can be run (Figure 1.4(v)) and the usual model outputs are computed (Figure 1.4(vi)). To reiterate, this is not a spatial modelling approach, but a punctual crop model applied at a different spatial resolution than its native design. The obtained output(s) (Figure 1.4(vi)) may not necessarily be at the scale desired by the user, as indicated in the pathway (a). Thus, the output data may also be subject to a change of spatial resolution. The approaches outlined in Figure 1.3 can be applied to the output data to change the support, coverage or extent of the model output. Some studies have compared strategies to aggregating input or output data and have highlighted only a few differences between these strategies (Angulo et al., 2013b; van Bussel et al., 2011).

Pathways (a) and (b) described in Figure 1.4 are the typical and shorter frameworks to spatialized predicted variables from crop models. They present simple versions where the change in spatial resolution is done either after (a) or before (b) the modelling step. A third simple pathway (not shown) could also be considered where changes in spatial resolution occur both before and after modelling, i.e. a cross-over approach in Figure 1.4 between pathways (a) and (b).

These are the simplest representations of model spatialization and more complex approaches are possible by adding other methods inside these pathways, because other data modifications may need to be done to obtain spatialized variables. For instance, in some studies, there is a resolution gap between different types of available data, especially weather-based inputs and other crop model inputs. Weather data is often designated at low resolutions (between 10 and 200 km), while crop and environmental inputs are designated at the field scale (Challinor et al., 2009). To bridge this gap, authors usually use scale transfer models, i.e. a corrective model on outputs from an upstream model, which involves upstream outputs that are generally too coarse in regards to the study area, to debias them using local variables from the area (Choukri et al., 2020; Huard et al., 2019).

An important point is how to perform model calibration when using a model in a spatialized context. The necessity for calibration is well known to improve crop model predictions, especially for variables estimated over large areas (Jagtap and Jones, 2002). Crop model calibration involves several key steps to improve predictions, whereas, practitioners do not necessarily have the same approaches (Seidel et al., 2018). Calibration requires a large amount of data when crop models are used on a large scale, but these data are often difficult to obtain at this scale. To tackle this issue of spatialized calibration, (Angulo et al., 2013a) tried three calibration strategies in an attempt to calibrate a crop model at the continental scale. Defining region-specific crop growth and phenology parameters, without considering output correction, improved the accuracy of crop model predictions on a large scale and seemed to be the best calibration strategy (Angulo et al., 2013a).

While this review is focused on methods to assess model spatialization, it is important to note the growing importance of data assimilation in the development of spatialized crop models (Jin et al., 2018). Data assimilation can be a method used inside the pathways (a) and (b). It is an approach used to recalibrate or to update a model to generate good short-term predictions. It is typically used for weather modelling, but is equally applicable to the shifting of strategic crop models to short-term tactical applications. To date, data assimilation has been mainly used for upscaling crop models to regional (Battude et al., 2016; Claverie et al., 2012) or national scales (for examples see Jin et al., 2018), but there is a growing interest in downscaling applications.

1.2.3.3. Uncertainty and error propagation when spatializing models

The uncertainty of a spatialized model will be a combined result of the model errors and the scaling errors, i.e. the uncertainty of the model itself plus the uncertainty of the scaled data plus the uncertainty of the spatialization method itself. Model uncertainty itself refers to parameter values and equations and will not be presented in this review.

1.2.3.3.1. Scaling errors

Scaling errors are linked to the methods used to scale model inputs and outputs. In some cases, a succession of scaling methods may be used and their combination will lead to an accumulation of uncertainty in the final result, which is often difficult to quantify (Ewert et al., 2011). The data aggregation effect (DAE) is a subject widely discussed in upscaling studies (Zhao et al., 2015) and this effect is linked to uncertainties introduced with the methods used to achieve aggregation. Many studies have focused on weather DAE (Ewert et al., 2011; Hoffmann et al., 2016, 2015; van Bussel et al., 2011; Zhao et al., 2015) because weather is an important driver in crop modelling and often observed at a large scale, whereas crop models are designed at the field scale. Therefore, using these large-scale data as an input in crop models could raise questions about the consequences of changing scale. Zhao et al. (2015) showed that DAE on weather inputs increased at coarser resolutions and was stronger with a higher spatial heterogeneity. However, some studies have shown that weather DAEs on crop yield and development are low (Ewert et al., 2011; Hoffmann et al., 2015). Apart from the weather, many studies have also focused on soil data. For instance, Grosz et al. (2017) showed that for soil organic content (SOC), DAE smoothed extreme values. As model inputs were aggregated to higher scales (from 10 to 100 km), the amount of heterogeneity in the model output(s) decreased. However, Grosz et al. (2017) demonstrated that aggregating to 50 km resulted in a higher variability than the reference aggregation (the computational scale was 1 km). Hoffmann et al. (2016) showed that, with soil data aggregation only, the bias of yield prediction was below 15%. However when weather data were aggregated in addition to these soil data, the bias increased (Hoffmann et al., 2016). This shows that the scaling may have a significant effect and can be complex because some variables can be overestimated at certain scales. Thus, a comparison between different aggregation scales could be a good approach to crop model evaluation (Al-Shammari et al., 2021).

An important issue to consider is how scaling errors will vary if scaling methods are applied on model inputs or outputs (Ewert et al., 2011). For instance, DAE on input data can be reduced using a coarse output resolution or aggregating model outputs, whereas these methods can only lead to a low reduction of model structure error, i.e. regrouping of model parameters and model equations (Grosz et al., 2017). If model manipulation (e.g. modifying model parameters, simplifying model structures or using nested models) is used to spatialize a model, using downscaling methods to match between the scale of upstream model outputs and the scale of downstream model inputs can increase the quality of downstream model outputs (Cammarano et al., 2017). In reality, scaling error impacts are a trade-off between a resolution fine enough to represent the spatial variability and an acceptable computational time (Grosz et al., 2017; Zhao et al., 2015). In some cases, spatial aggregation can reduce errors from deficient input data or model structure (Heuvelink, 2002).

1.2.3.3.2. Reduce model uncertainty by multi-model ensembles

Using multi-model ensembles (MMEs) rather than just one model is a quite new approach in crop modelling and has been enabled due to international cooperative modelling programs (Wallach et al., 2019, 2018). The more models there are, the more the prediction error decreases (Wallach et al., 2018). Studies using crop MMEs have shown that using indicators, such as the ensemble mean (e-mean) and ensemble median (e-median) of simulated data, produces better estimates than the use of indicators from a single crop model, even if it is the best available model (Martre et al., 2015; Wallach et al., 2018). These MMEs allow an increased accuracy of crop growth simulations (Martre et al., 2015) and so are useful to reduce the uncertainty introduced by error propagation. Improving the models used in MMEs, for example by re-calibration or incorporating or modifying simulated variables, can lead to a reduced number of models in these ensembles while simultaneously reducing uncertainty in the MME outputs (Maiorano et al., 2017). Examples of the successful use of small MMEs in agriculture include yield prediction and greenhouse gas emissions at the field scale (Ehrhardt et al., 2018).

1.2.4. Evaluation of model performances

1.2.4.1. Aim and importance of evaluating model performances

Model evaluation refers to the question of knowing how well model predictions are relevant to real observations, with the aim to ascertain the value computed by models. Moreover, this evaluation has to match with the proposed use of the model (Wallach et al., 2019). Model performances are case-dependent, so it is necessary to define at the beginning the model purpose (Bennett et al., 2013). The concepts of crop model ‘evaluation’ and ‘validation’ are slightly different. Validation refers to the process of determining if the model is adequate for its intended purpose or not and refers to the processes involved within the models (Tedeschi, 2006; Wallach et al., 2019). As argued by Wallach et al. (2019), crop models are never fully valid because they will always describe real world processes with assumptions and simplifications and thus are not identical to the real processes. Model evaluation is a black box concept and is not a question about the processes within the model but about the relevance of the model output (Wallach et al., 2019). Model inputs are subject to sources of uncertainty, such as measurement errors and inappropriate sampling resolutions (Crosetto et al., 2000). It is possible to consider parameter estimation when model evaluation is carried out; however, these are out of the scope of this review, and the focus in this review will be on output evaluation. Regarding output, evaluation can be performed qualitatively using graphs or quantitatively using indicators. Uncertainty and sensitivity analysis are part of the process of model evaluation (Wallach et al., 2019). These analyses aim to understand how variations in the output can be explained by variability in the model inputs.

To understand model evaluation, it is important to know what should be evaluated. To illustrate this, let’s take an example of irrigation decision-making using a water stress model compared to a threshold defined outside the model. The model simulates plant water stress and can have a wide uncertainty. However, this uncertainty will not ultimately change the final decision, which is to irrigate

or not, because the decision will depend on the model output relative to a threshold value identified by the decision-maker. So how should the performance of this model be evaluated? Only on the outputs from the predictive model? Or on the whole process that culminates with the final decision-making, which is ultimately the real action that is of interest to the agronomic community? These questions highlight that the method of evaluation has to match with the use of the model, i.e. if the model is used to estimate a variable then it is the variable that needs to be evaluated, but if the model is used to make a decision then it is the decision that needs to be evaluated, and not just the variable that was taken into account for the decision-making. This question is an important one when discussing model evaluation but it will not be detailed or discussed further in this review; indeed, this review is focused on evaluation of model output values.

1.2.4.2. Evaluation of methods used for evaluating spatialized crop model performance

1.2.4.2.1. Evaluation based on comparisons between observed and simulated data applied to crop models

The most common practice to evaluate a model is to compare observed data versus simulated data (outputs) using a metric or indicator that measures the distance between these observed and simulated data (Wallach et al., 2019). Various metrics exist for models in general, and many of these have been transferred for use in evaluating spatialized crop models. Table 1.1 reviews common metrics that have been reported in the spatialized crop model literature to date. In all these cases, model evaluation has been performed aspatially. There have been no spatial characteristics taken into account for model evaluation even though the crop models were being used in a spatialized context. This simplified utilization of aspatial indicators may affect the evaluation of spatialized crop models. Although multiple indicators are shown in Table 1.1, the indicators are not equally used. The RMSE was the most frequently used indicator in published studies.

1.2.4.2.2. Uncertainty and sensitivity analysis methods applied to crop models

Uncertainty analysis is used to quantify the global uncertainty in the model outputs in comparison to the uncertainty in model inputs (Crosetto et al., 2000). Sensitivity analysis (SA) is used to study how the model output variations can be assigned to different sources of input variations and how the model depends on its inputs (Crosetto et al., 2000). There are different ways of varying inputs: inputs can vary around a reference value, termed a local sensitivity analysis (LSA); or inputs can vary through and across a whole feasible domain, which is called a global sensitivity analysis (GSA) (Pianosi et al., 2016). SA can be used for different purposes, for instance evaluating the consistency of the model behavior or evaluating the robustness of model outputs depending on input uncertainty and model hypothesis (Pianosi et al., 2016). Thus, SA can be used as a form of model evaluation in various ways. For instance, SA can estimate if an input's impact on the model output is acceptable. It can also identify the key inputs with the most influence on the output(s) and can prompt users to consider if there is enough knowledge about these inputs to make a considered decision (Wallach et al., 2019). SA has been used in crop modelling studies in order to have a better understanding of uncertainty propagation and to determine impacts on simulated outputs (Acevedo-Opazo et al., 2010; Adam et al., 2011; Asseng et al., 2013; Baralon et al., 2012; Beaudoin et al., 2018; Duchemin et al., 2008; Teixeira et al., 2017). However, none of the reviewed literature on crop modelling has considered if there was a spatial component to the SA. Some questions arise from this observation such as: Is LSA suitable for assessing spatial effects? Should GSA be avoided in all situations?

Table 1.1 Indicators used to measure the distance between observed and simulated data.

Indicator measuring distance between simulated and observed data		References
Acronym	Description	Equation
Bias	-	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$
MAE	mean absolute error	$\frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $
MSE	mean square error	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
RMSE	root mean square error	\sqrt{MSE}
R ²	determination coefficients	$1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_{i, fitted})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
SEC	standard error of calibration	$\sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (y_{i,j} - \hat{y}_{i,j})^2}{(n \times m - 1)}}$
SEP _j	standard error of prediction (defined on spatial and temporal	$\sqrt{\frac{\sum_{i=1}^n (y_{i,j} - \hat{y}_{i,j})^2}{(n \times m)}}$
NSE	Nash-Sutcliffe efficiency	$1 - \frac{MSE}{MSE_{\bar{y}}}$
D-index	Willmott index of agreement	$1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y} + y_i - \bar{y})^2}$

y_i is the observed value, \hat{y}_i is the corresponding simulated value, $\hat{y}_{i, fitted}$ is the simulated value of y_i using a regression model between y_i and \hat{y}_i , n is the number of observations, m is the number of available data and \bar{y} is the average of observed values.

Some methods accounting for spatial characteristics have been used with environmental models. However, GSA was created to explain uncertainty in scalar outputs by variations of scalar inputs and so cannot directly be used with spatial models (Saint-Geours et al., 2012). To generalize GSA methods on spatial models, Saint-Geours et al. (2012) defined two sensitivity indices (SI): one on scalar inputs (i.e. a constant over the extent) and one on spatial inputs. Both are applicable as site SI (depending on an output point) and block SI (depending on the size of the spatial support defined by the upscaling process) in the case of upscaling process (average or sum of a point-based model) in the model output. This approach was applied to a point-based hydrology model. It was shown that the SI depends on the size of the spatial support and also that the uncertainty and the influence of the inputs on the outputs were spatially heterogeneous (Saint-Geours et al., 2014). Saint-Geours et al. (2014) showed that the SI of spatial inputs decreased with an increase in unit support size and the SI of aspatial inputs increased if the unit support size increased. Their study showed that a ratio of SI can be determined with the unit support size. Uncertainty analysis and SA on an environmental model at different scales have also shown that SI varies with the modeling scale (Şalap-Ayça and Jankowski, 2018). These results have raised important questions such as: Which unit size (and scale) for the output should the user choose to limit uncertainty propagation of spatial or aspatial inputs? What level of uncertainty are model users ready to accept to make a decision, and how is this uncertainty spatially distributed?

1.2.4.2.3. Why is an evaluation of spatialized models different from current model evaluation important?

As shown previously, the evaluation of spatialized crop models is currently done with aspatial indicators. Often, model evaluations are made without either accounting for a change of scale or the spatial character of the data. Tedeschi (2006) highlights that statistical analysis to evaluate predictive models is essential and needs to be appropriate for the model use in order to evaluate its precision and accuracy. For instance, there is an issue when input or output data are spatially autocorrelated, such that errors (i.e. difference between observed and simulated data) are not independent. The presence of this spatial autocorrelation can strongly reduce the reliability of many statistical metrics, including some popular ones shown in Table 1.1. Moreover, a lot of environmental variables present a continuity in their spatial structure so those variables are spatially dependent (Zhao et al., 2016).

Saint-Geours et al. (2014) showed that the output variance explained by spatial inputs decreases with an upscaling process, due to a data smoothing effect. This result highlights that evaluation should take into account changes of scale because model performance can depend on the scale at which it is run. The link between uncertainty propagation and scale change (upsampling and downscaling) is an area that requires more consideration (Saint-Geours et al., 2012).

To illustrate the issue and the need for new approaches to spatialized crop model evaluation, a simple case study is presented here. The aim is to demonstrate the limitation of aspatial statistics that have been widely used for evaluation of spatialized crop models in the recent literature (Table 1.1). In this case, the RMSE is used as the example statistic. In the case study, the intent is to define management zones (MZ) within a vineyard for precision viticulture. The predicted variable that is used to define these MZs is the predawn leaf water potential (PLWP). The purpose of this example is to show that with different theoretical spatialized models of PLWP, the outcomes of clustering based on PLWP predictions can be variable and independent of the RMSE.

This simulated example is built on observed data of PLWP on a 1.2 ha Shiraz vineyard in 2003. This vineyard is located in Pech Rouge (INRAE Gruissan, 43°08'47" N, 03°07'19" E) (See Acevedo-Opazo et al., 2010 for full details of the data set). To simulate the output from various theoretical spatialized crop models, three noise models were constructed, all built from the same values sampled from a normal distribution with a fixed mean (0) and variance (0.2). Various levels of spatial structure in the simulated PLWP were then obtained by altering how these noise distributions were associated with the observed PLWP values. Two of the noise distributions were dependent on the observed PLWP (some spatial structure), while the third distribution was random (i.e. independent from observed PLWP)

(Figure 1.5). These three noise models were added to the original data to simulate the output from three theoretical models. The original PLWP data were made into MZs based on a tertile analysis and the threshold values from this analysis was used to create the MZs in the simulated PLWP maps (Table 1.2). The agreement between the MZ maps was determined using Cohen’s Kappa statistic (Eq. 1.1) (Cohen, 1960).

$$kappa = \frac{P_o - P_e}{1 - P_e} \quad (1.1)$$

where P_o is the proportion of agreement observed (i.e. the proportion of agreement between MZs of observed and simulated data) and P_e is the proportion of a random agreement (i.e. the proportion of agreement in the case MZ are derived from observed and spatially reorganized data at random).

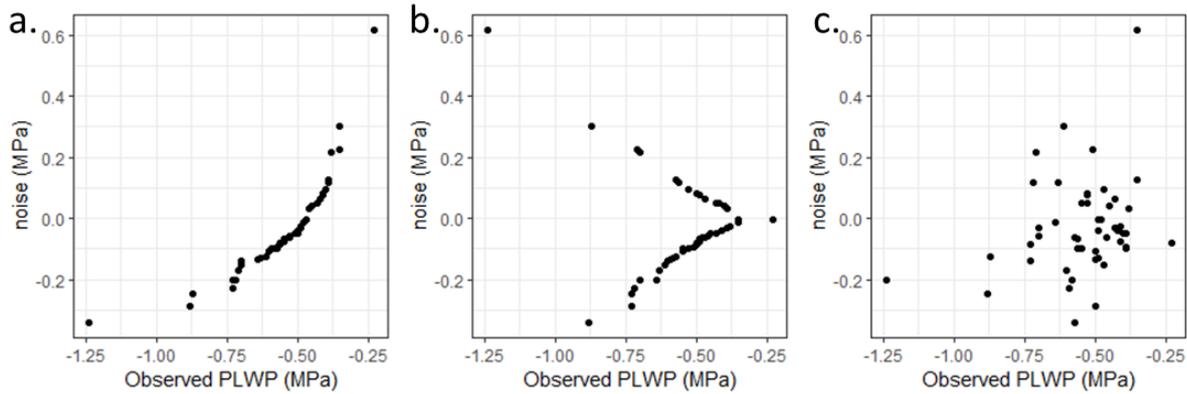


Figure 1.5 Attribution of noise depending on observed PLWP values: a Model 1 - Noise attributed is positively related to the PLWP values, b Model 2 - Noise absolute value are positively applied to the observed PLWP values, and c Model 3 - Noise randomly applied to the observed PLWP.

Table 1.2 Observed field data and three simulated Predawn Leaf Water Potential (PLWP) models classified into Management Zones (MZs) based on a 3-class classification of the observed data (and the same thresholds used for the simulated data). Metrics of model fit are shown as the RMSE between observed and simulated data ($n = 49$) and Cohen’s Kappa value associated with the similarity of the simulated MZ models to the observed MZ model.

	Observed PLWP		Simulated PLWP	
	Real Data	Model 1	Model 2	Model 3
MZ				
RMSE	-	0.16	0.16	0.16
Cohen’s Kappa	-	0.64	0.05	0.31

The RMSE was calculated from the simulated PLWP (i.e. sum of observed PLWP and attributed noise) and observed PLWP. Thus, the RMSE should identify which simulation is the best. However, because the simulated noise models are built from the same distribution (but different spatial structure) the RMSE in these cases was identical (Table 1.2). Therefore, the conclusion is that all three simulated models were equally good, and the defined MZs should be equally good. However, the resulting MZ maps for the three simulated models do not support this, nor do the Cohen's Kappa values (Table 1.2). Even though the RMSE was constant, the Model 1 spatial pattern was much closer to the original data (higher Cohen's Kappa value) than Model 2 or 3. Model 2 had the least similar spatial pattern to the observed data (lowest Cohen's Kappa value). Thus, even though the RMSE was the same on these three simulations, the derived MZs were significantly different between simulations. Selecting the best MZ (i.e. from the best spatialized model) cannot be decided with only the RMSE.

1.2.4.3. Perspectives and needs for crop model spatialization and evaluation in a precision agriculture context

Crop model spatialization is currently often used with upscaling methods. Published studies have aimed to apply traditional point crop models over larger areas, for instance, at the field scale (Acevedo-Opazo et al., 2010), at multiple field scales (Baralon et al., 2012), at the regional scale (Balkovič et al., 2013; Battude et al., 2016; Beaudoin et al., 2018; Therond et al., 2011) and at the continental scale (Adam et al., 2011; Teixeira et al., 2017). Upscaling crop models to a larger area is more common because crop models can be used by land managers and policymakers to make decisions on these large areas (Jones et al., 2017a). In contrast, there are considerably fewer studies that have aimed to use crop models at finer scales, i.e. attempting downscaling rather than upscaling. Precision agriculture is much more concerned with finer scale predictions and so downscaling approaches applied to crop models are of particular interest to the precision agriculture community.

Using crop models in a tactical management way represents a goal of precision agriculture. Nevertheless, to achieve this objective, crop models need to manage a large amount of ancillary spatial data (Chen et al., 2017). Chen et al. (2017) identify different kinds of spatial data: relatively stable data (e.g. soil type and depth), constantly changing data (e.g. LAI, soil moisture and temperature, solar radiation) and aspatial data (e.g. management activities, cultivar information). In addition to the nature of these data, the resolution of these data has to be taken into account and needs to match with the spatial footprint of the modelling (Adam et al., 2011). Spatialization, in the context of downscaling crop modelling approaches, leans heavily on using these high resolution data to define (relatively homogeneous) sub-units on which to apply a crop model (Basso et al., 2011; Cammarano et al., 2021, 2019; Guo et al., 2018).

If crop modelling using spatialized crop models is to become a common aspect of precision agriculture then new methods or statistical metrics that take into account the spatial characteristics of the data and models will be needed. Evaluation of spatialized crop models needs to be improved and evolve beyond aspatial metrics. These new statistical metrics could take into account some spatial characteristics of the data. For instance, systematically using variography to estimate the spatial structure of inputs and outputs could be a method to identify if there is an issue in input data, in the model structure or with an interaction between both of them. Using geostatistic metrics on residual crop models could be a solution to improve the spatialized crop model evaluation. At a minimum, an evaluation of the spatial autocorrelation should be performed, such as using Moran's autocorrelation coefficient (Moran, 1948) on the inputs, outputs or residuals to provide quantitative evidence of spatial autocorrelation in the data.

Furthermore, due consideration needs to be given as to how these spatial data can be best used as both inputs into the model and as data for the calibration and evaluation of the models. Spatial ancillary data are often derived data layers themselves with some level of error and uncertainty associated with them. The wide variety of gridded high-resolution digital soil property maps

(<https://esdac.jrc.ec.europa.eu/resource-type/soil-data-maps>, accessed 24/05/2021) now available are a good example of this. Soil property information is essential for many crop models, and better soil information is critical to expanding the uses of crop models. However, these soil property maps are estimates, derived from modelling approaches themselves. They are not directly measured soil properties that can be entered with confidence into the models, but the temptation is to treat these spatial ancillary data as ‘true’ data for modelling purposes. This temptation should be avoided and robust modelling approaches that explicitly take into account input uncertainty, such as Monte Carlo methods, should be routinely used in spatialized modelling applications.

Finally, the evaluation on any spatialized (or spatial) crop model will be affected by the number and spatial location of any real observations used for validation. This is true for any scale of application. However, for finer scale spatial modelling that is to be used for short-term predictive modelling to aid in-season management, the selection of correct validation sites is critical as there is limited time to resample before the crop model output needs to be used to make (spatial) management decisions. As for any modelling approach, if the validation sites do not cover the distribution of both model inputs and outputs then model evaluation will be restricted and diminished. In the case of using spatialized (or spatial) models, the spatial distribution and relevance of these validation sites must also be considered when they are selected. Related to this is the need for any validation data to respect the spatial footprint of the model outputs, either in its native form or after scaling. This in turn creates potential issues for crop model validation if the model outputs are multi-scalar in nature.

All of this comes back to the type of metric that is best suited to evaluate spatialized (and spatial) crop models. None of the metrics in Table 1 were developed for or are suitable to address these issues. How would a comparison between a well-performed model, with poorly selected spatial validation sites at the incorrect scale, and a poorly-performed model, with well-selected sites correctly sampled be properly made so that the better model was identified? Note that this question is not an issue of the quality of the analysis, but the location and the spatial footprint of the sampling. The assumption is that the analysis of the validation data is done equally well in both instances.

It is clear from the review of the literature performed here that there has not been a lot of consideration so far of spatial issues when applying crop models to precision agriculture. Despite precision agriculture being built on spatial data sets, spatial autocorrelation and its implications for statistical analysis, particularly for the assumptions behind many statistical methods, are often overlooked (Taylor and Bates, 2013). In many cases this is because precision agriculturists do not always fully comprehend the statistical implications behind spatial data (compared to ‘conventional’ agri-data sets). This is generally true across all aspects of agricultural science that are seeking to include spatial data in their domain, including the crop modelling domain. To ensure the correct use of these spatial data, agricultural scientists and modelers will continue to need the support of the statistical community, particularly the geo-statistical community, to develop new metrics to support this new area of crop modelling.

1.2.5. Conclusion

Existing crop models are point-based models and spatialization allows the use of these crop models to predict spatially across an area. Spatialization of a crop model is realized for different reasons, such as applying site-specific crop management, improving understanding of processes or to complete data sets. Most published studies have addressed spatialization from an upscaling objective to inform regional, national or global decision-making. However, in a precision agriculture context, downscaling methods need to be used for the spatialization of most crop models and only limited research has been performed in this domain so far. In addition to crop model uncertainty itself, scaling methods to spatialize models will add uncertainty to the model predictions. The present review raised questions about the current approaches to the evaluation of spatialized crop models. Current evaluation methods in published studies have used mainly spatial indicators. When spatializing crop models, spatial autocorrelation should be

considered and assessed, otherwise, crop model evaluation could be wrong. Additionally, spatialized predictive crop model evaluation will be influenced by the number, location and spatial footprint of validation data. To overcome those issues, indicators and coefficients that take spatial autocorrelation in account when evaluating the performance of the spatialized (or spatial) crop model are urgently needed and should be developed via a collaboration of the crop modelling and biometry (statistical) communities.

1.2.6. Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

1.3. Research positioning and scientific questions of the PhD project

1.3.1. Research objectives of the PhD project

1.3.1.1. Contextualization of the general research issue

As reported in the review article, crop model spatialization is currently mainly performed in upscaling studies, especially to help policymakers and land managers to make decisions (Jones et al., 2017b). When spatialization is used in upscaling studies, some issues, highlighted by the previous article, occur and crop modelers should be aware of them (e.g. scaling errors). If spatialization is used in downscaling studies, the same kinds of issues are likely to occur. However, our review showed that very few studies in the agronomic community addressed the question of downscaling processes to model agronomic variables at a within-field scale. Consequently, errors and issues with this approach have yet to be commonly reported. Even though the potential errors are still uncertain, the crop modeling community does recognize that there is a real opportunity to adapt existing crop models into decision support tools to help farmers in their precision agriculture management, i.e. at this within-field scale (Adeyemi et al., 2017; Cheng et al., 2023; Thorp et al., 2008). The availability of high spatial resolution data, as generated in the context of precision agriculture, remains recent, which certainly explains the lack of literature related to crop model downscaling. These high-resolution data will become more widespread as these data are increasingly available at low cost. Thus, their use to generate accurate agronomic knowledge or tactical decision support will necessarily become essential in the coming years. These possible new data acquisitions will highlight the growing interest of the question of downscaling crop models.

Currently used crop models are commonly used in predicting scenarios under climate change contexts or to estimate the impact of a change in cultural practices on crop production (Asseng et al., 2014). Moreover, these models are well documented, well recognized by the agronomic community and are known to work correctly if used in their correct contexts and with the correct input data. Spatialization of these models is a shift away from the target model footprint, i.e. it shifts the model usage out of its initial validity domain.

These spatialized crop models need to take into account the spatial structure of agronomic variables at the within-field scale in order to meet the challenges arising in precision agriculture. Rethinking the way of using existing crop models should permit the known advantages and recognized performances of these point-based crop models at the field scale to be transferred to other finer scales, without having to build new models. The alternative would be to design 'true' spatial crop models where the spatial modeling units interact with and impact on the modeling of neighboring spatial modeling units. This approach would ensure the reproduction of potential spatial interactions at the within-field scale but would require a fundamental change in the underlying crop model structure and equations. This could be desirable in very structured agrosystems, such as agroforestry systems for instance, but would also constitute a major endeavor in terms of time and cost for the crop modeling community and would negate a large amount of crop modeling work that has been achieved over the past 20-30 years.

For this reason, the spatialization of well understood, stable crop models is currently the preferred approach to scale changes with crop models.

A first approach to downscaling crop models would be to simply run the model independently over each spatial modeling unit as defined by the target spatial resolution. However, precision agriculture is mainly based on responding to the spatial variation of environmental and agronomic variables by differentially adapting cultural practices to increase the efficiency of inputs. In short, variability is inherent in the system and occurs at different scales for different variables, and it may not align with a fixed target spatial resolution. Additionally, when downscaling a crop model to an ever finer target spatial resolution, the influence of sources of stochastic variation in the model inputs are hypothesized to become more dominant. Crop models are designed to use ‘average’ field data so that the modeling is unaffected (or not overly affected) by sources of stochastic variation. For downscaled applications, the effect and limits of the stochastic variability in crop model inputs on the spatialized crop model outputs is still poorly understood. In addition, by operating the crop model as a point-based model individually on each spatial modeling unit, there is no consideration, nor any attempt in the modeling to reproduce the spatial structure of the agronomic variables of interest. The modeling remains aspatial. Given these two limitations, a finer resolution scale may not necessarily be the most relevant (accurate) modeling scale for within-field simulations. A trade-off between the accuracy of prediction against the ‘noise’ in the input data at different levels of downscaling is expected. In a similar vein, the resolution of available data to validate the downscaling processes and the targeted simulated scale will not always match, thus, a further trade-off needs to be found between both to ensure a consistency for simulation and evaluation at the within-field scale. Furthermore, to be able to model at within-field scales, spatialized crop models will typically need to rely on some high-resolution ancillary data to improve their performance, especially in the calibration step (e.g. using delineated of within-field zones, data assimilation or Bayesian inference approaches). These ancillary data must also be consistent with the downscaling processes if used.

These factors all feed into the linked question of concern relating to the evaluation of spatialized crop models performances, and comparing and contrasting outputs at different levels of scale. As highlighted in the review article, there is a lack of relevant metrics to be able to evaluate the performances of spatialized crop models at various scales. Given this context, it is difficult to determine if it is relevant to use existing crop models at the within-field scale and to identify the best performing simulation scale. This context has given rise to the general research issue on which this PhD project has been based:

Is the spatialization of existing crop models, by using downscaling processes, conceivable and relevant for their use at within-field scales?

The spatialization process highlighted in this general research issue has a major impact on the way to use existing crop models at finer spatial scales. Indeed, the considered spatialization method, added to the inherent stochastic character of high-resolution data, raises the question of the correct way to evaluate spatialized model performance. The performance of a spatialized crop model is not only based on the prediction accuracy of the numerical values of the considered agronomic variable but also on the ability to replicate the spatial pattern of this agronomic variable. Thus, the way these spatialized crop models are evaluated could influence the interpretation or comprehension of the model outputs if both the aspatial or spatial prediction errors are not taken into account. However, to have a well performing crop model implicitly requires having a good model calibration. Thus, the spatialization process evaluation involves issues related to model calibration as well as model prediction. The calibration issue for spatialized crop models also needs to find a trade-off between the stochastic and spatial variation of the considered agronomic variable. In particular, if model outputs need to be representative of the spatial characteristics of the agronomic variable, spatial calibration using minimal local data that is prone to stochastic error may lead to an incorrect (spatial) model calibration and induce bias in the modeling. There will be some trade-off between the level of model downscaling that can be

performed versus the aggregation of available data to spatially calibrate the model for a given level of downscaling. In this context, downscaling by creating management zones, and multiple levels of zones, could be effective for achieving a well-calibrated downscaled modeling solution. If done correctly, spatial calibration could retain the spatial characteristics of the variable and avoid the potentially strong impact of stochastic error in the modeling process.

1.3.1.2. Scientific questions and PhD project objectives

The objective of this PhD project was to elaborate preliminary work and to develop tools to be used for the spatialization of existing crop models at the within-field scale for precision agriculture purposes. A first theoretical step was considered necessary to establish the basis for performance analysis of spatialized (mechanistic) crop models used during the PhD project. The ways to evaluate such crop models at different scales needed to be interrogated and eventually rethought to consider both aspatial and spatial model performance. Then, to assess the relevance of using downscaling approaches with existing crop models, and in particular downscaling via a spatial calibration, several case studies were constructed to test various hypothesis associated with spatialized crop models. The intent was not to draw general conclusions on the spatialization of crop models, but to formalize this concept in a precision agriculture context and build a base for future research into the use of strategic crop models for tactical and spatial within-field management.

Regarding the general issue and highlighted points previously discussed, two main scientific questions were identified and detailed in several specific scientific sub-questions:

(I) How to perform a relevant evaluation and comparison of spatialized crop model performances across different spatial scales?

- (i) Are the currently used methods to evaluate spatialized crop models performances efficient?*
- (ii) Does taking into account the spatial component of the model outputs improve model evaluation?*

(II) Is the spatial calibration of selected crop model parameters an effective method of downscaling existing crop models to permit modeling at within-field scales?

- (i) Can the delineation of potential management units from high-resolution ancillary data assist with a spatial calibration approach? How does the relationship between the ancillary data and the agronomic variable of interest influence this?*
- (ii) Is there a trade-off between the spatial scale of modeling at within-field scale and the noise brought by the errors of measurement of the data used for the calibration and evaluation step of the spatialized crop models?*
- (iii) Does the complexity of a crop model affect the ability of the model to be downscaled for precision agriculture purposes?*

1.3.2. Research approach implemented for the PhD project

Spatialization, in this PhD project, is defined as applying a crop model at a spatial scale that is different to its native and target scale. The crop models are thus considered in this project as ‘black box’ tools. In particular, downscaling spatialization processes are identified for this project as an opportunity to use existing crop models that have been designed for ‘average’ field scale use at other finer (within-field) simulation scales without a modification of the internal structure of the crop model.

The PhD project was seen as a first approach for the study of the crop model spatialization via a spatial calibration downscaling process, rather than the much more common data assimilation process. Indeed, spatial calibration involves a spatial constraint in the calibration process compared to the data

assimilation. Data assimilation is based on the incorporation of data (often remote-sensed data or data from unmanned aerial vehicles) during the modeling phase, allowing a better consistency between observed and modeled data, leading to increased crop model performances. This later allowing a temporal adjustment of the crop model state variables during the modeling phase. Thus, certain parts of the PhD needed to be based on theoretical aspects. For these studies, the generation of simulated data (virtual fields) was necessary to have a greater representation of possible real cases and to control the properties, particularly the spatial properties, of agronomic variables in these virtual fields. During the PhD project, there was a desire to also carry out real case studies to exemplify the potential interests advanced in the theoretical aspects of the PhD project. Thus, whenever it was possible, research was based or informed on real data. However, data sets that are suitable for studies related to the spatialization of crop models at the within-field scale need to be collected at a resolution that is fine enough to have a proper representation of the within-field spatial variations. Unsurprisingly, these data sets are scarce as they are difficult, time-consuming and expensive to acquire. The collection of such data are not possible within the time-constraints of a PhD project, especially during the COVID health crisis that complicated field activities. Therefore, the decision was made to use two historical data sets that were already available.

The first available data set is related to spatio-temporal observations of vine water status within a vineyard. This data set was initially acquired for a previous PhD project realized at INRAE (Acevedo-Opazo, 2003). Predawn leaf water potential (Ψ_{PD}) was measured on a 1.2 ha Syrah (*Vitis vinifera* cv. Syrah) vineyard at Pech Rouge (43.144°N, 3.131°E, Gruissan, Aude, France), the experimental site of INRAE. The Ψ_{PD} was measured during two successive growing seasons in 2003 and 2004 and respectively on 7 and 6 dates (between June and September) in those years. In each year, there were 49 measurement sites within the vineyard block. Spatial ancillary data were also available for the vineyard related to the soil (soil apparent electrical resistivity), vine physiology (trunk circumference) and vigor (via the normalized difference vegetation index). This data set was used for Article 2 in Chapter 2 and Article 4 in Chapter 3.

The second available data set related to a durum wheat (*Triticum durum* Desf. cv. Gargano) production system. This research was funded by the Integrated system for development of southern cereal farming (SI.Cer.Me) a program for southern Italy development (CIPE 17/2003 -1.1 and 83/2003). Durum wheat yield was measured on a 12 ha experimental field of CREA (Research Centre for Cereal and Industrial Crops) near Foggia (41.462°N, 15.506°E, Italy). Yield was measured during three consecutive growing seasons from 2005-2006 to 2007-2008. In each year, there were 100 measurement sites within the field where phenology (Zadock stages) was measured throughout the growing season and variables related to soil properties were also measured (e.g. crop lower limit, drained upper limit, organic content, ...). Unfortunately, production in the middle year, 2006-2007, was severely affected by unusual weather patterns. This resulted in very low and atypical production data, which was outside the reasonable boundary expectations of the crop model. For this reason, the data from this year were omitted and only the data from 2005-2006 and 2007-2008 were retained for this study. These data were used for Article 5 in Chapter 4.

Studies realized during the PhD project were developed from two different crop models. Both of them were chosen according to two conditions. Firstly, the crop model needed to be able to model the target agronomic variables using the available (real and, where needed, simulated) data sets (i.e. vine Ψ_{PD} and durum wheat yield). Secondly, in order to assess more possible cases, the complexity level of various crop models was considered, with a desire to have an example of a simple crop model and a complex crop model (with complexity defined by the number of internal equations and the number of necessary inputs). Regarding these conditions, WaLIS (Water baLance for Intercropped System) (Celette et al., 2010) and APSIM (Agricultural Production Systems sIMulator) (Holzworth et al., 2014) were selected. WaLIS is considered a simple crop model to model Ψ_{PD} and APSIM is a relatively more complex crop model to model durum wheat yield.

1.3.3. Manuscript sectioning

This first chapter (Chapter 1) has described the general context (Article 1) and the specific scientific questions addressed in this PhD project. The following research component of the thesis document is divided into two main parts. The first part addresses theoretical concepts related to the evaluation of spatialized crop models and proposes a new metric to achieve this (Chapter 2). The second part (Chapters 3 and 4) is more practical and focuses on the relevance of downscaling at various within-field scales with a spatial calibration process when using existing simple and complex crop models.

The theoretical part deals with the question of a relevant evaluation of spatialized crop models (Chapter 2). In the first section (2.1), a non-peer reviewed conference article presents a study on the use of classical aspatial and spatial metrics to examine if it is reliable and relevant to use such metrics to assess the performance of a spatialized (simple) crop model related to vine water status (Article 2). In the second section (2.2), a peer-reviewed conference article introduces new metrics designed to evaluate the performance of a spatialized crop model at different spatial scales based on the assessment of aspatial and spatial error patterns (Article 3).

The practical part of the thesis is divided into two chapters. The first chapter (Chapter 3) is devoted to the use of spatial calibration of selected model parameters, with a relatively simple crop model, as a spatialization process for downscaling the crop model. The intent is to investigate and better understand the relevance of using a spatial calibration approach with crop models. Chapter 3 consists of an article that uses a mix of simulated and measured data to identify when spatial calibration is relevant. A relatively simple crop model for vine water stress in vineyards is used for this study (Article 4). The Chapter 4 addresses the spatialization and the use of a complex crop model at different within-field scales. Using the spatial calibration approach developed in Chapter 3 and the metrics defined in Chapter 2, a complex crop model is downscaled and evaluated using real spatial environmental and crop data that are related to spatially variable durum wheat yield in southern Italy over two years (Article 5).

The last chapter is dedicated to a discussion and perspectives of the PhD project in the field of crop modeling and the future use of spatialized crop models at within-field scales for precision agriculture purposes (Chapter 5).

To summarize, the manuscript is divided into five chapters. Chapters 1, 2, 3 and 4 are made of several scientific articles or conference communications. The articles in Chapter 1 (review journal article; Q1 journal) and Chapter 2 (two full conference papers, one of which is peer-reviewed) are published. The articles in Chapters 3 and 4 are currently under review in Q1 journals. Each chapter starts with an intention note and finishes with a general conclusion to place the articles and sections within the scientific context of the PhD project.

Chapter 2

Achieving a relevant assessment of spatialized crop model performance when simulating agronomic variables at changing scales

2.1. Intention note

Evaluation of a crop model is paramount to knowing its reliability to reproduce accurately the reality of a considered agronomic variable. In the literature, several metrics have been used to evaluate crop model performances. Therefore, the agronomic and modeler communities, who classically and routinely used these metrics, have emphasized the importance of selecting the right metric. Many of these metrics are based on evaluation of numerical errors, i.e. based on the numerical difference between observed and modeled data (Bennett et al., 2013; Wallach et al., 2019). The term ‘aspatial metrics’ is used in this manuscript to refer to such metrics. Other metrics used have focused on the characterization of the spatial structure and pattern of the variables (Leroux and Tisseyre, 2018; Luo et al., 2019), and the term ‘spatial metrics’ is used for such metrics. However, these spatial metrics have not necessarily been used for the evaluation of (crop) models but rather, have been identified as having the potential to fulfill the role of an evaluation metric. In particular, these metrics could assess the similarity of the spatial structures and patterns between observed and modeled data. However, the objective of a spatialized crop model should be to model both the numerical value and the spatial structure/pattern of the target agronomic variable. In this sense, the spatial characteristics of the data should be more important when evaluating them compared to the evaluation of existing point-based crop models. Thus, a first working assumption for the following work is that a relevant and reliable evaluation of a spatialized crop model should account for the both aspatial and spatial error in the modeling.

However, based on the proof realized in the simulated example in Section 1.2.4.2.3, a second working assumption is that currently used evaluation metrics, both aspatial and spatial ones, are not necessarily the most relevant for evaluating spatialized crop model performances. Thus, a new metric accounting for both aspatial and spatial error is needed. Such a metric could also be used to evaluate the spatialization process, i.e. the effect of changing the modeling spatial scale. In this chapter, such a metric will be proposed for application to spatialized crop models but, more generally, this metric could equally be applicable to any type of spatialized models (e.g. hydrological or environmental models) and even for ‘true’ spatial models (that account for neighborhood interactions within the model). The method used to compute and interpret the proposed metric will be relevant for the outputs of all of these types of models.

Before the work realized for this chapter, especially presented in the Section 2.3, little attention had been paid to the use of the terms ‘indicator’ and ‘metric’, and they were used interchangeably. However, as ‘metric’ is more frequently used in the literature, ‘metric’ has been preferred and will be used in the rest of this manuscript (but ‘indicator’ remains interchangeable with ‘metric’ in the previous sections). Spatial structure and spatial pattern also have different meanings, especially in Section 2.3, and these are not interchangeable terms. They will be clearly defined in Section 2.3.

The following chapter is divided into two different research papers. Based on the second working assumption of this chapter, the first paper expands on the initial proof in Section 1.2.4.2.3 to more clearly demonstrate that classically aspatial and spatial metrics are not necessarily the most relevant to evaluate spatialized crop model performances. This work is based on real data from an experimental vineyard in an operational context for precision viticulture management. Based on the results of this preliminary work and leading on from the first working assumption of this chapter, a new metric, spatial balanced accuracy (SBA) is proposed in the second paper that is able to account simultaneously for the aspatial and spatial pattern error of spatialized crop models (and more globally spatial models). The performances and relevance of this metric will be tested on simulated data.

2.2. Comparison of different aspatial and spatial indicators to assess performance of spatialized crop models at different within-field scales

2.2.1. Details about the paper

2.2.1.1. Title and publication information

This section was presented as an oral communication at the 15th International Conference on Precision Agriculture (2022, Minneapolis, USA) and the full paper was published without review in the conference proceedings. The full citation is:

Pasquel, D., Roux, S., Tisseyre, B. and Taylor J. A. (2022). Comparison of different aspatial and spatial indicators to assess performance of spatialized crop models at different within-field scales. In: Proceedings of the 15th International Conference on Precision Agriculture, Minneapolis, Minnesota, United-States, June 26-29.

2.2.1.2. Authors

Pasquel D.^{1,*}, Roux S.², Tisseyre B.¹ and Taylor J.A.¹

¹ ITAP, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

² MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

* Corresponding author: daniel.pasquel@inrae.fr

2.2.1.3. Abstract

Most current crop models are point-based models, i.e. they simulate agronomic variables at the spatial footprint on which they were initially designed (e.g. plant, field, region scale). Spatialization (i.e. using point-based crop models on a different scale than its native spatial footprint) represents a solution to use these crop models on a different scale. This is particularly interesting in a precision agriculture context where downscaling processes are involved to model agronomic variables on finer scale (e.g. within-field scale). To assess their performances, many indicators based on the comparison of estimated vs observed data, can be used. However, the use of classical, aspatial indicators may not be relevant to evaluate spatialized crop model performances. The objective of this work was to compare how different model performance indicators are able to evaluate the performance of a spatialized crop model at various within-field scales. The crop model spatialization processes were based on a spatial calibration of model parameters. This work focused on a case study using the crop model WaLIS (Water baLance for Intercropped Systems) to simulate vine water restriction (estimated through the predawn leaf water potential - Ψ_{PD}) for a vineyard in the South of France. The WaLIS model was employed at different spatial scales (field, site, within-field zone) to generate Ψ_{PD} maps. The management zones were generated from soil and vine ancillary data that are correlated with or directly influence vine water stress. Aspatial (RRMSE and D-index) and spatial (Cambardella index and Z-score) indicators were used to

evaluate model performances at these different spatial scales. Results showed that these different indicators generated different ‘best’ simulation scales and there was no clear result of model performance from the spatial and aspatial indicators. This confirmed that current approaches to crop model evaluation were not well suited to evaluation of the performance of spatialized crop models in a precision agricultural context. Evaluation in an operational context through decision-making evaluation and map comparison approaches provided a clearer understanding of model behavior and appeared to be a relevant method for evaluating downscaled spatialized crop model predictions for tactical, in-season and differential crop management.

2.2.1.4. Keywords

Spatialization, Spatial calibration, Downscaling, Evaluation indicators, Vine water restriction

2.2.2. Introduction

Most current crop models are point-based models, i.e. they simulate agronomic variables at the spatial footprint on which they were initially designed (e.g. plant, field, region scale) (Heuvelink et al., 2010). However, shifting model use from a strategic objective to tactical in-season management is becoming a significant issue for the agronomic community, especially in a precision agriculture context. Spatialization (i.e. using point-based crop models on a different scale than its native spatial footprint) represents a solution to address crop model use in a tactical and operational context. Calibration is an inevitable process to improve crop model performances (Seidel et al., 2018). This is especially true when using crop models over large areas (Jagtap and Jones, 2002), thus calibration is a critical step for models used in a spatialization context. In the case of precision agriculture applications that involve downscaling processes, it is important to consider how to calibrate crop models at the within-field scale. In this study, spatialization is based on a spatial calibration led at different spatial scales.

To assess crop model performance, many indicators based on the comparison of observed and modeled data, can be used, such as the relative root mean square error (RRMSE) or Willmott index of agreement (D-index) among others (Bennett et al., 2013; Wallach et al., 2019). However, the use of these indicators to evaluate spatialized crop models raises questions, and these indicators may not be relevant to evaluate their performances (Pasquel et al., 2022a). Evaluation on the raw error of the crop model (error between observed and simulated data) using these indicators may not be sufficient to assess the spatial efficacy of calibration and/or prediction. Preliminary results using simulated data have indicated that the indicators currently used to evaluate spatialized crop model performance are not the most relevant (Pasquel et al., 2022a). However, an evaluation and comparison of spatial and aspatial indicators on a real-world case study in a precision agricultural context has not yet been done, nor has there been any investigation of the effect of scale change (increasing levels of downscaling) on model performance. Note that if a spatialized crop model is used in an operational context associated with site-specific decision-making, it should be possible to mathematically define an error on the decision made using the simulations. Such indicators would clearly be the most adapted to assess model performances with respect to the targeted model use, but have not been strongly advocated to date.

Therefore, the objective of this work is to compare the evaluation of spatialized crop model performances using different indicators (aspatial: e.g. RRMSE or D-index; and spatial: e.g. Cambardella index or Moran index) for different simulation scales. The study is done on a vine water status crop model, WaLIS (Celette et al., 2010). As well as the different model statistics, the error on the decision taken is also used to evaluate the spatialized crop model performance. Model performance is defined by relationships between the observed and modeled data and the preservation of the spatial structure of the modeled variable in relation to the observed variable. It should be noted that the preservation of the spatial structure of the model output(s) is important in a precision agriculture objective.

2.2.3. Material and methods

2.2.3.1. Field description and observed predawn leaf water potential ($\Psi_{PD,obs}$)

The predawn leaf water potential (Ψ_{PD}) of the vine was considered as the reference data in this case study. Ψ_{PD} measurements were carried out on a 1.2 ha Syrah vineyard block at INRAE's Pech Rouge estate (Gruissan, Aude, France) on 49 within-field sites (Figure 2.1A) using a pressure chamber, these measurements were the observed data ($\Psi_{PD,obs}$). The $\Psi_{PD,obs}$ measurements were done for 7 dates in 2003 ($\Psi_{PD,obs,n-1}$) and 6 dates in 2004 ($\Psi_{PD,obs,n}$) (see Acevedo-Opazo et al. (2010) for full details on this data set).

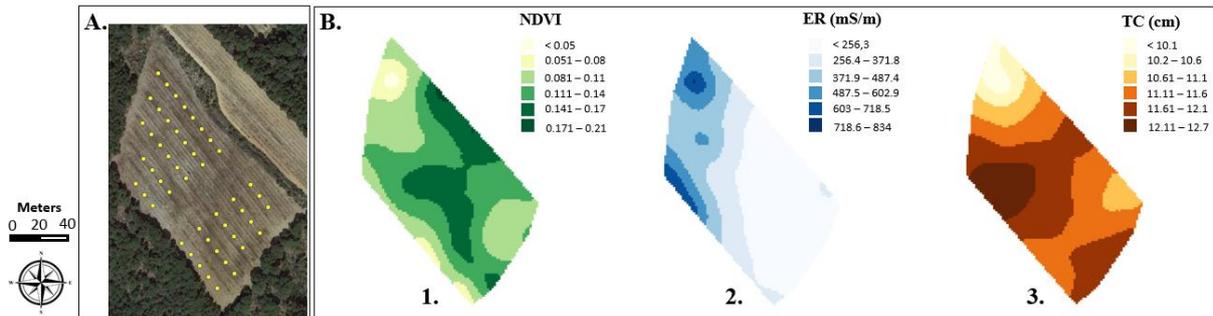


Figure 2.1 Experimental field with A. Locations of observed predawn leaf water potential ($\Psi_{PD,obs}$) within-field measurements in a 1.2 ha Syrah vineyard at INRAE Pech Rouge (Gruissan, Aude, France). B. Interpolated (kriged) maps of ancillary data used to define within-field zones: normalized difference vegetation index (NDVI) (a.), soil apparent electrical resistivity (ER, expressed as apparent electrical conductivity (EC_a)) (b.) and trunk circumference (TC) (c.) for the Syrah vineyard.

Three ancillary data were considered: soil apparent electrical resistivity (ER), trunk circumference (TC) and normalized difference vegetation index (NDVI). The ER and TC were both measured at the $\Psi_{PD,obs}$ measurement sites (Figure 2.1A) in March 2006. The NDVI values were derived from an airborne multispectral image obtained in August 2006. Note that these ancillary data were used for their spatial pattern and not for their absolute values. The ancillary data were not measured in the same year as the $\Psi_{PD,obs}$; however, for a perennial crop like grapevine, it has been shown that NDVI and ER spatial patterns in this vineyard are temporally stable over short-time periods (3-5 years) (Kazmierski et al., 2011; Tisseyre et al., 2008). Thus, these ancillary data are assumed to present the same spatial pattern even in a different year. Ancillary data were interpolated by ordinary kriging using GeoFIS (Leroux et al., 2018) (Figure 2.1B).

2.2.3.2. WaLIS and modeled predawn leaf water potential ($\Psi_{PD,mod}$)

The modeled predawn leaf water potential ($\Psi_{PD,mod}$) at multiple dates were simulated using a predictive model of vine water stress: Water baLance for Intercropped Systems – WaLIS (Celette et al., 2010). To run WaLIS, weather data for the years 2003 and 2004 were acquired through the weather station 11170004 (Gruissan) of the INRAE network via the Climatik application (Figure 2.2). Measurements realized in 2003 were used to calibrate WaLIS and measurements realized in 2004 were used to evaluate modeling performances. Note that weather conditions for both years were relatively close. Daily mean temperature (T_{mean}) and daily precipitation (P) were recorded and daily evapo-transpiration (ET) was computed using the Penman-Monteith equation (Allen et al., 1989; Pereira et al., 1999). In reality, WaLIS simulates the fraction of transpirable soil water (FTSW) and, by using a conversion, FTSW was transformed into $\Psi_{PD,mod}$ (Eq. 2.1) (Lebon et al., 2003). However, this conversion contains a logarithmic expression and FTSW can be equal to 0, so a realistic $\Psi_{PD,mod}$ minimum had to be defined. The same $\Psi_{PD,obs}$ minimum for the field and year was used, -1.1 MPa.

$$\Psi_{PD,mod} = \frac{\log(FTSW) - \log(C_a)}{C_b} \quad (2.1)$$

with C_a is a constant equal to 1.0572 and C_b is a constant equal to 5.3452.

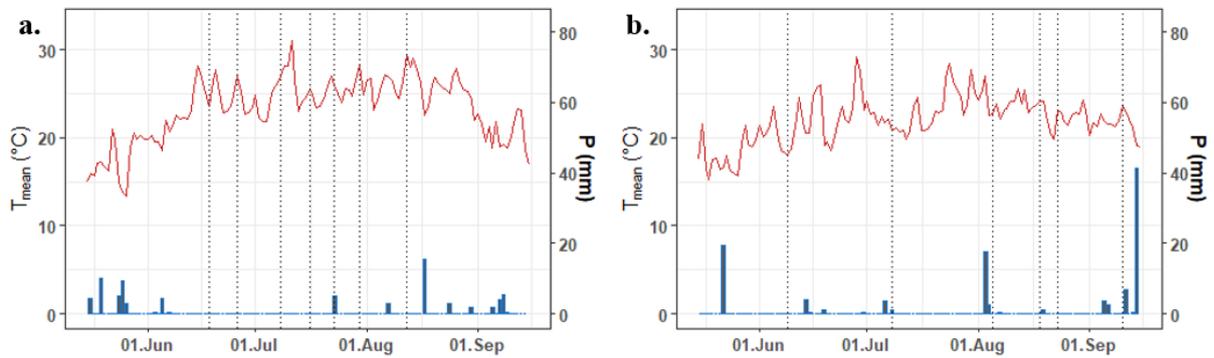


Figure 2.2 Weather conditions for 2003 (a.) and 2004 (b.) for the period of measurement dates used for respectively model calibration (2003) and performance evaluation (2004). The red line corresponds to daily mean temperature (T_{mean}) and the blue columns correspond to daily precipitation (P) events. Dates of measurement of predawn leaf water potential for calibration ($\Psi_{PD,obs,n-1}$) in 2003 were June 18th, June 26th, July 8th, July 16th, July 23rd, July 30th and August 12th. Dates of measurement of predawn leaf water potential for evaluation ($\Psi_{PD,obs,n}$) in 2004 were June 9th, July 8th, August 5th, August 18th, August 23rd and September 10th.

2.2.3.3. Spatial calibration of WaLIS

The intent of this study is to run the WaLIS model at different scales (at the measurement site-scale and at different within-field zone scales) than its native spatial footprint (i.e. field scale), thus WaLIS will be used in a spatialization process (Pasquel et al., 2022a). The ER, TC and NDVI data were used for the realization of within-field zones via a segmentation algorithm (Pedroso et al., 2010) included in the GeoFIS software (Leroux et al., 2018) with the aim to define a base grid for the analyses. All three types of ancillary data were considered as potential surrogate to explain Ψ_{PD} spatial variability. The field was divided into 2 to 5 zones by the segmentation algorithm with these ancillary data (Figure 2.3). Therefore, the spatial scales considered in this analysis are the individual measurement sites ($n = 49$) (Figure 2.1A), the different zoning levels ($z \in [2;5]$) (Figure 2.3) and the whole field (single value). The Ψ_{PD} is modeled by WaLIS ($\Psi_{PD,mod,n}$) on each of these considered spatial scales.

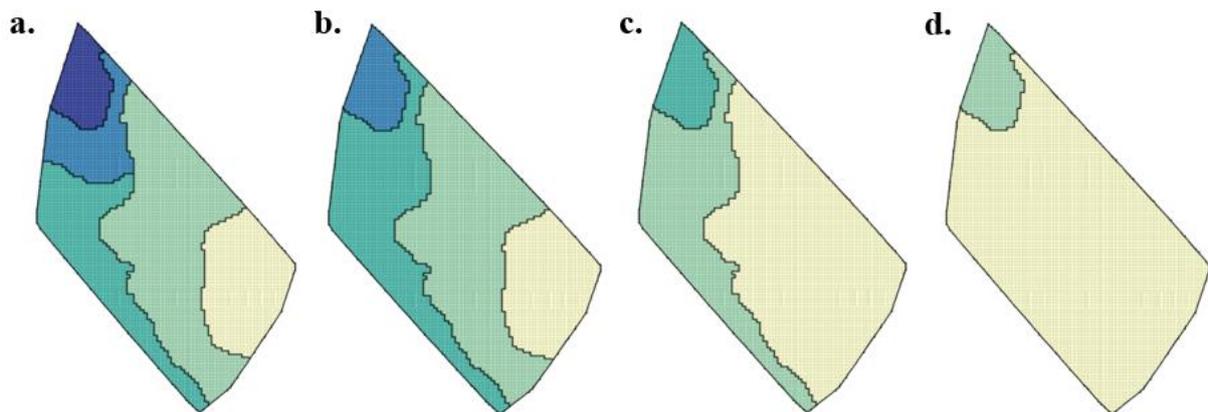


Figure 2.3 Maps of different within-field zones defined with ancillary data using a segmentation algorithm; 5 zones (a.), 4 zones (b.), 3 zones (c.) and 2 zones (d.) for the Syrah vineyard. Note that the 1 zone solution is equivalent to the whole field scale.

Defining sub-units (zones) using ancillary data to apply crop models in order to spatialize the model is a process already seen in the literature (Basso et al., 2011; Cammarano et al., 2019; Guo et al., 2018). Spatial calibration for spatialization of the crop model is driven by the zoning process based on

ancillary data, a process that is seldom seen in the literature. In an operational context here, a minimal size zone has been defined at 500 m², i.e. this is the minimum surface for which it is possible to set up a specific management action. With this constraint, 5 zones was the maximum number of zones that the segmentation algorithm was able to make (Figure 2.3). Beyond 5 zones, the segmentation algorithm was unable to find a solution with the minimum surface constraint.

To enable spatialization, within each delineated zone at each spatial scale, an aggregation of the $\Psi_{PD,obs}$ from the measurement sites was performed to generate mean $\Psi_{PD,obs}$ values at larger spatial scales (zoning levels for $z \in [2;5]$ and the whole field scale). The local WaLIS calibration was performed at either the original observation scale (for site-scale modeling) or after these aggregations (for modeling at scales larger than the site-scale), i.e. this is a kind of spatial calibration (Figure 2.4). The calibrated parameters in the WaLIS model were the total transpirable soil water (TTSW) and the maximum crop coefficient of the vine (K_c). Other WaLIS parameters were kept at their default value for a Mediterranean context. Thus, only TTSW and K_c have been calibrated differently at the different scales because these parameters are known to be variable from one vineyard to another even at within-block scales (McClymont et al., 2019; Verdugo-Vásquez et al., 2022). The TTSW and K_c were calibrated with the data from the previous year ($\Psi_{PD,obs,n-1}$), and an optimization was performed with TTSW values ranging from 55 to 210 mm and K_c values ranging from 0.35 to 0.5 to identify optimal parameter values. The retained parameter values were those that minimized the mean absolute error (MAE) when compared to $\Psi_{PD,obs,n-1}$. For all the analyses, the output scale is disaggregated to the site-scale to assess the model performance (Figure 2.4).

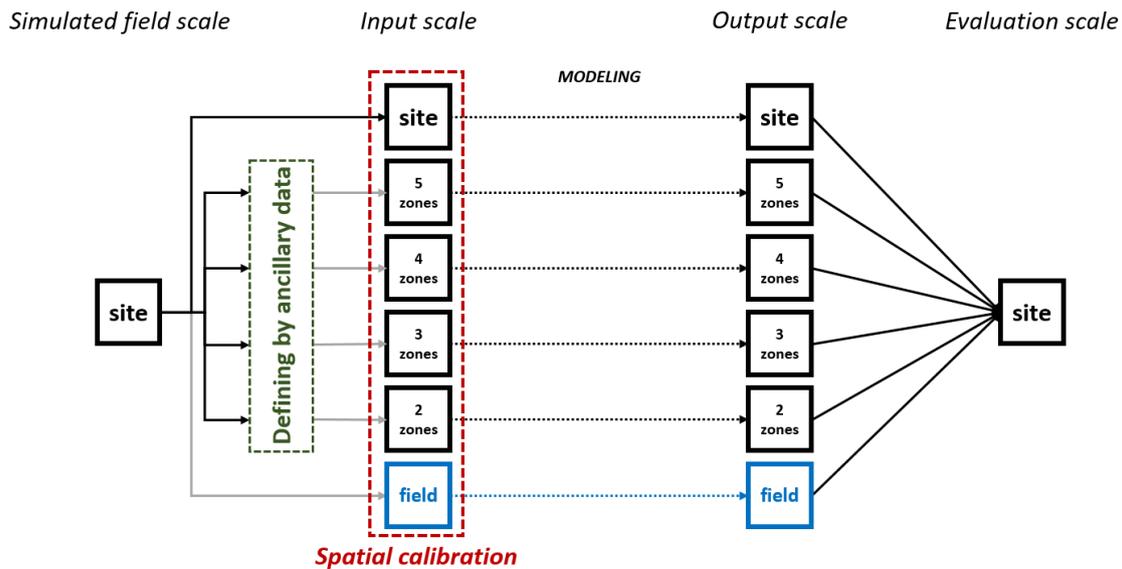


Figure 2.4 Different modeling cases of predawn leaf water potential (Ψ_{PD}) by WaLIS according to different spatial scales defined at the observation scale, the whole field scale or intermediate zonal scales (2 to 5 zones) based on ancillary data. The native WaLIS spatial footprint is shown in blue and corresponds to the field scale. Site-scale corresponds to the original observation scale. The grey arrows correspond to the upscaling process associated with aggregations of the observed data to a higher spatial scale as model input. The spatial calibration is performed at this input scale.

2.2.3.4. Simulation performance evaluation using aspatial and spatial indicators

In order to estimate the WaLIS performance for the different simulation scales a set of aspatial indicators: relative root mean square error (RRMSE) (Eq. 2.2) and the Willmott index of agreement (D-index) (Eq. 2.3); and spatial indicators: Cambardella index (C_i) (Eq. 2.4) and Z-score (Eq. 2.6) were computed. RRMSE and D-index are assessment indicators of the fit between observed and simulated data. If the D-index is equal to 1, there is a perfect match between observed and simulated data, if D-index is equal to 0, there is no match at all. The C_i and Z-score are indicators of the spatial structure of the data. These indicators were calculated on both the observed and modeled data and on the residuals

between the observed and modeled data. The C_i is derived from variographic analysis of $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$, with values of $C_i < 0.25$ corresponding to a strong spatial structure, values between 0.25 and 0.75 corresponding to a moderate spatial structure and values > 0.75 corresponding to a weak spatial structure (Cambardella et al., 1994). Global Moran's index test (I) (Eq. 2.5) was also used to estimate the spatial autocorrelation (Moran, 1948) and the values of I were then transformed into a Z -score. The Z -scores were interpreted to assess the magnitude and significance ($\alpha = 0.05$) of the spatial autocorrelation of Ψ_{PD} and residuals, knowing that $-1.96 < Z\text{-score} < 1.96$ corresponds to a non-significant spatial autocorrelation, i.e. there is no clearly identifiable spatial structure.

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - S_i)^2}}{\bar{O}} \quad (2.2)$$

$$D\text{-index} = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (2.3)$$

where O_i is the observed value, S_i is the corresponding simulated value, n is the number of observations ($n = 49$) and \bar{O} is the average of observed values.

$$C_i = \frac{C_0}{C_0 + C_1} \cdot 100 \quad (2.4)$$

where C_0 is the variogram nugget and C_1 is the variogram partial sill.

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.5)$$

$$Z\text{-score} = \frac{I - E(I)}{\sqrt{V(I)}} \quad (2.6)$$

where y_i and y_j are the variable of interest at different spatial locations (i and j), \bar{y} is the mean of the variable of interest, w_{ij} is a matrix of spatial weights quantifying the influence of j on i , n is the number of units indexed by i and j , $E(I)$ is the average of I and $V(I)$ is the variance of I .

2.2.3.5. Simulation performance evaluation from an operational context

The model outputs and the results of the indicators used to evaluate the performance of the modeling were also interpreted through the decision process to irrigate (or not) the vineyard at the within-field scale. The decision to irrigate was based on different Ψ_{PD} thresholds depending on the vine phenological cycle (Ojeda, 2007). The Ψ_{PD} thresholds were determined from the vine phenological stage assigned to the approximate corresponding dates (Figure 2.5). Thus, the intent was not to compare the Ψ_{PD} values (modeled vs. observed) but to compare the decision made (to irrigate or not). The balanced accuracy statistic (BA) was used to summarize if the set of decisions taken at the site-scale to irrigate with $\Psi_{PD,mod}$ corresponded to the set of decisions taken based on $\Psi_{PD,obs}$ (Eq. 2.7). When BA is equal to 1 there is a perfect classification.

$$BA = \frac{\text{Sensitivity} + \text{Specificity}}{2} = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (2.7)$$

where TP is true positive, TN is true negative, FN is false negative and FP is false positive.

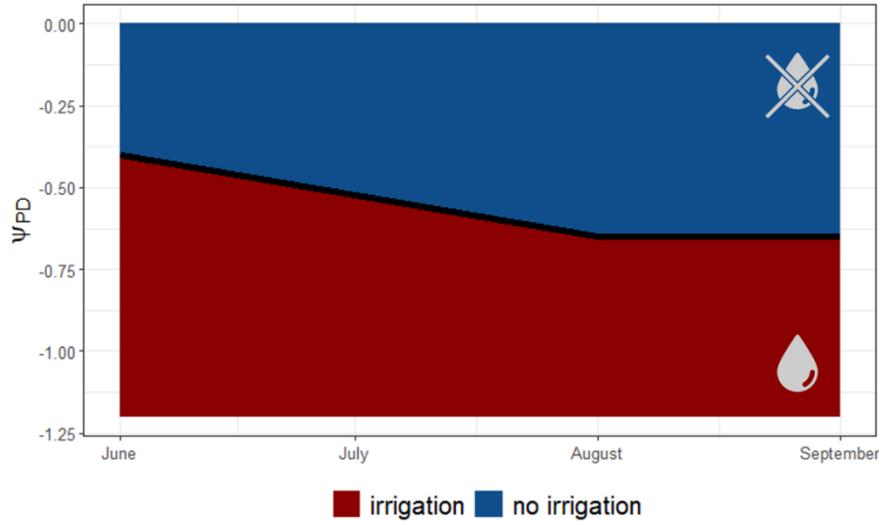


Figure 2.5 Ψ_{PD} values below which irrigation is required over the season. The equation between June and August is $\Psi_{PD} = -0.0033 \cdot \text{Day number} + 0.1$ and is constant after August 1st at -0.65 MPa (adapted from Ojeda, 2007).

2.2.4. Results and Discussion

2.2.4.1. Calibration values at varying spatial scales

Table 2.1 presents the results of the calibration optimization performed for the TTSW and K_c model parameters at different spatial scales. They represent the calibration of the model to the mean observation within each zone for the different levels of zoning ($z \in [2;5]$). The field-scale equated to a 1 zone scenario. The site-scale is not shown because it is too large to show here. However, for the site-scale calibration, the K_c values ranged from 0.368 to 0.5 and TTSW values ranged from 55 to 107.7 mm. The zone in the northern tip of the field (Figure 2.3) was constant to all zonal modeling approaches and returned constant parameter values for all calibrations. However, it was interesting to notice that for each zone resulting from the merging of two other zones (Figure 2.3), the calibrated parameter value was not necessarily an intermediate value of the parameter values of these two zones (Table 2.1). Notice that different zones within a spatial scale could have the same parameter values.

Table 2.1 Calibration values for each scale simulation (except site-scale) and for each WaLIS calibrated parameters: maximum crop coefficient of the vine (K_c) and total transpirable soil water (TTSW).

5 zones			4 zones			3 zones			2 zones			Field	
N°	K_c	TTSW	K_c	TTSW									
1	0.494	67.4	1	0.494	67.4	1	0.494	67.4	1	0.494	67.4		
2	0.458	67.4	2	0.5	73.6	2	0.5	73.6					
3	0.5	86							2	0.485	73.6	0.467	70.5
4	0.491	89.1	3	0.491	89.1	3	0.5	89.1					
5	0.5	86	4	0.5	86								

2.2.4.2. Spatial structure preservation between $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$

The $\Psi_{PD,obs}$ (in 2004) was moderately spatially structured over time in the Syrah vineyard (Figure 2.6a), with the exception being the August 8th observation that exhibited no spatial structure. This measurement occurred shortly after a large (17.5 mm) and unusual precipitation event that will have had a short-term effect of homogenizing the vine water status within the vineyard. Two weeks after the event, the spatial structure in these data had returned. Therefore, the spatial structure was present when the vines were differentially stressed due to the different soil types in the vineyard that permit access to more or less soil water. The spatial structure of $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$ realized at site-scale were comparable until early August, after which, the $\Psi_{PD,mod}$ no longer showed any spatial structure while the $\Psi_{PD,obs}$ continued to

exhibit a moderate spatial structure. The WaLIS model tended to simulate more negative $\Psi_{PD,mod}$ than the $\Psi_{PD,obs}$, which tended to homogenize the $\Psi_{PD,mod}$ and, consequently, it did not preserve the spatial structure. The C_i and Z-score interpretations were complementary and presented the same trends, with the Z-score allowing for an assessment of the significance of the spatial autocorrelation. The spatial structure estimated by C_i should be discussed with the significance of the spatial autocorrelation estimated by I (through the Z-score), as the two principles cannot exist without each other (Tiefeldorf, 2000).

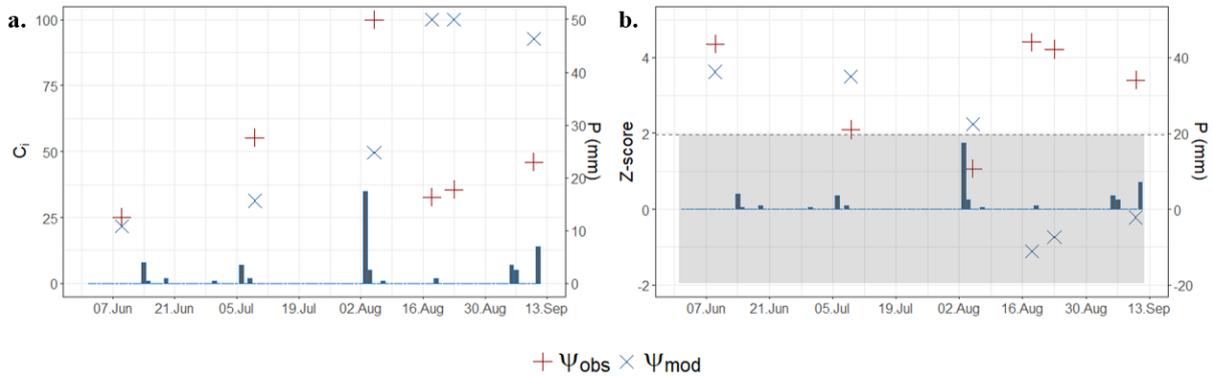


Figure 2.6 Comparison between spatial structures of observed predawn leaf water potential ($\Psi_{PD,obs}$) and modeled predawn leaf water potential ($\Psi_{PD,mod}$) using the Cambardella index (C_i) (a.) and Z-score (b.). The grey area on the Z-scores plot refers to an area of non-significance of spatial autocorrelation. Blue columns correspond to daily precipitation (P) events.

2.2.4.3. Which modeling scale is the most relevant regarding aspatial and spatial indicators?

Depending on the modeling date (seasonal timing), the model performance was not constant. At the beginning of the season, the RRMSE values were relatively close regardless of the modeling scale considered (Figure 2.7a). For late summer, when water stress was the highest, the RRMSE values increased because WaLIS underestimated the vine water restriction (overestimated Ψ_{PD}). In late summer, the RRMSE values were still the highest for the 2-zone and field scale simulations with 34% and 71% respectively. For the $\Psi_{PD,mod}$ generated at finer scales (3+ zones including individual sites), lower and similar RRMSE values were observed, showing a better agreement between $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$. When using the D-index for assessing the agreement between $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$, the finer the modeling scale at the beginning of summer, the better the agreement. However, this trend had disappeared by late summer (Figure 2.7b) with the site-scale modeling dropping from the highest D-index to one of the lowest values in mid-August. The D-index for the field scale simulations were always equal to 0, i.e. there was no agreement between $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$.

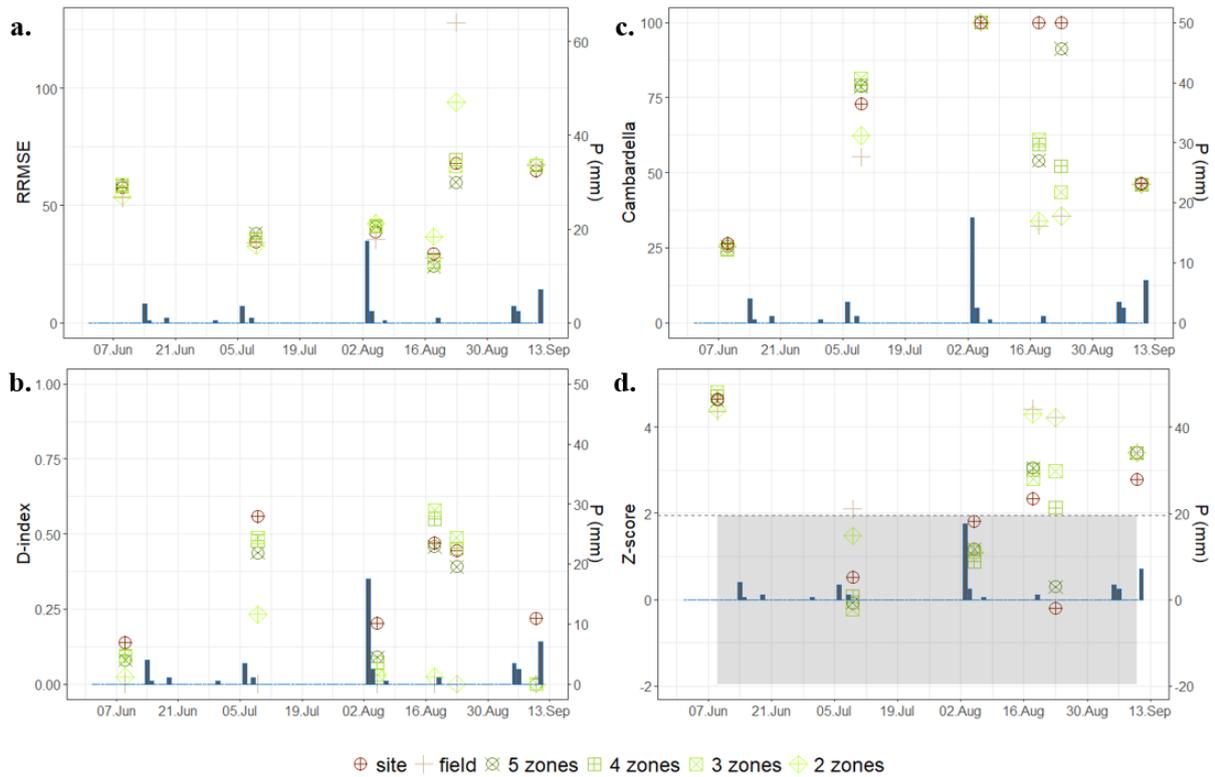


Figure 2.7 Evaluation of the WaLIS simulations of the predawn leaf water potential (Ψ_{PD}) according to the different spatial scales along the measurement period. The RRMSE (a.) and D-index (b.) are calculated between observed and modeled data. The Cambardella index (C_i) (c.) and Z-scores (d.) are calculated on the residuals between observed and modeled data. The grey area on the Z-scores plot refers to an area of non-significance of spatial autocorrelation. Blue columns correspond to daily precipitation (P) events.

The spatial structure of the residuals is an indication of how well the model has been correctly spatialized. The target variable, $\Psi_{PD,obs}$, is spatially structured during some periods of the season (especially during dry periods) (Figure 2.6a). If the modeling is able to replicate this, then the residuals should be spatially random (i.e. exhibit only a nugget effect). In the other extreme, as is the case for the whole field simulation, the removal of a constant (modeled) value from the observed data should retain the spatial structure in the $\Psi_{PD,obs}$ within the residuals. Therefore, a higher C_i value would indicate that the residuals were not spatially structured. A C_i value equal to 100 corresponds to a nugget effect of the variogram (random effects). Concerning the preservation of spatial structure with respect to $\Psi_{PD,obs}$, two trends were interesting to note.

The first trend concerned the period up until the beginning of August, with the spatial structure of the model residuals becoming less and less structured over time for all modeling scales, i.e. C_i increases (Figure 2.7c). Thus, the residuals were randomly distributed by early August, as indicated by an absence of spatial structure, and it appeared that the spatialized model was performing well. However, note the non-significance of the spatial autocorrelation with respect to the residuals (Figure 2.7d). The same trend was observed for the spatial structure of $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$ (Figure 2.6), this loss of spatial structure was likely influenced by the precipitation events, which occurred just before the measurement dates, and had a homogenizing effect on the $\Psi_{PD,obs}$ values within the field.

The second result concerned the late summer period (mid-August to mid-September), when the residuals were randomly distributed for the site-scale modeling but spatially structured for modeling at higher spatial scales (zonal or whole field). The residual spatial structure at the higher spatial scales can

be explained by the fact that the $\Psi_{PD,obs}$ were spatially structured and the $\Psi_{PD,mod}$ were not for this period (Figure 2.6). The $\Psi_{PD,mod}$ were not structured because for all sites constituting the within-field zones and the whole field simulation scale, the $\Psi_{PD,mod}$ values were the same for all sites. Thus, modeling was not able to replicate the $\Psi_{PD,obs}$ spatial structure except for the case of the site-scale modeling, which seemed to be the best simulation scale regarding spatial indicators for this period.

The spatial structures of $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$ were quite different, except at the beginning of the modeling period when the spatial structures were comparable, but a period that coincided with low vine water stress that is considered of little importance to producers. Additionally, even if the spatial structures reflected by the C_i were similar, it is important to note that the level of semivariance for $\Psi_{PD,obs}$ and $\Psi_{PD,mod}$ were not of the same magnitude. The $\Psi_{PD,mod}$ semivariance was nearly 75% lower than $\Psi_{PD,obs}$ semivariance, indicating a much smaller range of values in this dataset. The variance differences were not shown with the C_i value because it is a result of the ratio between the nugget and sill values. Another point of interest is the range of the variograms used to calculate the C_i . The $\Psi_{PD,obs}$ variogram range was twice that of the $\Psi_{PD,mod}$ variogram range. Thus, the $\Psi_{PD,mod}$ spatial patterns (using the site-scale modeling) were smaller than the $\Psi_{PD,obs}$ spatial patterns.

Looking holistically at the aspatial and spatial indicators, there is no clear pattern to identify the best modeling scale. The site-specific and 5-zone modeling (finer spatial resolution models) tended to indicate the best model spatialization, for example lower C_i and higher Z-score values in mid-August), but this was not supported by better aspatial metrics (lower RRMSE and higher D-Index) at this time and at these finer spatial resolutions (Figure 2.7). Thus, the evaluation of model performance cannot be estimated by looking at only one or a combination of the indicators, i.e. different indicators point to a different best modeling scales, and none can be selected with certainty. For this reason, the metrics have also been interpreted relative to the decision process for within-field irrigation. This aimed to be able to decide which modeling scale was the most relevant, such that the $\Psi_{PD,mod}$ indicated the correct decision to be made (i.e. the decision that would have been taken with the $\Psi_{PD,obs}$ data).

2.2.4.4. Which modeling scale is the most relevant in regards to an operational context?

Figure 2.8 shows the translation of the model output, at each date and at each spatial scale, into an irrigation decision based on the date and the recommendation of Ojeda (2007) (Figure 2.5). Note that the dates up until and including August 5th had a consistent non-irrigation decision for the real data and the modeling at all spatial scales regardless of the actual quality of the prediction (e.g. with the indices shown in Figure 2.7). The operational decision-making to change the irrigation situation evolved in mid to late August, and the observed data indicated that the northern tip of the vineyard should be irrigated from August 18th 2004, and by September 10th the majority of the vineyard, except the southern third, should be receiving irrigation. For dates on or after August 18th, the scale at which the crop model was applied affected the quality of the irrigation decision. The whole field and 2-zone scale modeling flipped the whole field from non-irrigated to irrigated, although on same dates. The WaLIS model tended to simulate Ψ_{PD} more negative than $\Psi_{PD,obs}$, so all sites were predicted as requiring irrigation. By August 23rd, the overestimation of vine water stress by the WaLIS model led to an effective irrigation decision for the entire field for modeling at all spatial scales. Consequently, the remainder of this discussion will focus on the August 18th results. The highest balanced accuracy statistic (BA) was obtained with modeling at the 5-zone scale, which outperformed the site-scale modeling according to the BA (Figure 2.8). The BA was also higher for the 3 and 4-zone modeling than the site-scale modeling. These results showed that modeling at a spatial scale between 3 and 5-zones generated a better decision. It was not possible to identify this when using the proposed aspatial and spatial indicators. Evaluation in an operational context allowed an identification of which scale was the most relevant to model Ψ_{PD} as close as possible to the $\Psi_{PD,obs}$ decision.

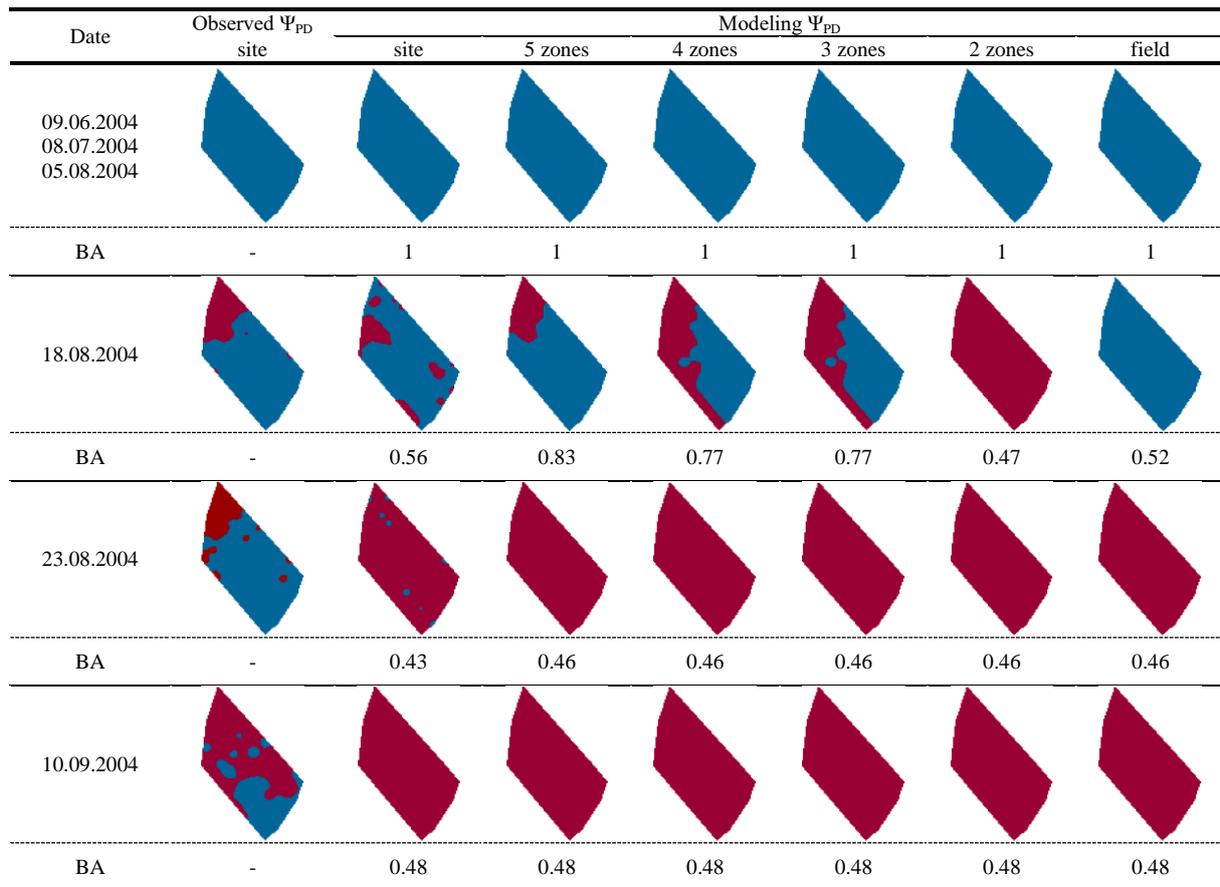


Figure 2.8 Maps of within-field irrigation according to the date and the value of predawn leaf water potential (Ψ_{PD}) for the observed point data and the modeled data at each considered scale of simulation. BA = Balanced Accuracy. Observed and site-scale modeled data maps were interpolated using inverse distance weighting.

2.2.5. Conclusion

In a precision agriculture context, using crop models at a finer scale than the model's native spatial footprint is of principal interest. Spatialization using downscaling processes is one of the methods that could be used to achieve this goal. Resulting spatialized crop models are currently often evaluated using aspatial and spatial indicators. However, interpreted individually, these indicators indicated different best simulation scales, thus, this study showed that using these indicators was not the most relevant method for assessing this kind of model application. The evaluation of spatialized crop models for precision agriculture in an operational context seemed to be a better evaluation method. Based on a decision-making approach, identifying the best simulation scale that was closest to the observed data was much easier and more relevant for assessing model performance. Ideally, a spatial indicator able to indicate if the zoning level is more relevant than another level to simulate an agronomic variable could be a great improvement. The spatial indicators used in this study are blind to this goal, which is why evaluation in an operational context was more relevant in this case. Spatial calibration is the process key here in the spatialization process, it would also be interesting to see to what extent the spatial structure of the agronomic variable also influences the relevance of this spatialization method.

2.2.6. Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

2.3. A new metric to evaluate spatial crop model performances

2.3.1. Details about the paper

2.3.1.1. Title and publication information

This section was presented as an oral communication at the 14th European Conference on Precision Agriculture (2023, Bologna, Italy) and published after peer-review in the conference proceedings. The full citation is:

Pasquel, D., Roux, S., Tisseyre, B. and Taylor, J. A. (2023). A new metric to evaluate spatial crop model performances. In: Proceedings of the 14th European Conference on Precision Agriculture, Bologna, Italy, July 02-06.

2.3.1.2. Authors

Pasquel D.^{1,*}, Roux S.², Tisseyre B.¹, Taylor J.A.¹

¹ ITAP, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

² MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

* Corresponding author: daniel.pasquel@inrae.fr

2.3.1.3. Abstract

In a precision agriculture context, the spatialization of existing crop models by downscaling processes to simulate agronomic variables at a within-field scale is of interest to better adapt technical decisions at this scale. The evaluation of spatial crop models needs to be based on both aspatial and spatial pattern error. However, current aspatial model metrics and existing spatial metrics have known limitations to evaluate the performances of spatial crop models. To address these limitations, a new metric, the spatial balanced accuracy (SBA), is proposed. The SBA is a novel metric, based on connectivity analysis that incorporates both aspatial and spatial aspects of model performance. The theory behind the metric development is presented here along with a comparison with existing model metrics applied to synthetic simulated data that covers a range of potential conditions.

2.3.1.4. Keywords

Downscaling, Connectivity analysis, Spatial patterns, Variograms

2.3.2. Introduction

The need for better information in precision agriculture (PA) is driving the increased use of existing strategic crop models for short-term and site-specific tactical applications (Pasquel et al., 2022a). This is enabled by an ability to spatialize these existing point-based crop models using data assimilation and spatial calibration approaches (Jones et al., 2017b). The spatialization of point-based crop models results in spatialized crop models. Equally, true spatial crop models are likely to become more available in agriculture (Pasquel et al., 2022a). For spatialized or spatial crop models (denoted globally as SCMs), having accurate spatial results is important to set up within-field management and to make correct in-season management decisions. Spatialization is considered as being correct if the observed data and the crop model output maps exhibit the same patterning.

To evaluate the performance of SCMs, particularly when applied at the within-field scale, metrics are needed to correctly assess the spatial pattern of the modelled agronomic variable (i.e. spatial organization of variable values) (Figure 2.9). A metric may have different goals: (i) to compare relatively different modelling approaches to understand model behaviour in varying theoretical circumstances (e.g. when changing model scale or the size of calibration/validation data sets) or (ii) to directly analyse how well a model has performed in absolute terms (i.e. obtaining real information on the model performance).

Regardless of the goal, in PA it is important to be able to quickly determine the effectiveness of a spatial or spatialization process and to identify the best performing SCM. Previously, classical aspatial metrics (e.g. root mean square error: RMSE) have been systematically used for model evaluation when models are applied spatially (see examples in Pasquel et al., 2022a). However, different spatial patterns in the model output can be achieved with a common RMSE (Pasquel et al., 2022a, 2022b). Thus, by themselves, existing aspatial metrics will not provide complete information of how well a SCM is spatially performing. An evaluation of SCM performance needs to account for both (i) the aspatial relationship between the values of observed and modelled variable and (ii) the preservation of the spatial pattern of the observed variable within the modelled variable. To achieve this, new metrics are needed that account for both the aspatial and spatial pattern error between the observed and modelled data.

In a PA context, the site-specific decision taken at a within-field scale will be directly based on the spatial pattern of the agronomic variable of interest (e.g. yield, vegetative vigour), but not directly on its spatial structure. The spatial pattern is a specific organization that derives from the spatial structure of the data, and multiple different spatial patterns can result from the same spatial structure (Figure 2.9). However, most geostatistical methodologies used in the PA community are based on the spatial structure (Leroux and Tisseyre, 2018), and for a given spatial structure will give the same result even if the agronomic variable shows differing spatial patterns. Therefore, any new proposed metric needs to be based on the spatial pattern error, i.e. the preservation of the variable spatial pattern, and resulting from an automated and robust approach.

To date, to the authors' knowledge, there has been no proposition of a metric in agronomic modelling that addresses both the aspatial and spatial pattern error. To address this gap, the proposition and theory behind a new metric for spatialized and spatial model evaluation is the main objective of this paper. Note that this metric could be applied in a general case of spatial modelling, but the focus here was made on downscaled PA applications. The proposed metric will (i) allow a relevant evaluation of SCM by assessing both aspatial and spatial pattern error, (ii) be based on an automated and robust approach, (iii) be intended to be used to identify which modelling approach is the best (and not to understand why the modelled data diverge from the observed variable) and (iv) be able to be used regardless of the agronomic context and/or modelling scale.

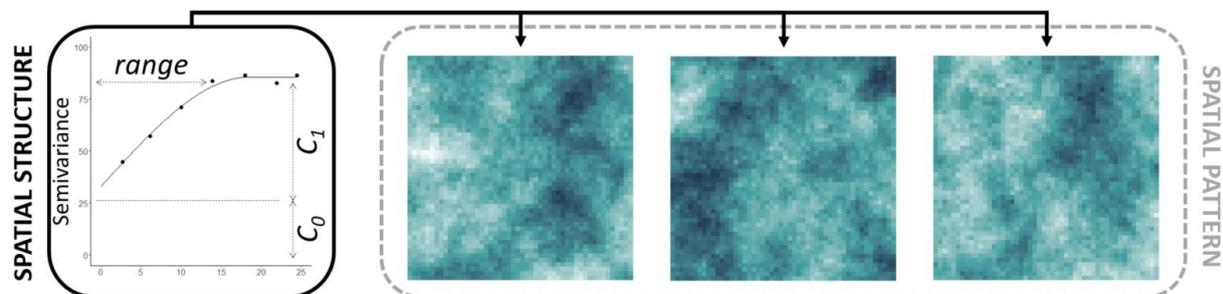


Figure 2.9 Difference between spatial structure and spatial pattern, the same spatial structure can result in different spatial patterns. Variogram parameters that are used to describe the spatial structure are C_0 : nugget, C_1 : partial sill, $C_0 + C_1 =$ sill and the range.

2.3.3. Material and Methods

Several metrics could be used to evaluate SCM performances. In this study, the decision was made to use the RMSE as an aspatial reference metric, $RMSE_{\text{vario}}$ as a variogram-based reference metric and $RMSE_{\text{con}}$ as a spatial pattern-based reference metric.

2.3.3.1. Aspatial metric: RMSE

The RMSE (Eq. 2.8) was chosen for the comparison as it is one of the most common aspatial metric used to evaluate SCMs (Pasquel et al., 2022a), by calculating the difference between observed and modelled data.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - M_i)^2} \quad (2.8)$$

where O_i = observed variable, M_i = modelled value and n is the number of observations.

2.3.3.2. Metric based on variographic analysis: $\text{RMSE}_{\text{vario}}$

Within the PA community, the use of geostatistics, particularly variograms, to evaluate the spatial structure of an agronomic variable is a well-known and accepted tool (Leroux and Tisseyre, 2018; Taylor et al., 2019). Variographic analysis evaluates the spatial autocorrelation structure between data points by computing an experimental variogram of the semivariance of the variable of interest at different distances (lags) (Eq. 2.9).

$$\gamma(h) = \frac{1}{2 \cdot |N(h)|} \sum_{(i,j) \in N(h)} |x(s_i) - x(s_j)|^2 \quad (2.9)$$

where h = the distance separating points, $N(h) = \{(i,j) : |s_i - s_j| = h\}$, $|N(h)|$ = the number of distinct elements of $N(h)$, $x(s_i)$ and $x(s_j)$ represent the agronomic variable respectively at location s_i and s_j .

Koch et al. (2017) suggested an adaptation to Eq. 2.9 to generate a metric that computes the difference between the semivariance of observed and modelled data at each lag (Eq. 2.10).

$$\text{RMSE}_{\text{vario}} = \sqrt{\frac{\sum_{h=1}^{N_{\text{lag}}} [\gamma(h)_{\text{obs}} - \gamma(h)_{\text{mod}}]^2}{N_{\text{lag}}}} \quad (2.10)$$

where $\gamma(h)_{\text{obs}}$ and $\gamma(h)_{\text{mod}}$ are respectively the semivariance computed at the distance h for the observed and modelled data and N_{lag} = the number of lags in the variogram.

As $\text{RMSE}_{\text{vario}}$ approaches 0, the spatial structure of the observed and modelled data tends to be the same. Therefore, the $\text{RMSE}_{\text{vario}}$ is of potential interest as a spatial or spatialized model metric as it integrates at least one aspect of performance, i.e. evaluation of spatial structure between observed and modelled data. However, it only addresses the spatial structure of the data, not the spatial pattern.

2.3.3.3. Metric based on connectivity analysis: RMSE_{con}

In PA, it is important to have SCM predictions that follow the real spatial distribution (pattern) of the variable of interest for management. Connectivity analysis is a method for assessing spatial pattern in hydrology modelling (Koch et al., 2017). It is adapted here to an agronomic context. Connectivity analysis is based on clustering neighbouring spatial model units of binary maps in order to compute the probability of connection according to Hovadik and Larue (2007) (Eq. 2.11).

$$\Gamma(X_t) = \frac{1}{n_{\text{tot}}^2} \sum_{i=1}^{N_{\text{clus}}(X_t)} n_i^2 \quad (2.11)$$

where X_t = the binary map obtained by thresholding a map X at threshold value t , $N_{\text{clus}}(X_t)$ = the number of distinct clusters in X_t , and n_{tot} (resp. n_i) = the number of spatial model units within X_t (resp. within the i^{th} cluster of X_t).

Several binary maps are considered in connectivity analysis by thresholding the initial data using the percentiles of the variable under study. To formalize Eq. 4 into a metric, Koch et al. (2017) further proposed the $RMSE_{con}$ to evaluate prediction performance by computing the difference between the probability of connection of observed and modelled data at each percentile (Eq. 2.12). Computation of thresholds based on percentiles make this metric insensitive to numerical bias.

$$RMSE_{con} = \sqrt{\frac{\sum_{q=1}^{100} [\Gamma(O_{t(O,q)}) - \Gamma(M_{t(M,q)})]^2}{100}} \quad (2.12)$$

where O and M are respectively the observed and modelled maps, $O_{t(O,q)}$ = the observed map binarized at threshold level $t_{(O,q)}$ defined relative to q^{th} percentile of O and $M_{t(M,q)}$ = the modelled map binarized using threshold $t_{(M,q)}$ defined relative to q^{th} percentile of M .

The closer $RMSE_{con}$ is to 0, the better the agreement between the (connected) spatial pattern of the observed and modelled data. Like $RMSE_{vario}$, $RMSE_{con}$ also integrates a spatial aspect of performance of a SCM. However, it has the advantage of directly assessing the spatial pattern, rather than the spatial structure (as in $RMSE_{vario}$).

2.3.3.4. Spatial balanced accuracy (SBA): a novel spatial pattern-based performance metric

In previous work (Pasquel et al., 2022b), the use of the balanced accuracy score (BA) (Eq. 2.13) was proposed to assess SCM outputs based on a form of map comparison. This approach required the SCM outputs to be carried forward into a decision system, i.e. it was not suitable for a rapid, robust, automated and direct assessment of the SCM performance. To overcome this requirement, an adaptation of the BA concept that incorporates part of the connectivity analysis methodology is proposed here. Neighbouring spatial modelling units are defined as in the connectivity analysis. Maps of observed and modelled data are used to generate several binary maps using a series of fixed thresholds. Considered thresholds are percentiles of the agronomic variable computed on values of observed and modelled data to be evaluated (Figure 2.10). For each percentile, the BA score is computed between the binary observed and modelled data to assess the spatial distribution and concordance of pixels below and above fixed threshold values for both maps. The different BA scores for all considered thresholds are averaged to generate the new metric, called the spatial balanced accuracy (SBA) (Eq. 2.14). When SBA is equal to 0, binary maps of observed and modelled data are identical.

$$BA = \frac{Sensitivity + Specificity}{2} = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (2.13)$$

where TP = true positive, TN = true negative, FN = false negative and FP = false positive, when $BA = 1$ there is a perfect agreement between observed and modelled map (and complete disagreement between maps when $BA = 0$).

$$SBA = \frac{1}{100} \sum_{q=1}^{100} [1 - BA(O_{t(O,M,q)}, M_{t(O,M,q)})] \quad (2.14)$$

where O and M are respectively the observed and modelled maps, $O_{t(O,M,q)}$ and $M_{t(O,M,q)}$ are respectively the observed and modelled maps at the threshold level $t_{(O,M,q)}$ defined relative to percentile q on the merging distribution of O and M .

The BA approach was preferred to accuracy to avoid misinterpretations in the case of unbalanced datasets. The closer SBA is to 0, the better the SCM performance. It can also be shown that $SBA = 0$ induces $RMSE = 0$ (i.e. the same values for all localizations), which is not guaranteed with the $RMSE_{vario}$ and $RMSE_{con}$ computation. In this sense, unlike $RMSE_{vario}$ and $RMSE_{con}$, SBA incorporates both aspatial and spatial pattern errors between observed and modelled data.

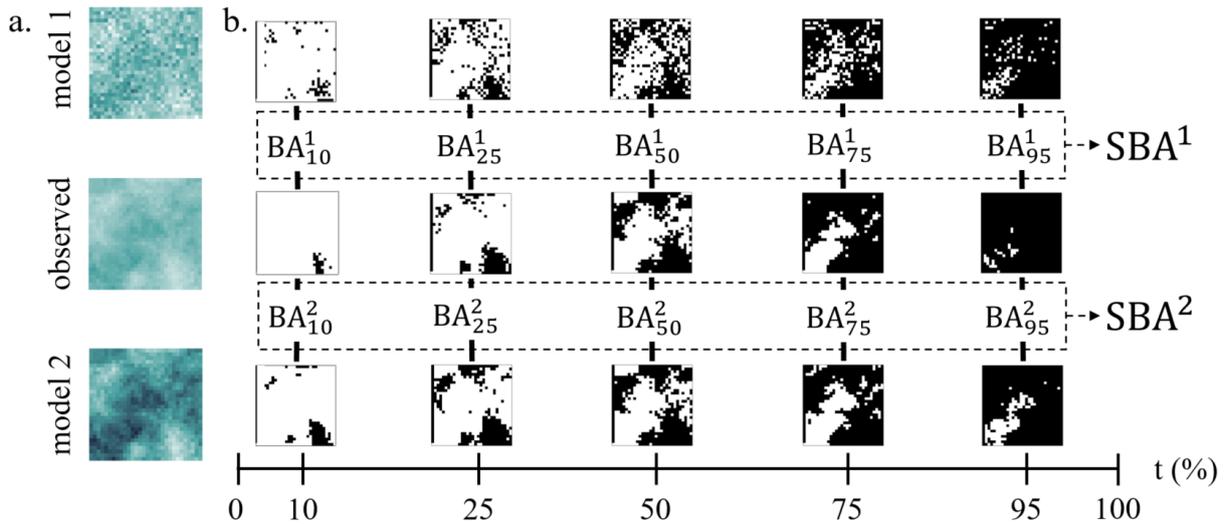


Figure 2.10 Representation of spatial balanced accuracy (SBA) computation. (a.) Maps of an observed and modelled agronomic variable resulting from different spatialized crop models. (b.) Binary maps generated for observed and modelled maps computed for the 10th, 25th, 50th, 75th and 95th percentile thresholds over the entire range of observed and modelled data. BA is computed for each considered percentile thresholds (subscript values) and SBA is computed for each different modelling approach (superscript values).

2.3.3.5. Simulation study

To assess the performance of the different metrics under controlled conditions, simulated data were generated for various scenarios considered relevant to PA applications. A virtual field (50 x 60 pixels) with a strong spatial structure of the simulated agronomic variable was generated using a spatialized Gaussian field with the *gstat* R package (Gräler et al., 2016; Pebesma, 2004) in R 4.2.0 (R Core Team, 2022). The 3000 pixel values were defined using a theoretical variogram fitted with a spherical model. The variogram parameters were chosen as follows: nugget = 1, partial sill = 100 and range = 20. The agronomic variable mean was fixed to 50. This field was then trimmed to 50 x 50 pixels to form the reference simulated observed field (Figure 2.11).

Five simulated SCM outputs were then generated. The first three were obtained by applying a form of error (or noise) distribution to the reference field. A constant normal distribution $N(0,5)$ of error was generated and then applied (A) randomly, (B) with a positive relationship and (C) as for B with a constant bias added (Figure 2.11.A-C). These cases represented situations where the theoretical SCM was performing (A) poorly, (B) well and (C) well but with a bias. The fourth and fifth simulated SCM outputs were generated by (D) shifting the reference field 10 pixels horizontally within the original 50 x 60 field, to obtain a new field from the simulated observed data with an identical spatial structure but an offset spatial pattern, and (E) by randomly generating a second field with identical variogram parameters of the original reference field and completely different spatial pattern (Figure 2.11.D-E). Thus, Model E has the same spatial structure as the simulated observed data (i.e. same variogram parameters) but results in a different spatial pattern. These represented situations where the theoretical SCM was performing (D) quite well but with a spatial bias and (E) poorly. A well-performed metric should be able to identify Model B as the best theoretical SCM.

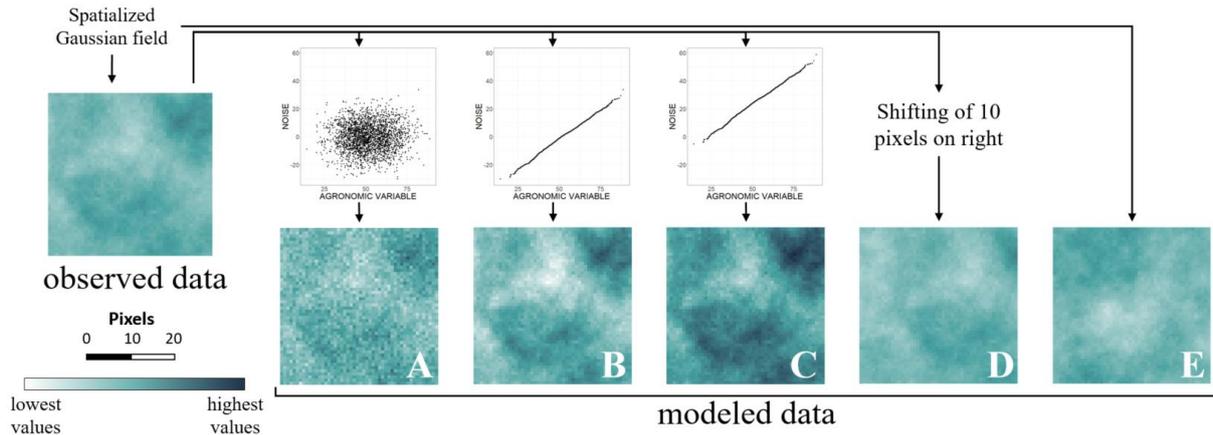


Figure 2.11 Illustration of the simulated spatial agronomic variable maps used to evaluate theoretical spatialized crop models. Observed data were generated from a spatialized Gaussian field. Modelled data were simulated by adding or creating different error distributions to the observed data: (A) random noise, (B) positively applied noise, (C) positively applied noise with a numerical bias, (D) spatial pattern translation of the simulated observed data and (E) same spatial structure but with a different spatial pattern.

2.3.4. Results and discussion

The results of the calculations for the five different simulated SCM outputs (Models A to E) relative to the simulated observed reference field are shown in Table 2.2. Using RMSE, Models A and B were identified as best and identical. The RMSE was unable to differentiate between them as the same error distribution was attributed in both models. RMSE is an aspatial metric thus only estimates the aspatial relationship between values of observed and modelled data. Model A should not be identified as equivalent to Model B as the spatial pattern is more distorted in Model A. This has previously been shown with other data sets (Pasquel et al., 2022b, 2022a).

Using $RMSE_{vario}$, Model E was identified as the best model, followed by Model D. The $RMSE_{vario}$ assesses spatial structure (with variogram parameters). Variograms of the reference observed data and Model E and D outputs were very close as they were generated from a spatial Gaussian field with the same variogram parameters, i.e. the same spatial structure. Given this, it was unsurprising that these models were identified as the best models by $RMSE_{vario}$. In contrast, Models B and C were the worst performing according to $RMSE_{vario}$, and they were unable to account for the numerical bias introduced in Model C. This was a poor result given that Model B should be considered as the best performing approach. It highlights the difference between assessing the spatial structure and spatial pattern in SCM outputs, and its effect on geostatistical metrics like $RMSE_{vario}$. Furthermore, estimating the spatial structure of an experimental variogram involved fitting a theoretical variogram, which required expert knowledge that is incompatible with the aim of having an automated and robust metric.

Table 2.2 Comparison of different metrics to identify the theoretical SCMs with the best performance. Each indicated value was computed between observed and modelled data.

Data	Model A	Model B	Model C	Model D	Model E
RMSE	9.90	9.90	26.63	14.10	17.90
$RMSE_{vario}$	106.49	317.87	317.87	21.67	9.14
$RMSE_{con}$	0.16	0.05	0.05	0.17	0.10
SBA	0.18	0.04	0.17	0.37	0.50

$RMSE_{con}$ identified Models B and C as the best models but was unable to differentiate between them. The spatial patterns of both models are the same, but the aspatial relationship is biased for Model C, i.e. the $RMSE_{con}$ is a spatial metric that is insensitive to numerical bias. Regarding the objective of

this study, $RMSE_{con}$ is more relevant than $RMSE_{vario}$ as it is responsive to spatial patterns, rather than spatial structures and is more easily automated. However, a drawback to using connection probabilities to evaluate SCM performances is that it detects the presence of patterns at the within-field scale but not their location. Thus, for a given output, such as Model B, the $RMSE_{con}$ would be identical if the output map was rotated 90° , 180° or 270° , i.e. if it had the same spatial pattern output but with different locations.

Like RMSE and $RMSE_{con}$, the result using the novel SBA metric also identified Model B as the best model. However, SBA clearly identified Model B as the single best model (and not equal to another model as in the case of both RMSE and $RMSE_{con}$). Model B is considered as the best because it has the closest aspatial relationship between value data and preserves a maximum of the spatial pattern relative to the observed data. The SBA methodology is mainly based on connectivity analysis, which assessed the variable spatial pattern. However, by also evaluating thresholds (i.e. percentiles for this study) across all values of observed and modelled data using the BA theory, the SBA also took into account the aspatial relationship between the observed and modelled data. The $RMSE_{con}$ was unable to do this because the thresholds were computed independently between the observed and modelled data. The methodology to assess the variable spatial pattern was kept (i.e. it computed a metric for different binary maps for different defined thresholds between observed and modelled data), but BA was computed instead of the probability of connection. By evaluating both the spatial pattern connectivity and the placement of these spatial patterns, the SBA was able to correct the second drawback of connectivity analysis. This was shown by its lower result (higher SBA value) of Model D compared to B and C. Thus, from this short study with limited simulations, the SBA appears to be the most relevant and promising metric among those tested for evaluating SCM performance. Further research will focus on testing the SBA via more in-depth sensitivity analyses and in real case studies to verify the ability of this metric to correctly characterise SCMs.

2.3.5. Conclusion

This work has proposed a new metric, spatial balanced accuracy (SBA), to address the issue of how to evaluate the performance of a spatial or spatialized crop model (SCM). The SBA may have accounted for both the aspatial relationship between the values of observed and modelled variable(s) and the preservation of the spatial patterns of the observed variable(s) within the modelled variable(s). It is based on connectivity analysis with a modification to correct drawbacks of this methodology to evaluate the outputs of a SCM. In this preliminary study, the SBA gave relevant results on theoretical SCMs using simulated data that encompassed a variety of conditions. More detailed simulations are planned to verify these results.

2.3.6. Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

2.4. General conclusion of the chapter

The objective of this chapter dealing with the theoretical part of this manuscript, was to investigate the way spatialized crop models are currently evaluated and the suitability of the metrics being used. The research work presented here has highlighted several scientific points. Firstly, it has clearly highlighted that current aspatial and spatial metrics used by the precision agriculture and crop modeling communities are not the most relevant metrics to assess the spatialized crop model performances. The simulated examples presented here have shown that in some particular cases, the interpretation of existing aspatial and spatial metrics could lead to a misinterpretation of the outputs of spatialized crop models. Therefore, different modeling approaches could be identified as performing equally well when this is not actually the case. Indeed, aspatial metrics only evaluate numerical error without taking into account spatial error.

Spatial metrics are mostly based on spatial structure and are designed to be used as descriptors and not as assessing metrics. Thus, using these metrics for evaluating spatialized crop model performances could lead to wrong interpretations.

The work presented in this chapter demonstrated that evaluating spatialized crop models on both aspatial and spatial errors is necessary to provide a relevant and reliable evaluation on the models. It is not possible to just use an aspatial or just a spatial indicator. Currently, the scientific community commonly accepts the use of multiple metrics in order to have a better interpretation of model outputs. This is especially the case for spatialized crop models which are evaluated on both aspatial and spatial error, even if the used metrics have been shown to be flawed. It was clear from the original literature review (Chapter 1) and the simulations in this chapter that a new and different type of metric was needed to help evaluate the performance of spatialized (or spatial) crop models. To this end, the second part presented in this chapter focused on the development of a combined aspatial and spatial metric to fill this gap, by proposing the spatial balanced accuracy (SBA) score.

Based on connectivity analysis used in the hydrology community, the SBA score allowed for a more relevant evaluation of spatialized or spatial (crop) models than existing crop model metrics. SBA score allows for a comparison between observed data and modeled data resulting from several spatial modeling scales. Thus, this metric represents a relative evaluation of spatialized crop models. The connectivity analysis is based on the assessment of spatial pattern instead of spatial structure. In a precision agriculture context, spatial pattern is the primary interest for decision-making about the spatial footprint of potential cultural practices. Interpretation of this metric could also be made through several spatial modeling scales allowing an identification of which spatial scale is the most relevant. This chapter demonstrated the contribution of SBA score and throughout the rest of this manuscript, this metric will be used to help to evaluate spatialized crop model performances (Chapters 3 and 4). Thus, SBA will be applied to different use cases to study its consistency. In the following chapters, the SBA score will be used in a relative evaluation between different spatialized crop model outputs to identify which scale is closer to the observed, validation data.

However, as this is a new metric, the suggestion from this work is to only use the SBA metric for relative comparisons until further analysis of the SBA score behavior is performed. The SBA was designed for comparison of modeling results from the same crop model, for the same crop, for the same agronomic variable and for the same field, but at different spatial scales. It should only be used in this context at the moment. For now, it is not possible or not relevant/reliable to use it for comparisons between completely different modeling contexts. Indeed, the proposed approach has not been tested or validated for this type of application and, until this has been done, results generated in such a context should be carefully interpreted regarding their relevance. In this manuscript, the SBA score was only employed for very specific cases, for a relative comparison of modeling accuracy. The value produced by the SBA is not as a RMSE value. SBA score interpretation relative to an operational use (e.g. difference in water or nitrogen supply regarding to a reference) remains difficult without more reference to help it. More works should be led to better understand how the two components (aspatial and spatial) are taken into account to be able to interpret SBA score correctly. Moreover, more research is needed to clearly defined a possible value domain to improve the interpretability of the SBA score in other context or for other purposes. The next steps on the SBA research are to investigate the relevance of this metric in any use-case, and to define the anticipated range of values and to propose an interpretation of the SBA values. These next steps do not exclude a modification of the SBA score if necessary.

Chapter 3

Focus on spatial calibration approach as a downscaling process to spatialize crop models

3.1. Intention note

Crop models are commonly used to reproduce the biophysical processes occurring through the soil-plant-atmosphere interaction. Generally, the overall aim is to predict an agronomic variable of interest (e.g. final yield, produced biomass, crop quality) under varying environmental and management variables. However, although these crop models are very useful for reproducing the conditions of crop growth and interactions with its environment, they are only models and cannot reproduce the world perfectly. As a result, they are subject to uncertainties. Several bias and uncertainties can affect crop model performances. The origin of these uncertainties can be grouped in three different sources regarding (i) the model structure, (ii) the input data and (iii) the model parameters (van Oijen and Ewert, 1999).

Reduction of these uncertainties is of primary interest for the agronomic and modeler communities to increase the reliability of crop model prediction. The uncertainty linked to the crop model structure is inherent to the design of the crop model. This source of uncertainty is mainly based on the representation of crop growth by the crop model equations. Crop models differ in the degree of equation complexity used internal, but regardless of their complexity, they cannot exhaustively describe all variables that influence plant development. Thus, depending on their complexity and parametrization (i.e. the internal parameter estimation), uncertainty linked to model structure can have a huge impact on the accuracy of crop model predictions (Kamali et al., 2022; Kawakita et al., 2020). The uncertainty related to input data is mainly based on the error of measurement and sampling. Thus, selecting the most suitable data, relative to the model equations needs, will increase the crop model performance. Error(s) in the input data will propagate through the modeling process. It is for this reason that many crop models are designed and constructed at the field-scale to use mean field input data, to ‘smooth’ the input data by minimizing stochastic errors in these data and ensuring that the data fit within observed and constructed norms that suit the model equations. For more accurate modeling (and this is particularly true for predictions at the within-field scale), the increase in the availability of high-resolution input data that could better reproduce the spatial pattern of the agronomic variables is of primary interest to end-users reduce this source of uncertainty (Laux et al., 2021). Such high-resolution data could be aggregated to generate mean field information to satisfy the native crop model structure. Aggregation smooths these data. However, when used in a data assimilation or, as proposed here, a spatial calibration approach, these data are used at their native spatial footprint, not the model’s footprint, and may contain different stochastic errors from more ‘typical’ model input data. The sensor type and the mode of data collection will influence the uncertainty in these high-resolution spatial data. In general, high-resolution sensor data are considered to have a higher level of stochastic error at an individual point, but this higher stochastic error is offset by the higher number of observations available within a given area when used for mapping or decision-making (Tisseyre et al., 2018). However, these high-resolution data are not necessarily directly the crop model input(s) or output(s) and can represent data more or less correlated with these crop model variables. Thus, depending on the level of correlation between the input/output of the crop models and the high-resolution data, uncertainties can occur when using these data for modeling processes. The influence of this effect when these data are used to downscale crop models is

still not well understood in the crop modeling community, it is referred here as the second source of uncertainty. The third source of uncertainty related to model parameters can be managed by an optimization of available data and the model calibration process to adjust model parameters, and this process has a key role to play and a great potentiality to reduce crop model uncertainties (Gao et al., 2020; Seidel et al., 2018). However, there are a diversity of existing methods that can be used to calibrate crop models and the consensus among model practitioners is that the optimum way to calibrate crop models is still an open, unanswered question among the scientific community (Confalonieri et al., 2016; Seidel et al., 2018). This means that beyond the data used for the calibration, the way the calibration proceeds will also have a significant impact on the crop model performances. This highlights the importance of careful selection of data and methods for the calibration process in order to optimize the crop model performance, and this will also be true for spatialized crop models.

The spatial calibration process presented in this research work is based on a methodology involving several steps. First, high-resolution ancillary data that has relatively similar spatial patterns with the considered agronomic variable to model is used to delineate within-field calibration zones using a segmentation algorithm. It is assumed that both are spatially structured, which justify the spatialization of crop model. Then, a frequentist method of calibration is applied to calibrate certain crop model parameters to the different zones, using zone-specific input data. Finally, the most relevant modeling scale is identified for the modeling process. It is important to note that this spatial calibration method does differ from data assimilation in the way that the high-resolution ancillary data are used. There is no ‘forcing’ of the model with these ancillary data. The ancillary data are only used to identify calibration zones, at various scales, which allows model parameters to be calibrated spatially. The calibration of the model parameters is done in a ‘classical’ sense, using typical, available input data averaged at the zone level, i.e. observed data not high-resolution, sensor-derived data. The amount of input data available for calibration is fixed. Therefore, as the number of zones increases, the data available per zone, on average, will decrease, which again will bring more stochastic (measurement) error effects into the zone means (compared to the field, or 1-zone, ‘average’ modeling), which needs to be investigated. The benefit of this spatial calibration approach, compared to a data-assimilation approach, is that the model can be calibrated and run from the start of the season, or even pre-season, to test management scenarios. Data assimilation requires in-season information to ‘force’ or ‘update’ the model predictions, thus can only be used once the growing season has begun. The drawback to this proposed spatial calibration is of course the need for relatively high-resolution, geo-referenced calibration data.

When developing the methodology explained in the following articles, there were several methods investigated and tested to propose the most successful spatial calibration approach (for these conditions). Concerning the within-field segmentation, other methods were tested, such as the *k*-means algorithm and its variants. However, conserving the spatial characteristics of the segmented data (zones) was found to be very important to the modelling process, i.e. contiguous zones, rather than spatially fragmented classes made more sense practically (for management) and for the modeling. Therefore, to delineate within-field zones, the segmentation algorithm (Pedroso et al., 2010) from *GeoFIS* R package (Guillaume and Lablée, 2022) was ultimately used for its ability to take into account the spatial characteristics of the data and define coherent zones. This approach was chosen because the data could possibly be on an irregular grid. The data could have been interpolated on a regular grid allowing the use of several different zoning methods but would have required kriging and thus would have modified the data distribution. Introduction of a smoothing (e.g. kriging) would not have allowed to specifically study the effect of the crop model but the effect of both crop model and interpolation method, which was not the intent of this work. Thus, the choice to not interpolate the data on a regular grid was made and this algorithm was used. Note that delineate calibration zones was based on different purposes tant management zones delineation. In addition, a relatively simple frequentist approach based on a mean absolute error was chosen to calibrate the crop model parameters. The main intent in the proposed spatial calibration approach was the spatial constraint (multi-scale zoning) enforced during the calibration.

Thus, a simple calibration method was chosen but the spatial calibration could obviously be done using more complicated approaches, such as a Bayesian approach for instance, that could introduce information related to the uncertainty associated with the calibrated parameter value and help improve model calibration.

The research work presented in this chapter aimed to investigate the relevance of using a spatial calibration approach depending on the considered agronomic variable. The objective was here to understand if a spatial calibration could improve and optimize the spatialized crop model performances. This relevance was investigated depending on the agronomic data characteristics: its spatial structure, its variability and its correlation with an associated ancillary data that was used to spatialize (zone) the field at various scales. Note that the range of the ancillary data correlation considered in this work was very spread (from 94% to 10%). Other levels of correlation were tested but have not been reported as the output results did not show any significant difference to those given already. However, this range of correlations was selected based on reported real studies from the literature, reflecting a realistic use of ancillary data correlations.

Similarly, this investigation was based on simulated data related to vine water status through the predawn leaf water potential. However, to ensure the temporal consistency for a reliable calibration process, these simulated data were based on real measured data from an experimental vineyard. A dedicated method of simulation data was designed to this aim. The relevance of the proposed spatial calibration was investigated using a relatively simple crop model WaLIS and its performance was evaluated using the RRMSE and the SBA score (Chapter 2). The spatial calibration process (involving all the different calibrations at each different considered spatial scale) and the evaluation of their performances using the SBA score required a lot of computing time, which led us to use a high performance computer server (MESO@LR-Platform) for this study.

3.2. What drives the performance of a crop model spatialized at the within field scale?

3.2.1. Details about the paper

3.2.1.1. Title and publication information

This section was submitted to *Precision Agriculture* in October 2023 and is currently under review.

3.2.1.2. Authors

Pasquel D.^{1,*}, Taylor J.A.¹, Tisseyre B.¹ and Roux S.²

¹ ITAP, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

² MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

* Corresponding author: daniel.pasquel@inrae.fr

3.2.1.3. Abstract

The calibration process is a necessary step to improve mechanistic crop model performances and to reduce the uncertainty related to model parameters. Modeling agronomic variables at the within-field scale needs to have an accurate reproduction of the spatial pattern to enable a suitable spatial management. Spatial calibration, as proposed here, is a downscaling method for crop model spatialization that includes spatial constraints by means of zoning and ancillary data. The objective of this study was to investigate the relevance of using spatial calibration at the within-field scale compared to classical calibration (without spatial constraint). To this aim, the impact of three driving properties was studied: the targeted variable variance, its level of spatial structure, and the correlation level of ancillary data. A method was designed to extrapolate various spatial structures from the one of an

existing vineyard whilst ensuring the temporal consistency of the synthetic data. The method was applied to a vineyard crop model (WaLIS) using predawn leaf water potential (Ψ_{PD}) measurements and synthetic ancillary data correlated to Ψ_{PD} . Results showed that spatial calibration improved the model performances at the within-field scale when Ψ_{PD} was strongly structured and highly correlated to the ancillary data. When Ψ_{PD} was either moderately or weakly spatially structured, the spatial constraint added more error instead of correcting errors, deteriorating the crop model performances. The ancillary data and the segmentation algorithm had a large impact on the spatial calibration performances. This study identified the cases where spatial calibration could be used to improve within-field modeling.

3.2.1.4. Keywords

Predawn leaf water potential, WaLIS, Spatial calibration, Precision agriculture, Spatial pattern

3.2.2. Introduction

Crop models are useful tools for predicting agronomic variables and interactions along the soil-plant-atmosphere continuum. Mechanistic crop models, especially the complex ones that require several parameters for their functioning, may involve unknown values or values with large domains of uncertainty in their parametrization. These uncertainties affect crop model outputs and thus reduce modeling performances. Therefore, investigating and understanding how to reduce the sources of uncertainty to increase the reliability of crop models is of primary interest to developers and users to have more relevant models (Wallach et al., 2016). The sources of uncertainty in predicting agronomic variables are well known and are based on (i) the model structure, (ii) the input data and (iii) the model parameters (van Oijen and Ewert, 1999). Seidel et al. (2018) highlighted suggestions to improve identified sources of uncertainty for crop modeling, especially for model parametrization via a calibration process. These sources of uncertainty are valid for both classical or spatialized applications of crop models. However, it is unknown how a spatialization process may affect these uncertainties.

The calibration process aims to adjust the crop model parameters to reduce the difference between the observed measurements and the modeled values. Initially, crop models are designed in a constrained domain that involves certain climatic and environmental contexts. The principal objective of a recalibration process is to be able to extend the crop model beyond this initial domain, because parameter values cannot be universally representative (Coucheney et al., 2015; Wallach, 2011). However, because the real world is so variable, a calibration approach is useful and even necessary to reduce uncertainties in crop model prediction regardless of the target domain. Consequently, calibration processes to reduce model parameter uncertainty are the second most important field in crop modeling uncertainty studies (Chapagain et al., 2022). Indeed, a large diversity of calibration methods and steps exist and, depending on the suitability of the approach selected, bias in the modeling may occur and can have a large effect on the crop model performances. Thus, the choice of a suitable calibration method is paramount. Most studies involving crop model calibration tend to use different methods that can be either frequentist, e.g. by Ordinary Least Square (OLS) (Ramirez-Villegas et al., 2017; Zeng et al., 2016), or Bayesian, e.g. Generalized Likelihood Uncertainty Estimation (GLUE) or Markov Chain Monte Carlo (MCMC) (Jha et al., 2022; Li et al., 2018; Sexton et al., 2016; Zhang et al., 2019). Gao et al. (2020) compared three calibration methods to perform crop model calibration. These authors concluded that OLS was a fast and efficient method although MCMC was reliable to investigate the uncertainty associated with the estimation of crop model parameters.

Regardless of the method used, the calibration of process-based crop models will involve many decision steps, and the best approach to realize these steps may have no consensus among the agronomic community (Wallach et al., 2021). The literature shows that many studies have already compared several calibration approaches for crop models, highlighting the diversity of methods used and the potential gap in performances between methods (Seidel et al., 2018). By focusing on the calibration of phenological parameters of complex crop models, which only represent a part of the whole model parameters that can

be calibrated, Wallach et al. (2021) highlighted the diversity of potential approaches that could be used and tried to propose some guidelines to homogenize practices. These guidelines directly refer to the careful selection of the method to optimize the crop model calibration. Many calibration methods are known to be efficient, but their efficiency can differ depending on the modeling goal. Thus, the right calibration method should be selected in accordance with the specific modeling goal. Moreover, using the right metric in agreement with the modeling objectives is also an important part of calibration (Yang et al., 2014). In other words, crop model calibration is not standardized among crop model practitioners and there are a lot of methods and tools to calibrate and evaluate them.

In a precision agriculture context, crop models could be used to identify the most relevant management strategy to be applied at the within-field scale. However, applying crop models at a the within-field scale involves some form of model spatialization to adjust the native spatial footprint of the model to a finer-scale prediction of agronomic variables (Pasquel et al., 2022a). However, downscaling crop models may increase input parameterization that may not lead to improve model performances at finer scales (Zhen et al., 2023). The spatialization of crop models can be achieved upstream or downstream of the crop model modeling and thus, does not fundamentally change the internal structure of the model, i.e. the crop model still operates the same way (Pasquel et al., 2022a) and changes are made before the model is run (upstream) or after the model has been fully or partially run (downstream). Calibration approaches to change the spatial footprint of crop models are considered to be upstream approaches. In the case of crop model downscaling for precision agriculture purposes and for modeling at the within-field scale, recalibration of model parameters is performed to adjust for local site-specific variability without changing the model equations.

While model calibration can be used for crop model spatialization (for either upscaling or downscaling applications) it has been rarely used (Basso et al., 2001; Cammarano et al., 2021; Leo et al., 2023). Spatialization has been more commonly achieved via data assimilation approaches, which are relevant when in-season data are available to adjust/correct the crop model. Data assimilation approaches are not strictly relevant for the pre or early-season use of crop models in precision agriculture. In contrast, if historical spatial data are available and they are known to influence certain crop model parameters then, by correctly selecting the right spatial data and the right crop model parameter(s), the targeted crop model parameter(s) could be spatially calibrated based on historical data. An example of this could be the adjustment of soil moisture characteristic according to soil type (particle size distribution) using high-resolution soil maps. The soil type will be stable over time so any spatial calibration in year n can be confidently carried forward into year $n+1$. This spatial calibration will allow the crop model to be spatialized (downscaled) before the season starts. Such an approach has been used recently to demonstrate the ability of spatial calibration to downscale the APSIM model (Holzworth et al., 2014) within a cereal field in Italy (Pasquel et al., 2023a). However, how effective this spatial calibration approach is and the scale of spatialization that can be achieved has been poorly investigated so far.

To address the lack of knowledge in this area, this study was conceived to investigate the relevance of applying a spatially constrained calibration approach, i.e. a spatial calibration (Pasquel et al., 2023a), at the within-field scale to preserve the spatial pattern of the agronomic variable. An assumption behind this previously existing and tested method is that constraining the spatial pattern of the agronomic variable will allow a spatially consistent smoothing of the calibration error relative to the measured data set. Given that fewer data, and data with a higher level of noise will be available for finer-scale calibration, it is not immediately apparent if a (spatial) calibration at a finer spatial scale will be able to efficiently reproduce the spatial pattern of the target agronomic variable (Tisseyre et al., 2018). A spatial calibration approach could lead to higher noise in the calibration resulting in a poorer overall model performance, i.e. a poorly calibrated local crop model is likely to be less useful than a properly calibrated field-scale model for management. How to determine the correct scale at which to perform a

spatial calibration is an obvious research question that is still to be addressed by the crop modeling community.

Within this broad context, the objective of this study is twofold. Firstly, and most importantly, the work is realized to better understand how relevant it is to spatialize crop models by using spatial calibration to predict agronomic variables at the within-field scale using a relatively simple crop model. More specifically, the goal is to investigate the effect of three components hypothesised to have a significant impact on the spatial calibration process : (i) the characteristics of the spatial structure/pattern of the agronomic variable of interest at the evaluation date, (ii) the total variance of the agronomic variable at the evaluation date, and (iii) the degree of correlation between the agronomic variable of interest and the ancillary data used to delineate within-field calibration zones. All three were investigated to understand how these effects influenced the spatial calibration process of a simple crop model. In this case the model was WaLIS (Celette et al., 2010) that is designed to simulate predawn leaf water potential (Ψ_{PD}) in vineyards as a season progresses (with the intention of managing water stress). As it is a temporally evolving model, to be able to calibrate the WaLIS spatialized crop model, Ψ_{PD} data from several dates were necessary. Consequently, to investigate the relevance of the three identified effects, a large number of different fields were needed and it was therefore necessary to generate specific temporally coherent synthetic data¹. This represented a secondary objective, namely, how to properly simulate data, like Ψ_{PD} , that temporally evolves over the season. This is fundamentally different to other agronomic variables, such as yield, which have a fixed value at a certain time point (e.g. harvest). To respect the temporal evolution of Ψ_{PD} , it was necessary to generate spatially rearranged fields (SRFs) from real field data that represented Ψ_{PD} at different phenological stages. Issues arising from the generation and the use of synthetic data in this way for study, and for future studies, will be also discussed along with the impact of the spatial structure/pattern of the target variable, temporal variability in the target variable and the role of ancillary data in defining calibration zones.

3.2.3. Material and Methods

3.2.3.1. General methodology

The whole method applied for this study is described in Figure 3.1 and is further detailed in the following sections. As multiple different hypotheses were being tested, several different fields with identified spatial structures of model inputs and outputs were needed. However, such an agri-dataset was unavailable and in reality would be quite complicated to obtain, thus unlikely to ever be available. Therefore, synthetic data were generated in order to control all the desired properties of the fields and to study the relevance of the spatial calibration in specific different cases. These synthetic data should also present a temporal consistency, and this was achieved by using available data from a real case study (one single field over several years) as a template for the generation of the synthetic data.

The classical calibration approach was based directly on the synthetic vineyards and applied at two scales. Firstly, a classical model calibration was performed using field averages (which is how the model is designed to work). Following this, each point was individually calibrated, using the classical approach, whereby the observed values at each point were effectively considered to be ‘field averages’ (site scale calibration was not considered as spatial calibration because there was no direct spatial constraint in the calibration process thus directly equivalent to a classical calibration method). This resulted in crop model outputs at either the field or the site scale.

The spatial calibration was defined as a calibration process based on the delineation of calibration zones from ancillary data that were representative of the spatial pattern of the considered agronomic variable. For further details about the spatial calibration approach please refer to Pasquel et

¹ The term ‘synthetic data’ and ‘simulated data’ refers in this article to the virtual data generated for the need of this study. The term ‘modeled data’ refers to the data resulting from the modeling using the crop model, i.e. modeled data are considered here as crop model outputs.

al. (2023a). After the delineation of the calibration zones, the mean value of the measured sites within each zone was used for the spatial calibration, i.e. the modeling was performed at the spatial scale of the zones. Crop model outputs resulting from the classical calibration approaches and the spatial calibration approaches were compared to estimate the relevance of a spatial calibration to spatialize crop model at the within-field scale. Regarding these results, several key hypotheses were investigated: (i) a stronger spatial structure of the target variable will provides a better option for spatial calibration, (ii) a higher variance of the agronomic variable will generate a better opportunity to have contrasting calibration zones (and ultimately more effective modeling with spatial calibration) and, (iii) the effectiveness of the calibration zone delineation is dependent on the correlation of the target variable with the ancillary data. Indeed, the zones derived from ancillary data are assumed to be more relevant to define calibration zones when the correlation between both information sources is higher.

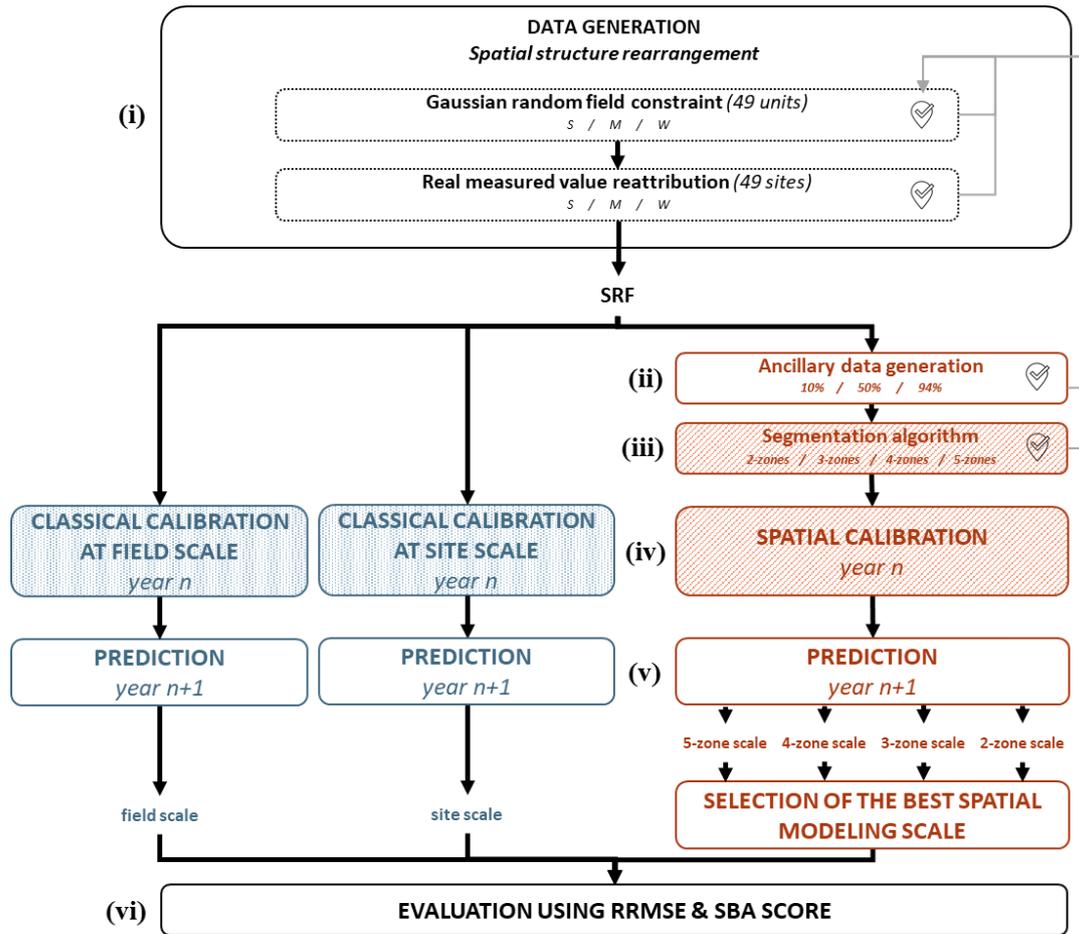


Figure 3.1 The method used to compare the classical calibration approach and the spatial calibration approach and to assess the relevance of this latter. (i) Data generation by rearranging the real measured sites to match with the targeted spatial structure corresponding to three spatial structure modalities: S (strong), M (moderate) and W (weak). Each check mark corresponds to a verification step, if the synthetic field do not fill the considered modalities for the considered step, this field is discarded and another is generated. (ii) Simulation of the ancillary data used for the spatial calibration approach with three different levels of correlation with the rearranged site of measurement: 10%, 50% and 94%. (iii) Ancillary data segmentation performed with the segmentation algorithm to delineate different resolution of within-field zones: 2-zone, 3-zone, 4-zone and 5-zone. (iv) Calibration of the crop model using either a classical calibration approach (blue paths) or a spatial calibration approach by identifying the spatial parameters (red path). (v) Modeling the agronomic variable at the different spatial modeling scales considered: field scale, 2-zone scale 3-zone scale, 4-zone scale, 5-zone scale and the site scale. (vi) Evaluation of the calibration approaches using the relative RMSE (RRMSE) and the spatial balanced accuracy (SBA) score.

3.2.3.2. Field data and crop model

3.2.3.2.1. Temporal consistency by using real data field observation

The experimental vineyard used for this study was a 1.2 ha Syrah vineyard located in the south of France (43.144°N, 3.131°E, Gruissan, Aude) on the INRAE Pech Rouge experimental site. The agronomic variable of interest was the predawn leaf water potential (Ψ_{PD}). Ψ_{PD} was measured for two consecutive years (2003 and 2004) on several dates (respectively 7 and 6 dates) at the within-field scale on 49 measurement sites using a pressure chamber (more details are available in Acevedo-Opazo et al. (2010) for this data set) (Figure 3.2). The measurements realized in 2003 were used to ensure the temporal consistency of the temporal in-season evolution of Ψ_{PD} for the synthetic vineyards described hereafter. (Section 3.2.3.3).

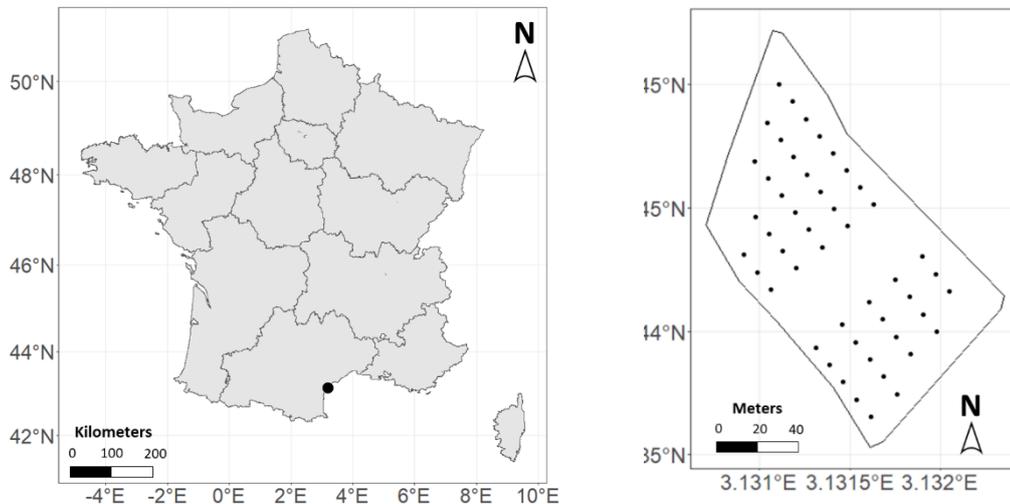


Figure 3.2 Location of the experimental Syrah vineyard at INRAE Pech Rouge (Gruissan, Aude, France) (left) and the distribution of the 49 sampling points within the vineyard, where each point represents the location of measured predawn leaf water potential (Ψ_{PD}) at multiple dates during the season.

3.2.3.2.2. WaLIS and predawn leaf water potential (Ψ_{PD})

The crop model used was a predictive model of vine water stress: Water baLance for Intercropped Systems – WaLIS (Celette et al., 2010). WaLIS was able to model the Ψ_{PD} , which was the agronomic variable of interest for this study. For the WaLIS modeling, the 2003 and 2004 weather data acquired through the weather station 11170004 (Gruissan) of the INRAE network via the Climatik application were used as inputs (Figure 3.3). These data included the daily mean temperature (T_{mean}), the daily precipitation (P) and daily evapotranspiration (ET). The latter was computed using the Penman-Monteith equation (Allen et al., 1989; Pereira et al., 1999). Note that WaLIS actually simulates the fraction of transpirable soil water (FTSW) and, by using a conversion, the FTSW is then transformed into Ψ_{PD} (Eq. 3.1) (refer to Lebon et al. (2003) for further details). Therefore, a realistic Ψ_{PD} minimum had to be defined because this conversion contained a logarithmic expression and the resulting FTSW value could be equal to 0. This minimum was assumed to the lowest value observed within the 2003-2004 dataset, which was -1.1 MPa. All measured dates in 2003 were used to calibrate WaLIS and the prediction of Ψ_{PD} was carried out in 2004 focusing primarily on only two dates when water stress in the vineyard is changing relatively rapidly. As the agronomic variable of interest is the vine water stress in a non-irrigated vineyard in this study, it is common to have no stress (uniform response) in the early to mid-early season period. This is then followed by a period where stress increases (mid to mid-late season), and the rate of increase may be spatially variable, until the entire vineyard reaches a constant level of high water stress and again exhibits a fairly uniform response. It was not relevant to predict in

times of uniform stress (either low and high), hence the focus on the two mid to mid-late dates (July-August) in this study.

$$\psi_{PD.mod} = \frac{\log(FTSW) - \log(C_a)}{C_b} \quad (3.1)$$

where C_a is a constant equal to 1.0572 and C_b is a constant equal to 5.3452 (from Lebon et al., 2003).

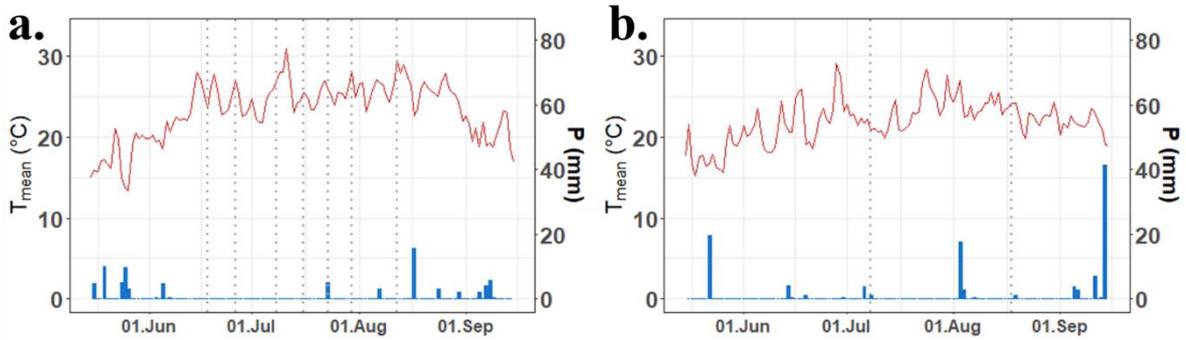


Figure 3.3 Weather conditions for the production season for (a.) 2003 and (b.) 2004 with the red line and the blue bars corresponding respectively to the mean temperature (T_{mean}) and the daily precipitation (P). The dashed gray lines correspond to the dates of measurement of Ψ_{PD} considered for this study in 2003 (June 18th, June 26th, July 8th, July 16th, July 23rd, July 30th and August 12th) and 2004 (July 8th and August 18th).

3.2.3.3. Data generation

3.2.3.3.1. Spatially rearranged field simulation

Spatially rearranged fields (SRFs) are defined in this study as synthetic data obtained by spatially rearranging existing data, i.e. forming new virtual vineyards. All the SRFs were derived from the 2003 observed data. The geographic locations of the sampling points was retained but the observed values (for a given date) were redistributed to different sampling points to achieve different spatial structures and patterns. The redistributed fields, i.e. SRFs, were then considered as reference data to evaluate the spatialized crop model performances following the process described in Figure 3.1. To calibrate the WaLIS spatialized crop model, Ψ_{PD} data from several dates were necessary. In non-irrigated vineyards, the spatial pattern of Ψ_{PD} is known to show a temporal consistency over time as water stress increases (Acevedo-Opazo et al., 2010), and all the Ψ_{PD} point measurements realized in 2003 exhibited such a temporal consistency and a distinct Ψ_{PD} evolution. Therefore, in order to have a consistent simulation and a realistic modeling from WaLIS, it was necessary to generate the same temporal consistency in the SRFs, i.e. for a given set of SRFs that represented a given year used for the calibration process. Thus, the Ψ_{PD} data needed to have a coherent temporal variability. To achieve this, each unique measurement site had a distinctive temporal profile in Ψ_{PD} that represented the observed temporal distribution in 2003. When a value was shifted to a different sampling point in a SRF, its temporal profile was also shifted, i.e. each measurement site kept its own temporal profile. Thus, the evolution of Ψ_{PD} value measured at a specific location was considered independent from the evolution of Ψ_{PD} value measured in its neighborhood to allow this spatial rearrangement.

The SRF characteristics were constrained by the selected evaluation dates, especially with the spatial structure of Ψ_{PD} (SS_{Ψ}) and the total variance of Ψ_{PD} (σ_{Ψ}). In the evaluation year (2004), two dates were selected for the study according to their SS_{Ψ} and σ_{Ψ} : July 8th 2004 and August 18th 2004 (Table 3.1). These dates were originally relatively spatially structured compared to the other measured dates, so rearrangement of measured sites could be done more easily to simulate different levels of SS_{Ψ} . Note that in the following study, resulting synthetic vineyards from July 8th 2004 and August 18th 2004 were respectively referenced as low σ_{Ψ} and high σ_{Ψ} .

Table 3.1 Characteristics of the predawn leaf water potential (Ψ_{PD}) measured for possible evaluation dates related to their variability and spatial structure. The dates in bold correspond to the considered dates kept for the study having desired coupled characteristics. σ_{Ψ} = standard deviation of Ψ_{PD} (characterizing the variability), $C_i\Psi$ = Cambardella index of Ψ_{PD} (summarizing the variance spatially organized and the spatial structure).

Date	σ_{Ψ} (MPa)	σ_{Ψ} characterization	$C_{i,\Psi}$ (%)	$C_{i,\Psi}$ characterization
09.06.2004	0.05	low	25	moderate
08.07.2004	0.10	low	51	moderate
05.08.2004	0.08	low	98	weak
18.08.2004	0.16	high	32	moderate
23.08.2004	0.15	high	37	moderate
10.09.2004	0.21	high	60	moderate

For this study, a choice of three different modalities to characterize the SS_{Ψ} was made based on the Cambardella index (C_i) (Cambardella et al., 1994). These were strong (S) ($C_i < 20\%$), moderate (M) ($40\% < C_i < 60\%$) and weak (W) ($80\% < C_i$) SS_{Ψ} . To model the different SS_{Ψ} , a reference spatial structure with known parameters was simulated using a theoretical variable. This theoretical variable was generated using Gaussian random fields (GRFs) on theoretical fields of size 500 x 500 pixels to ensure the targeted spatial structures were realized and that the theoretical variogram approximated the experimental variogram. To ensure the temporal consistency, the same number of units was needed for the synthetic vineyards, thus, new 49-unit GRFs were generated (corresponding to the 49 initial measurement sites) with the same field shape as the real vineyard field. These new GRFs were selected to approximate the targeted spatial structure (i.e. defined by the 500x500 GRFs) to ensure the spatial structure characteristic conservation of each considered SS_{Ψ} modalities (S, M and W). Every spatial structure was estimated by theoretical variograms fitted with a REML method, more suitable when limited data are available to compute variograms (Kerry and Oliver, 2007). Thus, when both theoretical variograms of targeted GRFs and 49-unit GRFs were close (i.e. with a similar C_i and a weak RMSE computed between each lag), spatial structures were assumed to be similar, i.e. the generated 49-unit GRFs were considered to have the desired SS_{Ψ} modality and the considered 49-unit GRF was kept. Then, the same approach was used to reassign the Ψ_{PD} values from the real data set (Figure 3.4) from the evaluation year to the 49-unit GRFs. Reassignments were based on value ranking between the 49-unit GRF and the real 49 measurement sites. This reassignment ensured that the spatial structure/pattern characteristics were kept by sorting the values and attributing them according to their rank, i.e. real 49 measurement sites values were orderly reassigned to the 49-unit GRF values (this method aimed to conserve a similar spatial structure) (part (i) Figure 3.1). Temporal consistency was ensured because the reassigned sites kept their Ψ_{PD} evolving through the calibration data set. Finally, there was 120 SRFs simulated (40 per SS_{Ψ}).

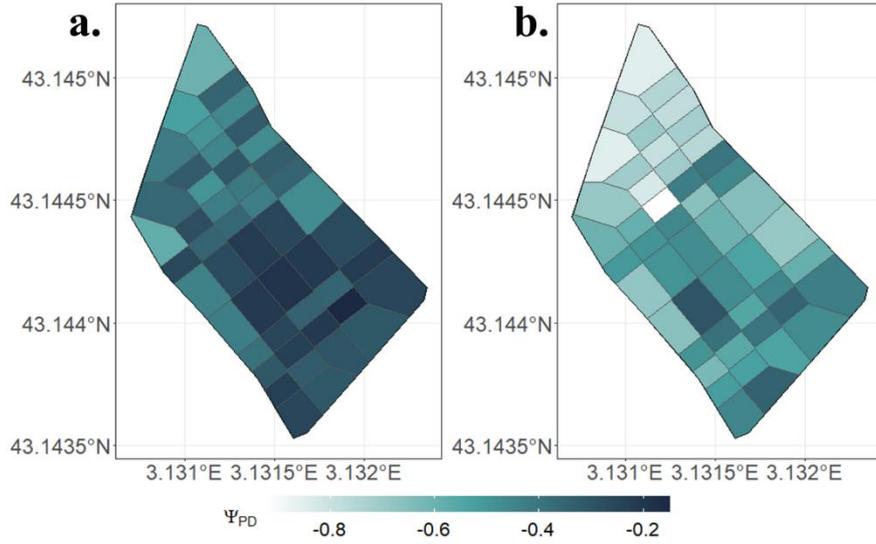


Figure 3.4 Predawn leaf water potential (Ψ_{PD}) measured on the experimental vineyard on (a.) July 8th 2004 corresponding to the date with low total variance of Ψ_{PD} (σ_{Ψ}) and (b.) August 18th 2004 corresponding to the date with high σ_{Ψ} . Voronoi tessellation was used to convert each of the 49 measurement sites of Ψ_{PD} into polygons.

3.2.3.3.2. Ancillary data simulation

Ancillary data were also simulated for each generated SRF at three different levels of correlation with the corresponding Ψ_{PD} (Q_{AD}) at the evaluation dates (part (ii) Figure 3.1). As a result, these ancillary data were spatially correlated with the SRF. To be able to modify Q_{AD} , the method from Oger et al. (2021) was adapted to simulate the ancillary data as needed for this study (Eq. 3.2 and 3.3)

$$AD_i = \psi_{PD_i} + \varepsilon_i \quad \text{with} \quad \varepsilon_i \sim N(0, \sigma_{AD}^2) \quad (3.2)$$

$$\sigma_{AD}^2 = \frac{\sigma_{\psi}^2}{(Cor(\psi_{PD}, AD))^2} - \sigma_{\psi}^2 \quad \text{with} \quad Cor(\psi_{PD}, AD) \in [0, 1] \quad (3.3)$$

where AD_i is the ancillary data value at the location i , ψ_{PD_i} is the Ψ_{PD} value at the location i , σ_{AD}^2 is the variance of simulated ancillary data, σ_{ψ}^2 is the variance of predawn leaf water potential reassigned and $Cor(\Psi_{PD}, AD)$ is the Pearson correlation between both variables.

In other words, with the assumption of second-order stationarity, the variance of Ψ_{PD} and ancillary data corresponded to the sill values of the theoretical variogram. Thus, depending on the Q_{AD} desired, the variance to be added to the Ψ_{PD} values was determined by Eq. 3.2. The correlation levels considered were 94%, 50% and 10% for Q_{AD} within the SRFs. These three levels were selected based on correlations observed in the literature between normalized difference vegetation index and vine physiology parameters (Bramley et al., 2019; Hall et al., 2011) and to provide strongly contrasting levels.

3.2.3.4. Spatial calibration approach

The objective of this study is to identify in which situations the use of a spatial calibration approach is likely to be more relevant compared to classical calibration. The intent was to model an agronomic variable at the within-field scale, and to do so, within-field calibration zones derived from ancillary data were considered. The results of the spatial calibration approach were compared to the results of a classical calibration approach corresponding to the site scale modeling (where all sites are individually calibrated) and the field scale modeling (where a unique value is computed over the whole field) (part (v) Figure 3.1). Although, the classical calibration approach was applied to the site-scale modeling, the calibration was independent of the ancillary data, so there was no spatial pattern of the agronomic

variable explicitly conserved through the modeling (i.e. sites are considered as independent modeling units).

3.2.3.4.1. Segmentation

Within-field zones were delineated using a segmentation algorithm (Pedroso et al., 2010) included in the *GeoFIS* R package (Guillaume & Lablée, 2022) applied to the synthetic ancillary data for each generated SRF (part (iii) Figure 3.1). The considered number of within-field zones were between 2 and 5 zones. Concerning the Q_{AD} , the hypothesis was that the higher the level of correlation of the ancillary data with the Ψ_{PD} , the more relevant the segmentation will be. Thus, the relevance of the segmentation was evaluated by a one-way analysis of variance (ANOVA) linking Ψ_{PD} with the delineated zones. The more the delineation explained the variation in Ψ_{PD} , the more the segmentation was considered relevant. Therefore, SRFs were kept depending on the delineation relevance, i.e. based on the proportion of variability explained by the delineated within-field zones. For the modeling scales higher than the measurement site scale, the spatial calibration was performed by an aggregation of the observed Ψ_{PD} within each zone (based on a mean of the constituting sites of each zone) from the rearrangement of the measured Ψ_{PD} .

Thus, considering the 120 SRFs, the 3 modalities of Q_{AD} , the 7 levels of spatial modeling scales and the classical calibration scales (i.e. field and site scale), there were 2,520 simulations for the whole study.

3.2.3.4.2. Spatial calibration of WaLIS

The WaLIS parameters identified as likely to have some spatial variation were the total transpirable soil water (TTSW) and the maximum crop coefficient of the vine (KC) (McClymont et al., 2019; Verdugo-Vásquez et al., 2022). Other WaLIS parameters were aspatially defined and their values were kept constant at a level representative of the Mediterranean context from where the original data were measured (part (iv) Figure 3.1). The TTSW and KC were spatially calibrated by finding the optimal parameter values on a 2-dimension grid using the measured data from Ψ_{PD} 2003. The TTSW values ranged from 55 to 210 mm in increments of 5 mm and the KC values ranged from 0.35 to 0.5 in increments of 0.05. Retained values were those that minimized the mean absolute error (MAE) compared to all Ψ_{PD} measurements realized in 2003.

3.2.3.5. Modeling performance evaluation

For all the modeling, outputs were disaggregated to the site scale level to assess the relevance of the spatial calibration. To evaluate outputs of spatialized WaLIS, the root mean square error (RMSE) (Eq. 3.4), the relative root mean square error (RRMSE) (Eq. 3.5) and the spatial balanced accuracy (SBA) (Eq. 3.6) (Pasquel et al., 2023b) were used (part (iv) Figure 3.1). The RMSE and the RRMSE were used to complement the SBA score in order to have a quantified numerical error. The SBA score was used to account for the preservation of the spatial pattern of the outputs. It provides an assessment of both aspatial and spatial pattern errors. Thus, SBA was used on each spatial modeling scale.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - M_i)^2} \quad (3.4)$$

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - M_i)^2}}{\bar{O}} \quad (3.5)$$

where O_i is the observed Ψ_{PD} , M_i is the corresponding modeled Ψ_{PD} , n is the number of observation ($n = 49$) and \bar{O} is the average of observed Ψ_{PD} .

$$SBA(O, M) = \frac{1}{100} \sum_{q=1}^{100} [1 - BA(O_{t(O,M,q)}, M_{t(O,M,q)})] \quad (3.6)$$

where O and M are respectively the observed and modeled maps, $O_{t(O,M,q)}$ and $M_{t(O,M,q)}$ are respectively the observed and modeled maps at the threshold level $t(O,M,q)$ corresponding to percentile q on the merging data distributions of O and M .

The SBA score was used to identify the best modeling scale from the spatial calibration approach, i.e. calibration at within-field zonal scales (RMSE was used to compared results obtained with SBA score). The best performed spatial modeling scale (n-zones), for each synthetic vineyard, was then used to compared and evaluate the relevance of using a spatial calibration compared to a classical calibration approach at either the field or site scale. The SBA score is currently designed to perform a comparison of spatialized modeling processes within the same field. It is not designed to compare a spatial modeling scale between fields (Pasquel et al., 2023b). However, the rearrangement of the sites in the generation of the SRFs meant that there were effectively many different fields (even though they had the same shape and the same number of observations). Therefore, it was not possible to compare between different SRF simulations, but only to different spatial scales within an individual SRF. To overcome this limitation and to avoid misinterpretation, the difference in the SBA score (ΔSBA) was used to compared results across the SRFs. The ΔSBA allowed an assessment of the relevance of a spatial calibration approach by comparing the results from a classical calibration at either the site scale (ΔSBA_{site}) (Eq. 3.7) or the field scale (ΔSBA_{field}) (Eq. 3.8) against the best performed spatial calibration scale. If the ΔSBA was positive, the spatial calibration approach had a better performance, i.e. a closer numerical value with a closer spatial pattern, than the classical calibration approach and vice versa if ΔSBA was negative.

$$\Delta SBA_{site} = SBA(O, M_{site}) - \min_{z \in \{2,3,4,5\}} [SBA(O, M_{zone})] \quad (3.7)$$

where M_{site} is the Ψ_{PD} modeled using a classical calibration approach at the site scale and M_{zone} is the Ψ_{PD} modeled using a spatial calibration approach for each considered within-field zones.

$$\Delta SBA_{field} = SBA(O, M_{field}) - \min_{z \in \{2,3,4,5\}} [SBA(O, M_{zone})] \quad (3.8)$$

where M_{field} is the Ψ_{PD} modeled using a classical calibration approach at the field scale and M_{zone} is the Ψ_{PD} modeled using a spatial calibration approach for each considered within-field zones.

3.2.3.6. Statistical analysis

All statistical analysis were performed using the R software (R Core Team, 2022). Multiple-way ANOVA was used to evaluate the significance of SS_{Ψ} , σ_{Ψ} and Q_{AD} in the determination of ΔSBA . Before applying ANOVA, residual normality and homoscedasticity were tested using the Shapiro test and the Levene test respectively. A post hoc test, realized using a Tukey's Honest Significant Difference, was used to identify significant pairwise differences among the modalities and to identify which groups were different from the others.

3.2.4. Results

3.2.4.1. Performance of WaLIS in predicting Ψ_{PD}

WaLIS modeled Ψ_{PD} with its own intrinsic error directly related to crop model assumptions. For the date with high σ_{Ψ} , the mean of the different RMSEs obtained for each spatial modeling scale and each Q_{AD} , was equal to 0.16 MPa (Figure 3.5). The spatialized WaLIS performances were dependent of the segmentation level and the correlation between the ancillary data and the Ψ_{PD} (refer to the Supplementary Figure S.A1, S.A2 and S.A3 to see the WaLIS performances for each level of segmentation and each level of correlation for the considered virtual vineyards in Figure 3.5).

WaLIS performances were similar for both modeling dates (high and low σ_{Ψ}). Note that for all modeling results from the ‘classical’ model calibration, the field scale modeling RMSEs were always equal (as the average is always the same). Similarly, the site-scale modeling RMSEs were always equal regardless of the spatial arrangement of the sites as the temporal consistency was always constant. Indeed, only the measurement site locations were different between each of them (modification of the SS_{Ψ}), and calibrations were thus identical and consequently the modeled Ψ_{PD} as well. The RMSE at the within-field scales (2-5 zones) were different depending on the modeling because, depending on the spatial arrangement of the measurement site, the calibration zoning was different and the resulting calibrations differed leading to a differently modeled Ψ_{PD} . Evaluation using RMSE allowed an estimation of numerical errors in the spatialized WaLIS prediction. However, RMSE only assessed the aspatial error but was not able to assess the spatial pattern error.

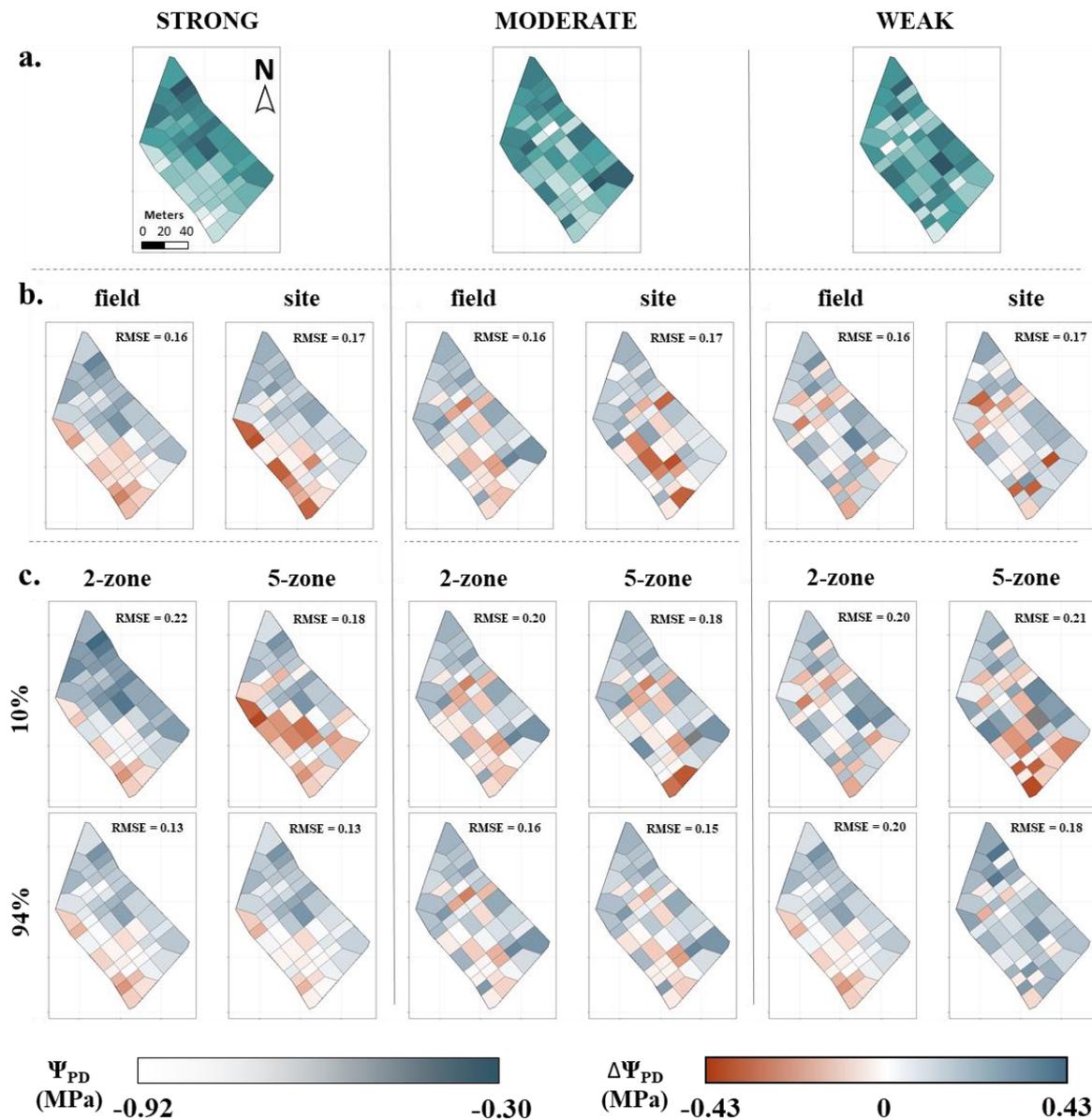


Figure 3.5 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) regarding each considered spatial structure of Ψ_{PD} (strong, moderate and weak) between the (a.) observed (synthetic) Ψ_{PD} and modeling at different spatial scales: (b.) field and site scale, (c.) at within-field scale (2-zone and 5-zone scales) for different level of correlation between Ψ_{PD} and the synthetic ancillary data (10% and 94%). The presented modeling date were based on August 18th with a high Ψ_{PD} variability. Results are only shown for one synthetic vineyard for each spatial structure modality.

3.2.4.2. Relevance of using a spatial calibration approach

In Figure 3.6, the more positive the ΔSBA value, the more the spatial calibration approach was relevant compared to a classical calibration at field or site scale. The ΔSBA dispersion was different for both dates (high and low σ_Ψ). The variability of ΔSBA was higher for the date with a high σ_Ψ (August 18th). For both dates, ΔSBA were higher when Q_{AD} was higher, except for simulations that generated weak spatial structures where the results were less constrained. Thus, in all cases, increasing Q_{AD} increased the relevance of using a spatial calibration approach. The violin plots (Figure 3.6) showed that most of the ΔSBA values were relatively well grouped for all combinations of spatial structure and ancillary data correlations, but most did present some outliers. The results of Tukey's test (indicated by a letter on the top of the Figure 3.6) identify any significance differences between different combinations of spatial structure and ancillary data correlation.

For a high σ_Ψ , comparisons between using spatial calibration or classical calibration at the site scale, i.e. ΔSBA_{site} results, showed that spatial calibration was significantly relevant when SS_Ψ was strong with Q_{AD} at 94% (Figure 3.7) but that the relevance of using spatial calibration compared to the site scale calibration decreased with a decrease in Q_{AD} . The 50% and 10% approaches were equivalent and had a tendency to have a deteriorated performance when spatial calibration was used. For moderate or weak SS_Ψ , the ΔSBA_{site} showed that it was not significantly relevant to use a spatial calibration approach. The comparison between using spatial calibration or classical calibration at the field scale, i.e. the ΔSBA_{field} results, showed that the use of spatial calibration was significantly relevant when SS_Ψ was strong or moderate with Q_{AD} at 94%. When SS_Ψ was strong with Q_{AD} at 50%, there was a tendency to improve the modeling performance using spatial calibration, but this was not significant. When SS_Ψ was moderate with Q_{AD} at 50% or 10%, there was a tendency to deteriorate the modeling performance using spatial calibration. For a weak SS_Ψ , using spatial calibration or classical calibration at field scale approaches were equal, except with the Q_{AD} at 10% where a tendency for the spatial calibration to deteriorate modeling performances was observed.

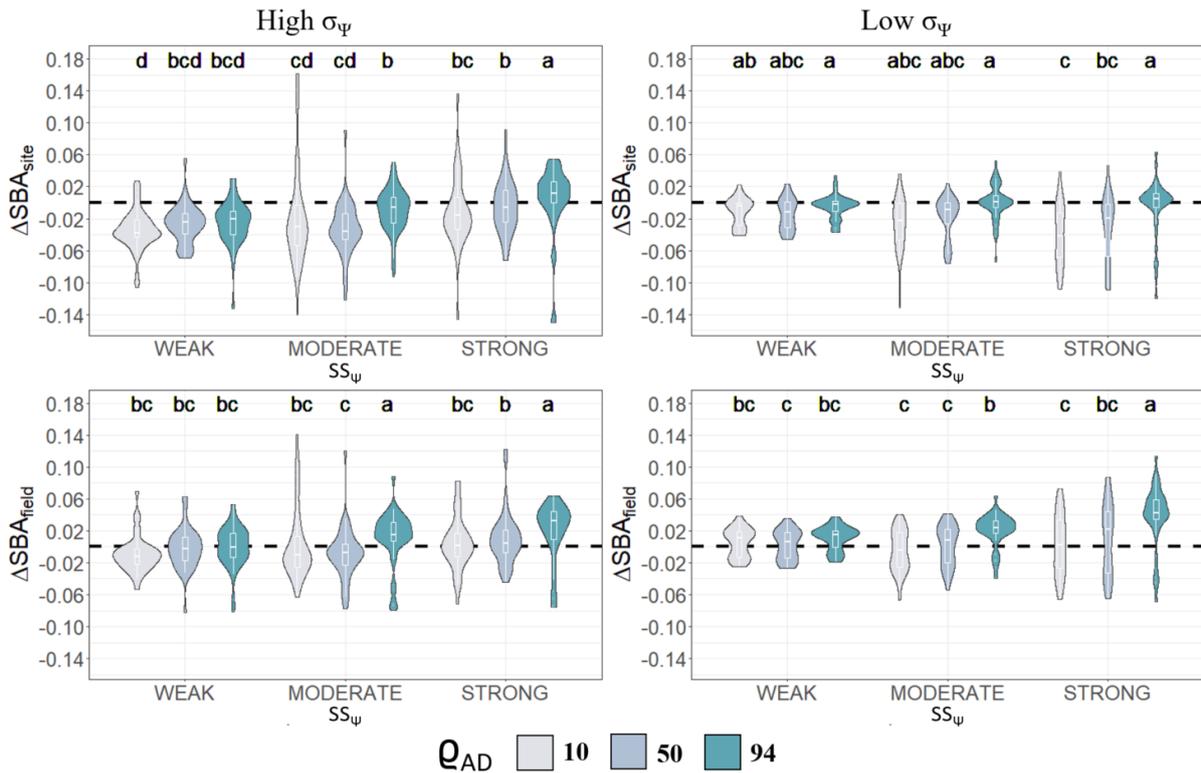


Figure 3.6 Change in the SBA distributions using a spatial calibration approach compared to a classical calibration approach (at field scale at the bottom and at site scale on top) depending on the spatial structure of the agronomic variable (SS_{Ψ}) (strong, moderate and weak spatial structure), the variability of the agronomic variable (σ_{Ψ}) (high variability on right and low variability on left) and the correlation between the agronomic variable and the ancillary data (ρ_{AD}) (Pearson’s correlation of 10%, 50% and 94%). A positive ΔSBA corresponds to better results with a spatial calibration approach, respectively a negative ΔSBA corresponds to a better result with classical calibration approach. Letters on the top are based on Tukey’s test of significance. Modalities with the same letters are not significantly different at 0.05 probability level.

For a low σ_{Ψ} , the ΔSBA_{site} results showed that using spatial calibration was significantly relevant when SS_{Ψ} was strong with ρ_{AD} at 94% (Figure 3.7). For ρ_{AD} at 50% and 10% with a strong SS_{Ψ} , using spatial calibration had a tendency to deteriorate modeling performance, which was also observed for a moderate SS_{Ψ} with ρ_{AD} at 10%. For other modalities, the ΔSBA_{site} showed that using spatial calibration or classical calibration at site scale approaches led to equal performances. Regarding the ΔSBA_{field} results when σ_{Ψ} was low, the results showed that spatial calibration was significantly relevant when SS_{Ψ} was strong and moderate with ρ_{AD} at 94%. For all other modalities, the ΔSBA_{field} showed that spatial calibration or classical calibration at field scale approaches led to equal or better performances. Overall, for conditions with a low σ_{Ψ} , the results were less contrasted than with a high σ_{Ψ} .

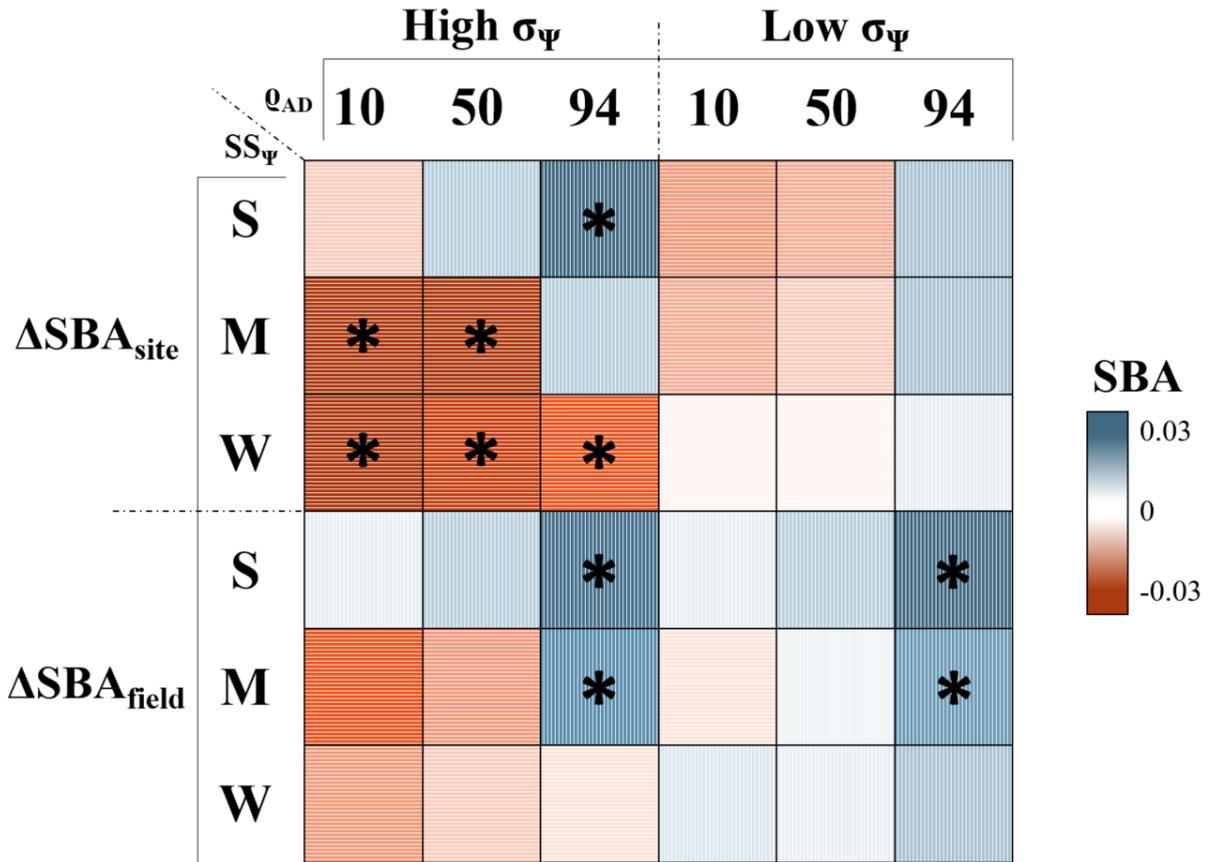


Figure 3.7 Relevance of using a spatial calibration compared to a classical calibration regarding the difference of SBA score for different spatial scales of modelling. Significances were estimated with a Tukey test (modalities with different letters) and correspond to a significant improvement or deterioration of the spatialized crop model (represented by *). Blue cases correspond to a better performance of a spatial calibration compared to a classical calibration. Red cases correspond to a better performance of a classical calibration compared to a spatial calibration. σ_Ψ corresponds to the variability of Ψ_{PD} , S, M and W correspond to the spatial structure modalities SS_Ψ (respectively strong, moderate and weak), 10, 50 and 94 correspond to the ancillary data correlation q_{AD} (respectively 10%, 50% and 94%).

3.2.4.3. Assessment of aspatial and spatial errors

Results obtained using ΔSBA were compared with results obtained using $\Delta RRMSE$. Only a comparison for ΔSBA_{site} and $\Delta RRMSE$ computed between spatial calibration and classical calibration at the site scale approaches are shown here. This choice was made to highlight the situation of using a classical calibration at the site scale instead of a spatial calibration to represent the spatial pattern of Ψ_{PD} at within-field scales. However, in some cases, the ΔSBA_{site} showed that a spatial calibration could improve the spatialized crop model performances especially when σ_Ψ was high (Figure 3.7).

Compared to ΔSBA_{site} (Figure 3.7), $\Delta RRMSE$ showed more contrasted results (Figure 3.8). When the correlation with ancillary data was high (q_{AD} at 94%), spatial calibration was significantly better performed for all levels of spatial structure ($\Delta RRMSE > 0$; blue coloring). However, for the other two levels of q_{AD} (10% and 50%), spatial calibration was significantly worse than classical calibration for all levels of spatial structure ($\Delta RRMSE < 0$; red coloring). The $\Delta RRMSE$ results were more sensitive to q_{AD} than SS_Ψ in this study.

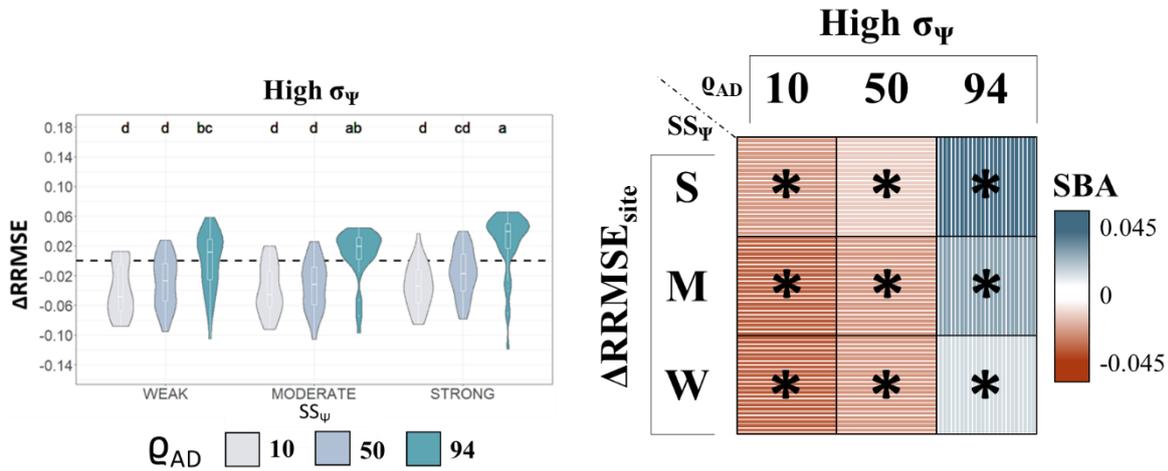


Figure 3.8 Relevance of using a spatial calibration compared to a classical calibration regarding the difference of RRMSE for different spatial scales of modeling for a date of high σ_Ψ (August 18th 2004). Significances were estimated with a Tukey test (modalities with different letters) and refer to a significant improvement or deterioration of the spatialized crop model performances (represented by *). σ_Ψ corresponds to the variability of Ψ_{PD} , S, M and W correspond to the spatial structure modalities SS_Ψ (respectively strong, moderate and weak), 10, 50 and 94 correspond to the ancillary data correlation Q_{AD} (respectively 10%, 50% and 94%).

3.2.5. Discussion

3.2.5.1. Cases when spatial calibration should be preferred compared to classical calibration approach

The relevance of using a spatial calibration approach compared to either classical calibration approaches (whether at the field or site scale) decreased with a decreasing Q_{AD} . The spatial calibration was identified as significantly relevant when Q_{AD} was high (equal to 94%) for both dates, i.e. regardless of whether the σ_Ψ was high or low. When Q_{AD} was equal to 94%, the spatial pattern of both the Ψ_{PD} and ancillary data were very similar. Consequently, the segmentation of the ancillary data into calibration zones faithfully reproduced the spatial pattern of the target agronomic variable. Thus, it was unsurprising that spatial calibration was able to improve the calibrated model performances. However, when Q_{AD} dropped to 50%, which is a value that is often observed in spatial agri-datasets, results were more contrasted and the degree of variability in the target variable became important as to whether the spatial calibration was more relevant or not. With a lower level of correlation, it appeared that a greater variance in the target variable was needed to justify a spatial calibration approach. When there was very little correlation between the ancillary data and the target variable ($Q_{AD} = 10\%$) the segmentation method could not define relevant calibration zones and the spatial calibration generated worse results than a classical calibration approach at both the field and site scale.

For modeling under conditions of high σ_Ψ , it was more complicated to know when the use of a spatial calibration approach was relevant. Indeed, even with a high σ_Ψ , there were simulations with a significant deterioration of prediction quality. For example, when SS_Ψ was weak or moderate, using spatial calibration significantly deteriorated modeling performances compared to classical calibration at the site scale, even with a high σ_Ψ . Without a strong SS_Ψ , the segmentation of the ancillary data was generally unable to accurately reproduce the spatial pattern of the target variable. Therefore, incorrect calibration zones were defined that added more error into the modeling compared to a non-spatial calibration approach. These results indicated that for moderate and weak SS_Ψ , a spatial calibration approach should be avoided. When the target variable exhibited a strong spatial structure, the spatial calibration improved performances compared to a classical calibration either at field or site scales, but

the significance of the improvement decreased with decreasing Q_{AD} . In such cases, spatial calibration improved modeling performances because it corrected errors that could occur when applying classical calibration at site scale because it better reproduced the spatial within-field variance.

For modeling with a low σ_{Ψ} , the results from the spatial calibration and classical calibration approaches were almost equivalent. The low σ_{Ψ} resulted in simulated Ψ_{PD} values that were close together, generating very little differentiation between the calibration processes and consequently resulting in the same (or similar) modeled Ψ_{PD} values. This was particularly evident for moderate to low Q_{AD} where modeling performance was relatively poor. Indeed, for low σ_{Ψ} , the segmentation of calibration zones is less relevant for the spatial calibration. However, when considering a low σ_{Ψ} , using either a spatial calibration or a classical calibration resulted in the same quality of modeling.

Nevertheless, when performances of spatial calibration and classical calibration at the site scale were equal, the spatial calibration approach may still be preferred because it allows the reproduction of the spatial pattern of the agronomic variable to some degree. Spatial calibration, relative to site-scale classical calibration, also reduces the number of calibrations/parametrizations to be achieved to reproduce the spatial patterns of the agronomic variable. This is likely to be more relevant when using complex crop models, such as APSIM or DSSAT (Hoogenboom et al., 2019; Jones et al., 2003) that have many more parameters to estimate when compared to simpler crop models, such as WaLIS. A spatial calibration approach, based on calibration zones, keeps the spatial pattern but reduces the number of calibrations needed in complex situations and thus will reduce the effort and time needed in the calibration processes. As it also involves some form of data aggregation, calibration zones could also, in specific cases, reduce the potential uncertainties resulting from the calibration process (compared to a site-by-site classical calibration approach).

Finally, spatial calibration showed globally better results than the classical calibration at the field scale and better results, on some identified combinations, than the classical calibration at the site scale. Applying spatial calibration by applying a spatial constraint via segmentation into the calibration process, allowed the calibration process to correct a part of the calibration error that occurred when ignoring the spatial pattern of the agronomic variable. By reproducing the spatial pattern of the agronomic variable, the aim was to smooth the calibration error and thus reduce the uncertainty of the prediction. Typically, the goal was to reduce the part of calibration uncertainty coming from poor quality data observed at high resolution. However, using a spatialized crop modeling approach will still not be able to correct any errors originating from the model itself, because spatial calibration does not alter the internal structure (equations) within the crop model.

3.2.5.2. Spatial calibration performances using a spatial pattern based metric

Assessment of the spatial calibration approach using the RRMSE was more contrasting than using the SBA score. Results were only shown for a comparison between spatial calibration and classical calibration at the site scale (Figure 3.7 vs 3.8). Classical calibration at the site scale is intuitively the natural method when reproducing spatial patterns of the agronomic variable is the objective, i.e. the highest available resolution should be preferred. Thus, it was of primary interest to know if this was true and when and to what degree of downscaling was relevant for a spatial calibration approach. When comparing results from the RRMSE and SBA scores, it was clear that these two metrics were not converging regarding the weak SS_{Ψ} scenarios. The RRMSE showed that using spatial calibration with Q_{AD} at 94% improved performances for all considered SS_{Ψ} while, the SBA showed that for weak SS_{Ψ} , the use of spatial calibration was not the best approach because spatial patterns could be forced depending on the considered agronomic variable. The RRMSE also showed that using spatial calibration on strong SS_{Ψ} significantly deteriorated performances with Q_{AD} values of 50% and 10%, although the SBA analysis showed equivalent results for Q_{AD} at 50% and a tendency to deterioration at 10%. Although the SBA score was unable to numerically quantify the error, its use has been shown to be relevant to assess the spatial pattern error in the evaluation of spatialized crop model outputs (Pasquel et al., 2022a).

3.2.5.3. Spatial calibration dependence regarding the segmentation of the within-field zones

In this study, the spatial calibration approach was based on using calibration zones as a means of downscaling the crop model. The simulations certainly revealed that there was a huge dependency on the considered ancillary data and the segmentation algorithm. However, in this preliminary study, it was difficult to know which of these factors had the strongest impact on the determination of the relevance of spatial calibration. Ancillary data introduced uncertainty depending on its level of correlation with the considered agronomic variable and in a real case scenario will also have uncertainty associated with its acquisition. There are many different ways to segment or classify agri-data into classes or zones, and the relevance of any given approach will be linked to the type of data acquired and the objective of the zoning. However, each method will yield different outcomes and will have some inherent level of uncertainty in the resulting delineation.

Calibration zones are not the same as management zones. They should be derived specifically to support spatial calibration and the spatialization of the crop model. It is possible, and likely expected, that the spatialized output from the crop model could be used with other data to generate management zones, i.e. that calibration zones help to inform, but are not necessarily equal to, management zones. It is clear that further work on the best methodology to delineate calibration zones is needed. However, from this preliminary work, it appeared that having ancillary data that accurately reflected the target variable should be the first prerequisite when considering spatial calibration. If a relevant ancillary data for the target variable is available, then it is likely that the choice of classification/segmentation algorithm will be less important. Consequently, when the ancillary data and the target variable are less well correlated, the method of calibration zone delineation should become more important. However, this question of the relative incidence of the zoning approach versus the quality of ancillary data is still unclear and should be investigated in further research.

For this preliminary study, only a single ancillary data layer was chosen to simplify the assumptions and the simulation of the ancillary data used. In reality, for a relevant and reliable use of ancillary data, several types of data should be used to better describe the spatial pattern of the agronomic variable (Derby et al., 2007). Data fusion is necessary to integrate several sources of ancillary data to have a more accurate delineation of within-field zones (Castrignanò et al., 2019). The use of ancillary data with relatively stable spatial pattern characteristics (e.g. soil texture) coupled with variables with changing spatial pattern from year to year depending on weather conditions (e.g. yield maps, crop coverage) has been identified as possible alternative to fully integrate the spatial pattern of production at the within-field scale (Nawar et al., 2017). In the case of this study, having several ancillary data layers would have been preferable but would also have been very computationally challenging and could possibly have led to bias in the results interpretation, particularly if the level of autocorrelation within and between multiple synthetic ancillary datasets was not well controlled. This question of how to best define calibration zones using multiple datasets remains an open question to be addressed in future work.

The main goal of applying a spatial calibration approach was to reproduce the spatial pattern of the agronomic variable at the considered modeling date. However, spatial information of the agronomic variable at the beginning of the production season is not always possible. Moreover, spatial patterns, on which the calibration zones are assumed to be delineated, may be temporally dynamic for many agronomic traits. In other words, the spatial calibration approach was not designed to take in-season temporal aspects into consideration because calibration zones are fixed at the beginning of the modeling. The spatial calibration is a method with a similar objective to data assimilation processes, which are often advocated for downscaling a crop model (Jin et al., 2018). Data assimilation represents a wide variety of methods to correct or adjust the crop model parameters multiple times during the growing season based on observed in-season data (e.g. remotely-sensed data) to predict a crop model state variables. The data assimilation approach is mainly used for its accurate temporality to estimate crop growth status, but is limited as it cannot be used pre-season for predicting the upcoming season, like the spatial calibration method used here. Furthermore, applying a data assimilation method does not

guarantee the consistency of the agronomic variable spatial pattern, which can generate spatial issues to correctly model the agronomic variable. In other words, there is a potential to take advantage of both approaches, spatial calibration and data assimilation, to complement each other. Thus, the temporal characteristics of data assimilation and the spatial characteristics of spatial calibration may allow an in-season recalibration to better take into account the spatial pattern of the agronomic variable. The temporal consistency of the synthetic data was a huge issue in this study. It was a challenging objective to model an agronomic variable with both spatial and temporal consistency, and was achieved by constraining the simulations to the properties of existing real data. This has an intrinsic limitation as this relative small real data set may not be a correct representation of the actual population. A better accounting of spatio-temporal changes in the target agronomic variable could be an interesting source of improvement that should also be investigated in further research.

3.2.6. Conclusion

The study investigated the effect on a spatial calibration approach of the spatial structure characteristics of the considered agronomic variable at the evaluation date, the total variance of the same agronomic variable, and the level of correlation between this agronomic variable and the ancillary data used to delineate the within-field zones. Spatial calibration significantly improved modeling performances when the agronomic variable had a strong spatial structure, a high level of variability and when ancillary data were highly correlated. However, a special attention should be given to cases with moderate and weak spatial structure with high variability that did not benefit from spatial calibration in this example. When the variability of the target variable was low, there was no difference in the model performance by using a classical or spatial calibration approach. The modeling performances were assessed by both a classical (RRMSE) and a new spatial metric (SBA) to allow an evaluation of both aspatial and spatial pattern errors. As a preliminary study, it was noted that the results were strongly influenced by the synthetic ancillary data and, potentially, the segmentation algorithm used to generate the calibration zones. This was a preliminary study into the potential to use a spatial calibration process as a means of spatializing a simple crop model. The approach taken was to downscale predictions by defining within-field calibration zones. This work has generated new knowledge on the way to increase the resolution of the spatialized crop models at the within-field scale. The objective of this downscaling method is to improve cultural practices towards more sustainable ones in a precision agriculture context. Thus, while spatial calibration appears to be relevant in certain situations, considerably more research, particularly using sensitivity analysis related to key data characteristics is still need to better understand this novel approach to crop model downscaling using a spatially constrained calibration process. Data assimilation also represents a real, alternative approach to crop model spatialization and further research should also explore the potential for coupling spatial calibration with data assimilation approaches.

3.2.7. Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004. This work has been realized with the support of MESO@LR-Platform at the University of Montpellier.

3.3. General conclusion of the chapter

The purpose of the research work presented in this chapter was to investigate the relevance of using a spatial calibration approach to optimize the spatialized crop model performances. The spatial calibration proposed here was used as a downscaling method to spatialize WaLIS predictions. This method allowed to identify which modeling scale was the most relevant to apply WaLIS at the within-field scale on our study. The spatial calibration approach is presented as a downscaling method, i.e. as a way of using tools and data differently for modeling at the within-field scale. However, this downscaling method is relatively different from usually downscaling processes used to spatialize climatic models (e.g.

correction algorithms). In that study, using SBA score was appropriate because it involved comparing modeling scales for each simulated vineyard, which correspond to the intent of the SBA score. However, regarding the limits highlighted in Chapter 2, interpretation of SBA was completed with RRMSE to have a more common interpretation of the aspatial error within spatialized WaLIS outputs.

The results of this chapter revealed the relevance of applying a spatial calibration approach based on the specific characteristics of the considered agronomic variable such as its spatial structure/pattern, its variability, and its correlation with ancillary data used for delineating calibration zones at within-field scale. The results indicated that spatial calibration approach improved the crop model performances compared to a classical calibration approach (at field or site scale), especially when the agronomic variable exhibited a strong spatial structure with a high correlation with the simulated ancillary data. Using either a classical calibration approach or spatial calibration approach was in most cases equivalent (especially for agronomic variable with low variability). However, a specific attention should be paid when agronomic variable is highly variable and the spatial structure is moderate or weak. In this case, spatial calibration could significantly deteriorate crop model performance compared to a classical approach. As a result, a classical calibration approach should be preferred in such cases. Beyond the results related to relevance of using spatial calibration approach, the methodology to simulate data from already measured data represent a result too. Indeed, it was really challenging to simulate at the same time both desired spatial structure and a temporal consistency to ensure a realistic dataset. The temporal consistency constrained to use already measured data and assumed that the evolution of the temporal profile of each measured sites were spatially independent. Moreover, dataset for studies at the within-field scale for precision agriculture purposes can be time-consuming and expensive to realize. Thus, this method represents a potential to make profitable any dataset already measured to create new field(s) for demonstration without carrying other measurement campaigns.

In this work, only 49 points of measured data were available to spatially calibrated WaLIS at the within-field scale. Simulated ancillary data were computed on the same measurement locations, as it was the case for some ancillary data in the real data set to be the closer from a real study. Thus, limitations on the measured data for the calibration were also affecting the delineation of the within-field zones necessary for the spatial calibration. Indeed, this number of points can be a limitation to delineate reliable calibration zones. These ancillary data from the real case study corresponding to time-consuming measurements. Thus, this limitation was probably a brake to the augmentation of measured data allowing them to a better spatial calibration. High-resolution sensor data represent a real opportunity to manage this limit on the condition of having enough observations describing each calibration zone to be able to correct the individual measurement error of each sensor data. For that reason, number of delineated calibration zones considered for that study was between 2 and 5 zones. Indeed, regarding the number of available data for the calibration and the size of the vineyard (1.2 ha), more delineated zones could lead to more stochastic error. Increasing the number of calibration zones to 10 or 20 (has it was made in Chapter 4) would have tended to site scale calibration performances. For modalities with structured characteristics, this could deteriorate performances, as it was the case for classical calibration at the site scale. However, for modalities with moderate spatial structure, spatial modeling scale between 5-zone scale and site scale could have better performances (as the spatial pattern might be more fragmented). However, based on the simulated ancillary data, the segmentation algorithm was not able to make more than 5 zones. This assumption on spatialized crop model for these modalities were not possible and are thus only suppositions. Thus, the number of zones that appeared to make most sense should be minded on a trade-off between field size and number of measurements (data available for the within-field delineation and calibration).

This presented work highlighted an important point regarding the methodology concerning the use of ancillary data and delineated within-field zones. A selection of the most relevant ancillary data should be carefully made. Here, only one ancillary data was used due to consistency in simulating data set. However, using several ancillary data should be preferred for delineation of the whole field. This

approach could be beneficial, especially regarding the potential temporality of the generated within-field zones. Encompassing the spatiotemporal characteristic of the field variability is not a trivial process and is still a huge scientific domain of research for decades. Segmentation algorithm should also be carefully selected. Two important aspects of the selected algorithm are based on the consideration of the spatial characteristic to have a spatial consistency at within-field scale and its performance to have a reliable segmentation even with small data set or for a small field. Obtaining high-resolution data could allow to increase segmentation algorithm performances and increase spatial calibration relevance.

Moreover, the selected parameters identified to be spatially variable at the within-field scale also have a role to play in the relevance assessment of delineated zones. Indeed, an important question related to the spatialization of existing crop models is the reliability to use them at the within-field scale. The proposed spatial calibration method is based on within-field zones and the calibration of certain crop model parameters at these spatial zone scales. However, the relevance of these delineated zones is mainly based on the parameters considered as having a significant spatial impact, i.e. likely to vary at the within-field scale and having an impact on spatialized crop model outputs. More research should be led on the parameters considered as spatial parameters to spatialize crop models. This could be made by carrying sensibility analysis for instance to optimize the parameters selection and thus increase crop spatialized crop model performances.

Resolution of the measured data to calibrate the crop models could be coarser than the delineated zones that can lead to calibration inconsistencies. Increase of sensors and probes use at the within-field scale for precision agriculture will allow to obtain more data with a finer resolution to improve spatialized crop model performances. High-resolution sensors are usually affected by a higher stochastic error and the higher number of measurements that give a better representation of the considered variable spatial pattern often overcomes this drawback. However, depending on the within-field zone size, the stochastic error at the individual location of these sensors could significantly deteriorate the spatialized crop model performances if not enough sensors can be aggregated for the considered calibration zones. This could finally lead to an unbalanced performance according to the zones of calibration for a considered modeling scale.

Chapter 4

Simulation scale impact on the relevance of utilizing downscaled spatialized crop models at the within-field scale

4.1. Intention note

Mechanistic crop models are used by a large range of practitioners, from the agronomic community to the policy-makers, and are applied over a diverse range of applications. These applications range from technical applications directly linked with cultural practices, e.g. nutrient fertilization (Cornet et al., 2022; Goffart et al., 2008), irrigation scheduling (García-Vila and Fereres, 2012; Mairech et al., 2021) and pest management (Rasche and Taylor, 2019); to the prediction of climate change impact on major crops to select the best modes of adaptation to face this challenge across the world (Ewert et al., 2015; Rosenzweig et al., 2013). A diversity of potential uses has also led to a diversity in the development of crop model types, as there is no single model type that is suitable for all uses. Most of the currently used crop models were originally designed for crops of worldwide interest, especially cereal crops that are essential to feed the world's population (Silva and Giller, 2020). The importance of these models for scenario testing around food security issues and the interest in developing these crop models has deepened over the last decades. As a consequence of this, the crop modeling community has invested a lot of time and knowledge in improving several large crop models to make them more efficient over a variety of different key crop types and, as a result, usually more complex too. On the opposition, for many other crops or agricultural systems that are much more specific in their cultivation and agronomy (e.g. tuber or root crops, tropical crops, perennial crops), there has been a trend to develop more simplistic, crop-specific models. These less complex crop models are generally used in a more specific context, which is more focused on understanding a specific biophysical process associated with production, rather than being applied to understand the impact of climate change on crop production at worldwide scale for instance.

In this manuscript, a complex crop model corresponds to a crop model with several tens of parameters and several coupling modules. For instance, the STICS (Brisson et al., 2003, 2002, 1998), DSSAT (Hoogenboom et al., 2021, 2019; Jones et al., 2003) and APSIM (Holzworth et al., 2014) models are among these complex crop models. The complexity of these crop models demands high-quality input data to permit an accurate and reliable modeling performance (Pasley et al., 2023). For strategic applications based around hypothesis testing and scenario testing in potential futures, input data can be managed to achieve the desired level of input quality. However, if the objective is to model an agronomic variable at the field or the within-field scale for tactical, in-season management purposes, then the input data needs to be derived from the target agro-ecosystem. Obtaining high quality data from local, agricultural production systems is not always possible and the stochastic error associated with in-field observations may be problematic for model operation. Moreover, the complexity of these crop models also generates a complexity of interactions among the model equations that describe the bioprocesses related to the soil-plant-atmosphere. The general understanding of the effect of introducing 'noisy' data into these complex crop models is limited, especially for processes modeled at a finer scale than the native spatial footprint of the model (Porwollik et al., 2017). Improving the understanding of this effect

is of important interest for the agronomic and modeling communities that wish to adapt crop models for short-term, tactical, management uses.

Despite the general lack of knowledge of crop model performance at finer spatial scales with lower quality input data, there is a working assumption that the scientific knowledge embedded in these existing complex crop models, and the development and knowledge gained from running the models, could be taken advantage of by practitioners to support more efficient, site-specific, crop production practices. The hope is that crop models can be modified for production uses at different spatial scales without the need to fundamentally redeveloping new ‘spatial’ crop models. Following on from the research work presented in previous chapters, the consideration of the spatial patterns of the predicted agronomic variable seems to be of primary interest when using downscaling crop models for tactical, within-field decision-making. Therefore, a working assumption arising from the previous chapters is that a spatial calibration approach is also a relevant downscaling method for complex crop models to predict the spatial patterns of an agronomic variable of interest.

The work presented in this chapter aims to aggregate all the previous knowledge realized in the PhD project and to apply the concepts and methods detailed previously in this thesis on a real case study involving a complex crop model. Thus, this chapter is less theoretical than previous chapters. However, as the proposed methods are still under experimentation and to be improved. This this work was carried out during a scientific exchange at Aarhus University in Denmark (Tjele).

The following chapter present a research work to study the behavior of a complex crop model, modified by a downscaling method to simulate an agronomic variable at the within-field scale and to see if there is an opportunity to use the model predictions in precision agriculture. This work is based on real data from a durum wheat experimental field. The aim was to model the durum wheat yield at the within-field scale by using a spatial calibration approach as downscaling method on APSIM as a complex crop model.

4.2. Downscaling the APSIM crop model for simulation at the within-field scale

4.2.1. Details about the paper

4.2.1.1. Title and publication information

This section was published in *Agricultural Systems* as a scientific review in September 2023. The full citation is:

Pasquel, D., Cammarano, D., Roux, S., Castrignanò, A., Tisseyre, B., Rinaldi, M., Troccoli, A., Taylor, J.A., 2023. Downscaling the APSIM crop model for simulation at the within-field scale. Agric. Syst. 212, 103773. <https://doi.org/10.1016/j.agry.2023.103773>.

4.2.1.2. Authors

Pasquel D.^{1,*}, Cammarano D.², Roux S.³, Castrignanò A.⁴, Tisseyre B.¹, Rinaldi M.⁵, Troccoli A.⁵ and Taylor J.A.¹

¹ ITAP, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

² Department of Agroecology, iClimate, CBIO, Aarhus University, Tjele, Denmark

³ MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France

⁴ Department of Engineering and Geology (InGeo), Gabriele D'Annunzio University of Chieti-Pescara, Chieti, Italy

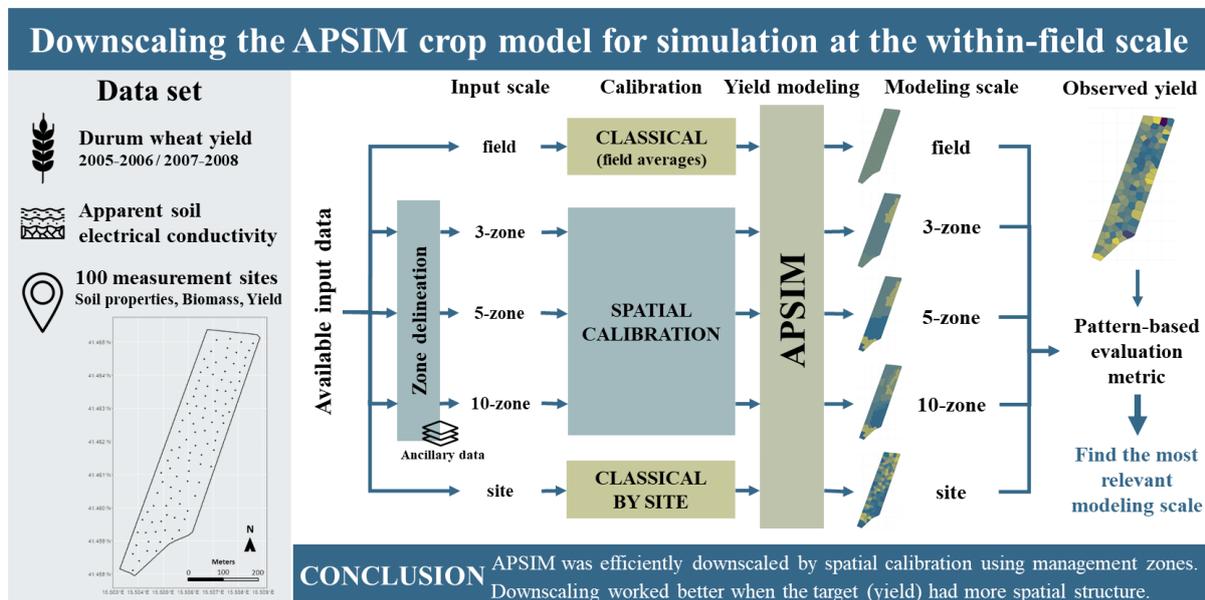
⁵ CREA Research Centre for Cereal and Industrial Crops, Foggia, Italy

* Corresponding author: daniel.pasquel@inrae.fr

4.2.1.3. Abstract

Most crop models are designed for point-based modeling and to simulate agronomic variables on their native spatial footprint, i.e. typically as a uniform field-scale value. Precision agriculture needs crop model simulations at sub-field scales to support differential management application. Spatialization processes are used to change the simulation scale of crop models. The objective of this study is to investigate the spatialization of a complex crop model by using a spatial calibration approach to modify its native spatial footprint and to evaluate if it is relevant to use this kind of crop model at the within-field scale. APSIM was spatialized to simulate durum wheat yield at different spatial scales (field, within-field and site-scale) on an experimental field under Mediterranean conditions in southern Italy. Ancillary soil data were used to derive potential management (modeling) zones at different scales, which were then used to spatially calibrate soil and biomass parameters in APSIM to spatially predict yield in two different production years (one year was used for calibration and the other for evaluation). Spatialized crop model performances were evaluated using the spatial balanced accuracy (SBA) score, a metric to evaluate the global preservation of patterns between maps. The spatial structure of the yield data influenced the effectiveness of the spatial calibration process. When the agronomic variable (durum wheat yield) was spatially structured, a spatialized APSIM approached performed best (5-zone modeling scale, SBA = 0.17) and outperformed the field-scale (native footprint) model (SBA = 0.19). In contrast, when the target agronomic variable was more random (less spatially structured), the uniform field-scale modeling performed best and spatial calibration had no benefit. The spatialized APSIM performances were mainly based on the reliability of the delineated zones that undeniably affected the quality of the spatialized model outputs. Thus, more research is needed on how best to model scale-dependent processes to have more reliable modeling at the within-field scale. Based on the example of a complex crop model like APSIM, this study showed that spatial calibration can be effective and has a role to play in the spatialization of complex crop models.

4.2.1.4. Graphical abstract



4.2.1.5. Keywords

Spatialization, Precision agriculture, Spatial calibration, Durum wheat yield

4.2.2. Introduction

Precision agriculture (PA) represents an opportunity to use site-specific management to increase input efficiency and reduce agriculture's environmental footprint (Khanal et al., 2017; van Evert et al., 2023).

In this context, spatial decision support tools (DSTs) are important to help farmers to adapt their cultural practices in a spatially and temporally uncertain environment to increase resource efficiency and maximize sustainability in time and space (Jones et al., 2017a). Crop models, especially mechanistic crop models, can be useful DSTs because they can take into account various variables (e.g. climatic conditions, soil properties, management options) that affect production. For this reason, mechanistic crop models are widely used to simulate various ‘what-if’ scenarios and to describe and understand how some factors (e.g. environment, weather conditions, etc.) may affect crop growth and development.

Most existing crop models simulate agronomic variables (e.g. wheat yield, vine water status, fruit nutrient content) on their native spatial footprint, i.e. the spatial footprint on which they were initially designed (e.g. plant, field, region scale) (Pasquel et al., 2022a). Crop models in current use are generally designed for modeling at the field-scale and they simulate agronomic variables by using homogeneous (average) field conditions (You et al., 2022). Even if the modeling is at the field-scale, such models are referred to as point-based models because they simulate agronomic variables at a specific spatial scale over a spatial footprint that is considered a homogeneous spatial modeling unit (Heuvelink et al., 2010). However, simulating agronomic variables at the field-scale is no longer sufficient to tackle the issues arising in the agronomic and modeling communities related to PA applications. The incorporation of crop models into PA applications is shifting the use of crop models from long-term strategic uses to short-term in-season tactical (and spatial) uses. Spatialization processes are used to change the simulation scale of crop models (Pasquel et al., 2022a). The concept of spatialization is not new and has previously been well defined by Faivre et al. (2004) in order to be able to use crop model at scales other than their native spatial footprint, particularly at larger scales to predict at regional, national and global scales. The use of these crop models in PA assumes a spatialization process that simulates agronomic variables at a finer scale than the field-scale. This way of modeling at the within-field scale for PA will be directly related to the spatial pattern of the agronomic variable. Therefore, the relevance of using such crop models at finer spatial scales will be dependent of the considered agronomic variable and its spatial distribution.

Many studies, mostly linked to the impact of climate change on crop production, have used crop models at a larger scale than their native spatial footprint. Thus, the most common form of crop model spatialization has been achieved by using upscaling methods to simulate agronomic variables at regional, national or even international scales (Asseng et al., 2018; Challinor et al., 2009; Villa et al., 2022). Most upscaling methods of applied crop models are performed on a defined grid and the crop model is run using the grid points (or pixels) as the modeling unit. Each modeling unit is individually calibrated (Hochman and Horan, 2018; van Ittersum et al., 2013). In crop-climate ensemble model studies, approximately half of the studies have used an upscaled data aggregation crop model approach, even if these crop models were initially designed at the field scale (Challinor et al., 2017). Others studies, using statistical crop models based on historical datasets, have aimed to upscale crop models to larger spatial scale (e.g. national scales) to predict the impact of climate change on crops (Lobell et al., 2008). In contrast, there have been very few studies investigating crop model uses on spatial scales smaller than their native spatial footprint, i.e. downscaled crop modeling processes. Despite this, the spatialization of crop models at a within-field scale is of great interest for PA purposes to both model and manage within-field spatial variability.

Within-field spatial variability is well-known to be highly significant to production and is caused by local interactions between several spatially variable biotic (e.g. pests, soil microorganisms) and abiotic (e.g. soil properties, weather conditions, anthropogenic consequences, topography) factors (Corwin and Lesch, 2005). Within-field production variability can be mapped by remote and/or proximal sensing data (Jin et al., 2018; Weiss et al., 2020; Zhang and Kovacs, 2012). To tackle spatial variability, fields can be divided into within-field management zones, i.e. sub-field areas that tend to have more homogenous production characteristics. Commonly in these studies, observed data are aggregated at a selected within-field zone scales. Some studies have spatialized crop models by

downscaling processes based on this management zone concept (Basso et al., 2001; Cammarano et al., 2021; Leo et al., 2023), which are also known in modeling terms as ‘functional units’ (Launay and Guerif, 2005). In this approach, the intent is to model a zonal response. The same kind of approach, by segmenting the modeling extent by simulation zone partitioning, has also been applied on upscaling studies at larger spatial scales than the native spatial footprint of existing crop models (Guo et al., 2018; Zhuo et al., 2022). Of the published works in this area, most have focused on data assimilation approaches, whereby observed spatial data sets are used to update or replace intermediates/variables within the model (Jin et al., 2018). A common example of this is the use of remotely sensed imagery as a surrogate for LAI (or biomass/vigor) within a crop model (Hu et al., 2019; Huang et al., 2019; Jin et al., 2018). Alternatively, ancillary data could be used within a spatial calibration approach to locally correct/adjust model parameters. The aim of spatial calibration is to calibrate model parameters that are likely to vary spatially based on the delineation of within-field zones that are representative of the spatial pattern of the agronomic variable. This can be done using in-season information using a data assimilation approach or, in a more classical sense, spatial calibration could be performed *a priori* (pre-season) using historical data sets. In either case of spatial calibration, the spatial pattern is hypothesized to determine the number of delineated within-field zones on which to perform the spatial calibration.

Spatial calibration studies performed *a priori* are much less common than forcing data assimilation studies, but they have the advantages of allowing the production to be modelled (and potentially managed) from day one of the season. This is in contrast to data assimilation approaches that need crop development and data collection/processing to occur before the data assimilation and crop model spatialization can be performed. One previous study of downscaling by the spatial calibration of crop model parameters using historical (rather than in-season data) was effectively performed on a relatively simple crop model (Pasquel et al., 2022b), WaLIS (Celette et al., 2010). The WaLIS model is a simple model to simulate water partitioning between the vine and cover crop using water balance equations and vine and cover crop growth equations. However, the most commonly used crop models in agriculture are more detailed and complex than WaLIS, involving more equations and inputs to better account for atmosphere-soil-plant water movements and crop physiology (Soltani and Sinclair, 2015). Examples of such models include STICS (Brisson et al., 2003, 2002, 1998), DSSAT (Hoogenboom et al., 2021, 2019; Jones et al., 2003), APSIM (Holzworth et al., 2014), AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009), WOFOST (de Wit et al., 2019), MONICA (Nendel et al., 2011) or Daisy (Abrahamsen and Hansen, 2000; Hansen et al., 2012, 1991). To date, to the authors’ knowledge, there has been no published work on how a spatial calibration approach to downscaling would (or could) perform with these complex mechanistic crop models at the within-field scale. The spatial calibration approach advocated previously by Pasquel et al. (2022b) ran the spatialized crop model on a modeling unit by modeling unit scale while maintaining a spatial consistency at the within-field scale through the delineation of within-field zones. Thus, Pasquel et al.’s (2022b) approach to spatial calibration was not at a predefined grid/pixel size, but informed by the zoning of existing and relevant data. The transfer of this approach to a much complex crop model using real-world data for validation represents one of the main innovations of this study.

Most crop models, whether simple or complex, are based on the assumption that the model parameters are homogenous over the spatial footprint that they are run on, regardless of the spatial scale. All model parameters exhibit no spatial variability, i.e. they are aspatial parameters, regardless of the type of model parameter. Moreover, they are commonly tested at the field-scale (Zhen et al., 2023). Thus, in a PA context, it is important to test the relevance of these crop model assumptions when applied at the within-field scale. For instance, if the same wheat cultivar is sown in a field, model plant parameters are not expected to change spatially. However, other model parameters relating to water and energy balances or soil dynamics, which are known to be spatiotemporally variable, would be expected to change spatially. The objective of this study is to investigate the spatialization of a complex crop model by using spatial calibration to modify crop model resolution (spatial footprint) for PA purposes. The crop model APSIM was selected for this study because it is a well understood model among the co-

authors and meets the above requirements of being a more detailed and complex crop model. This is realized by (i) using ancillary data and a segmentation algorithm to delineate within-field zones, and (ii) spatially calibrating certain model parameters at different resolutions defined by these within-field zones. The purpose is to better understand how a complex crop model responds to such a spatialization process (i.e. spatial calibration) and whether within-field scale modeling is relevant using a spatially calibrated crop model. This study represents an exploratory and preliminary work to understand how a complex crop model could work regarding a change in spatial modeling scale.

4.2.3. Material and Methods

4.2.3.1. Site description and collected data

A 12 ha experimental field of CREA (Research Centre for Cereal and Industrial Crops), near Foggia (41.462°N N, 15.506°E), south-eastern Italy was used in this study. This site has previously been used for comparing the performance of crop models under conditions of varying within-field soil properties in Wallor et al. (2018). Briefly, data in the original study were collected over three production seasons (2005-2006, 2006-2007 and 2007-2008) with durum wheat (*Triticum durum* Desf. cv. Gargano) grown in all three seasons. However, the 2006-2007 year was drought affected (poor subsoil moisture at the start of the season and low rainfall during the crop season) with very low yields recorded. Since 2006-2007 was atypical, the decision was made not to use these data in this study. The rationale for this was that the production conditions in 2006-2007 were likely to be at the limit or beyond that for which the APSIM model was designed. It is not a drought model. Spatialization of a model operating under known sub-optimal conditions was considered to be of little value as the source of any errors and effects would be ambiguous, i.e. would an observed effect in 2006-2007 be caused by modeling under drought conditions or by the model spatialization? For the other two years, meteorological conditions were typical of a Mediterranean climate, i.e. hot and dry summer (May to September) with precipitation concentrated in the autumn-winter period that coincides with cool-cold temperatures (October to April). Precipitation (P), minimum and maximum temperatures (T_{\min} and T_{\max} respectively) and solar radiation (Srad) were recorded daily at a weather station 300 m from the experiment field (Figure 4.1). Rainfall patterns of both growing seasons were different and 2007-2008 was identified as a drier growing season. The experimental field is located on a wide plain ('Tavoliere' that means flat table) so weather conditions were assumed to be homogeneous over the whole field. Management practices for the two considered years are reported in Table 4.1 and were applied uniformly on the whole field (see Wallor et al. (2018) for more details). Phenological stages were recorded for each year at seedling growth, tillering, flowering and ripening. Site-specific yield was recorded for each harvest using a John Deere combine equipped with a yield monitoring system that was calibrated prior to harvest. For these two production seasons, the mean harvested yields were similar, but they exhibited different spatial structures (Figure 4.2). In 2005-2006, the yield was spatially structured within-field, while in 2007-2008, it was more random across the whole field, as shown by the shape of variogram model as pure nugget effect (Figure 4.2d). The soil is a deep silty-clay Vertisol of alluvial origin, classified as a Fine, Mesic, Chromoxerert (IUSS Working Group WRB, 2015). Lateral water redistribution at the within-field scale was assumed to be negligible in this flat landscape. Soil properties were measured within-field at 100 optimized georeferenced sites in order to obtain an even spatial distribution (Buttafuoco et al., 2010) (Figure 4.3A). For this study, only the values of crop lower limit (CLL) and drained upper limit (DUL), resulting from texture measurements on the shallow soil layer (0-0.2 m) and computed using a pedotransfer function (Hollis et al., 2012), along with measured soil organic carbon (OC), at the same 100 locations and at the same depth, were used. Aboveground biomass and soil water content (TDR measurements) were measured at each of the 100 measurement locations, each year, at harvest and during the growing season.

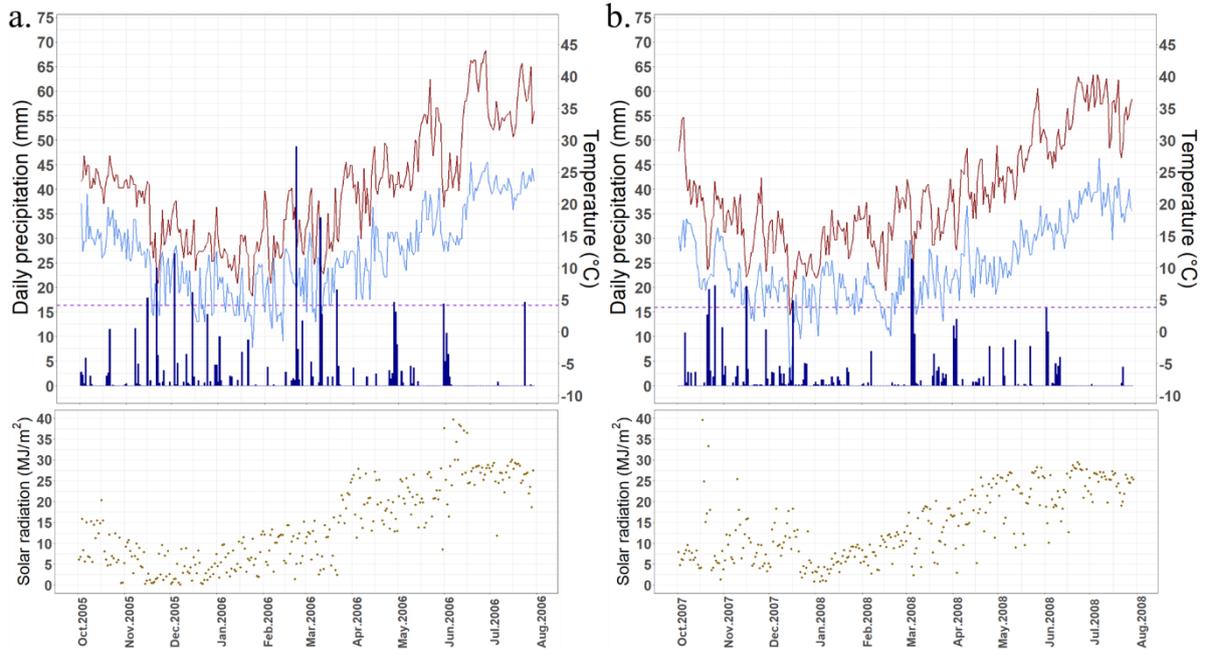


Figure 4.1 Daily precipitation, air temperatures (maximum = red lines and minimum = blue lines) and solar radiation for (a.) 2005-2006 and (b.) 2007-2008 growing seasons for the experimental field. Purple dashed lines correspond to the mean temperature over the years 2005-2006 and 2007-2008. Respectively for 2005-2006 and 2007-2008 growing seasons, cumulative rainfalls were 510.4 mm and 429.2 mm, mean solar radiation was 14 MJ/m² and 14.2 MJ/m² and mean temperatures were 14.7°C and 13.5°C.

Table 4.1 Crop practices of the experimental field used as management information for the modeling. Adapted from Wallor et al. (2018).

Date	Crop practice
2005-2006	
30.10	Ploughing (0.40 m)
13.11	Disc-harrowing (0.15 m)
07.12	Disc-harrowing (0.15 m)
15.12	Sowing winter wheat (cv. Gargano) + fertilization with ammonium phosphate (30 kg N/ha)
20.02	Fertilization with ammonium nitrate (60 kg N/ha)
26.06	Harvest winter wheat
2007-2008	
14.12	Sowing winter wheat (cv. Gargano) + fertilization with ammonium phosphate (30 kg N/ha)
28.02	Fertilization with ammonium nitrate (60 kg N/ha)
05.07	Harvest winter wheat

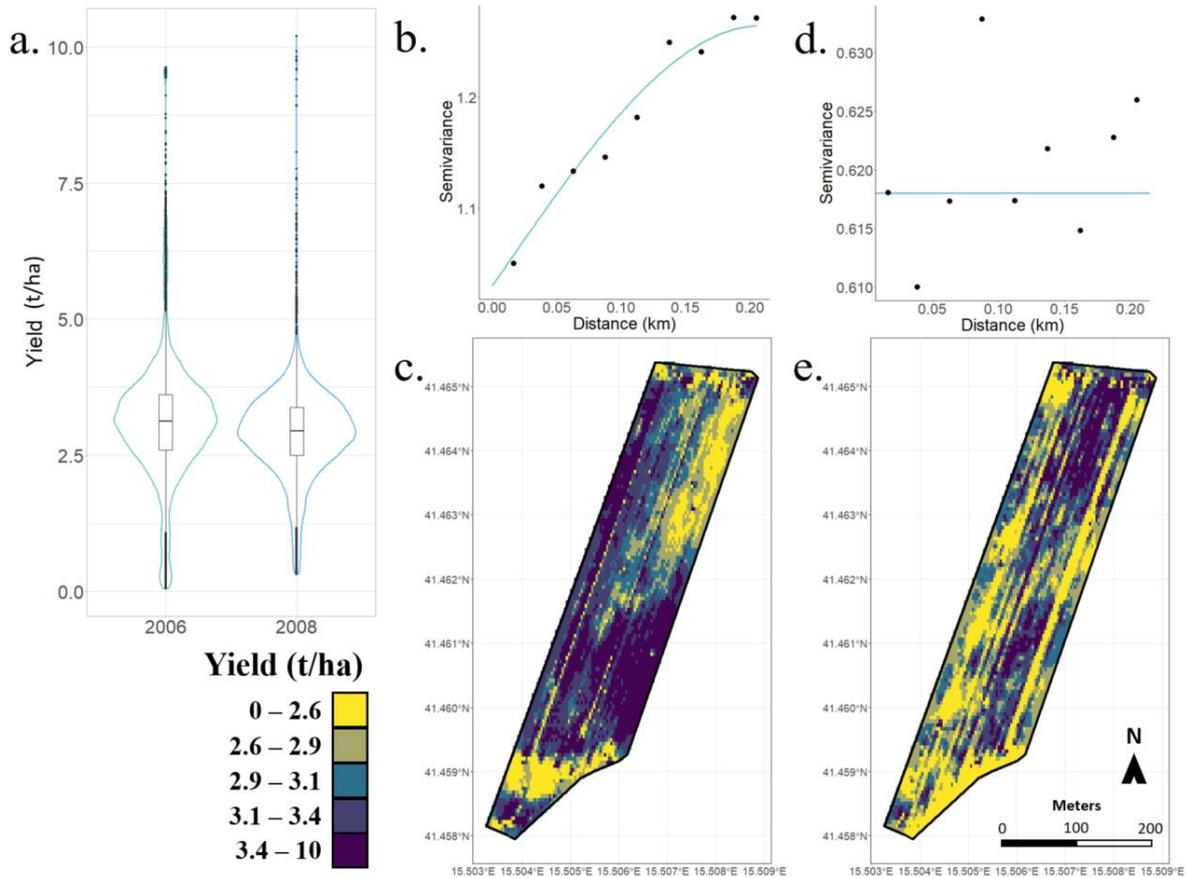


Figure 4.2 Durum wheat yield characterization. (a.) Distribution of yield values recorded by yield monitor system at harvest, mean yield in 2005-2006 and 2007-2008 were respectively 3.0 t/ha and 2.8 t/ha. (b.) Experimental and theoretical variogram of durum wheat yield in 2005-2006, yield is spatially structured at within-field scale. (c.) Durum wheat yield map in 2005-2006. (d.) Experimental and theoretical variogram of durum wheat yield in 2007-2008, yield exhibits no spatial structure at the within-field scale as shown by the shape of variogram model as pure nugget effect. (e.) Durum wheat yield map in 2007-2008.

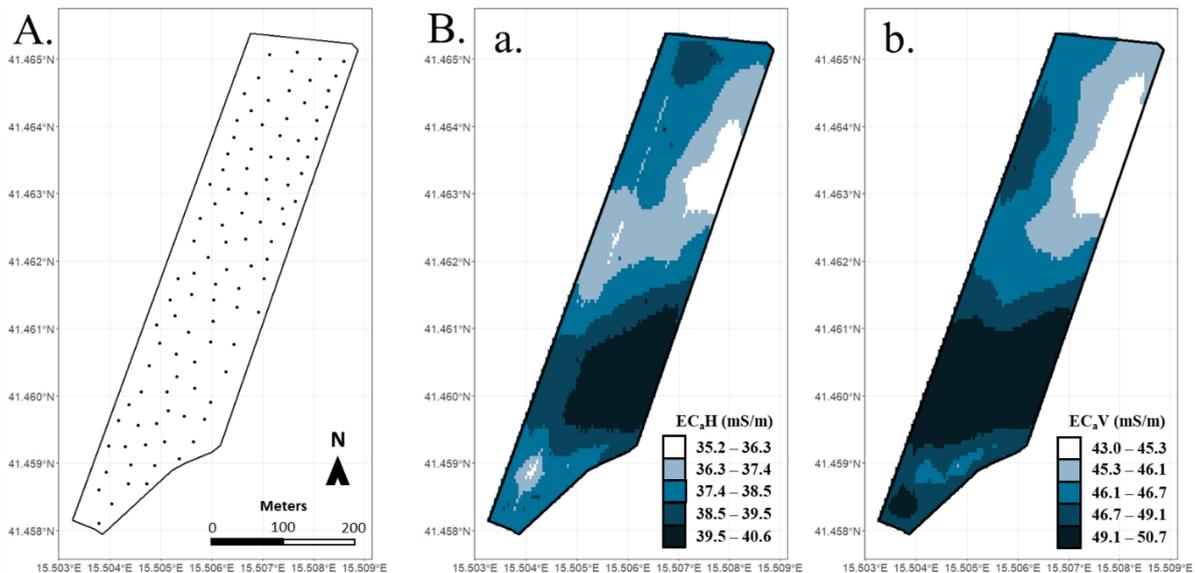


Figure 4.3 Experimental field with A. Location of 100 measurement sites at within-field scale. B. Interpolated maps of (a.) apparent soil electrical conductivity in horizontal mode (EC_aH) and (b.) in vertical mode (EC_aV) using inverse distance weighting.

A soil apparent electrical conductivity (ECa) survey was performed in 2010 using an electromagnetic induction ground conductivity meter (EM38DD, Geonics, Ltd, Ontario-Canada). The ECa was simultaneously measured in two polarization modes that explored different depths depending mostly on soil moisture conditions and textural properties (Sudduth et al., 2001). The EM38DD was set up to provide a depth of exploration equivalent to the topsoil layer in the horizontal mode (ECaH, maximum ECa sensitivity at 0-0.10 m) and to the expected, typical rooting depth in the vertical mode (ECaV, maximum ECa sensitivity at ~0.40 m depth) (Figure 4.3B).

4.2.3.2. APSIM and modeled durum wheat yield

Durum wheat yield was modeled using the crop model Agricultural Production Systems sIMulator - APSIM 7.14 (Holzworth et al., 2014) through the *apsimx* R package (Miguez, 2022) in R 4.2.0 (R Core Team, 2022). APSIM required soil properties as input for different soil layers, and for this study it was decided to define soil input up to a depth of 2 m (for the entire soil profile). However, observed soil properties were only available for the first 0.20 m. Therefore, subsoil soil properties needed to be estimated. To do this, the soil profile was first divided into seven layers to ensure the correct functioning of the ground modeling (Figure 4.4). A pedometrics approach was used to model soil hydraulic limit (SHL) values, i.e. air dry moisture content (AD), lower limit soil moisture at -1.5 MPa (LL15), crop lower limit (CLL), soil water (SW), drained upper limit (DUL) and soil water at saturation (SAT) at each layer from the topsoil observations (Figure 4.4). In particular, SHL values at each measurement site profile were estimated from the observed topsoil measurements and known SHL shapes for Vertisol soils described in (Dalgliesh et al., 2016) in order to initialize model parameters. The SW for each layer i was calculated from the modeled CLL and DUL of each layer using Eq. 4.1.

$$SW_i = 0.25 \times ASW_i = 0.25 \times (DUL_i - CLL_i) \quad (4.1)$$

where SW_i , ASW_i , DUL_i and CLL_i are respectively the soil water, the available soil water, the drained upper limit and the crop lower limit for the i^{th} layer.

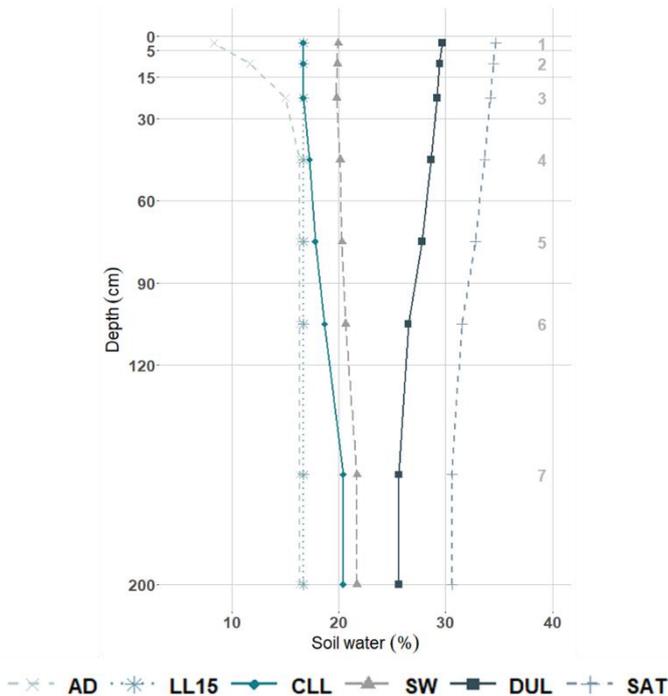


Figure 4.4 Example of a general soil profile shape for soil water properties used to model durum wheat yield with APSIM for one of the considered measurement sites. Right-hand side numbers refer to the soil layer number. Lines refer to soil hydraulic limits: AD = air dry moisture content, LL15 = lower limit soil moisture at -1.5 MPa, CLL = crop lower limit, SW = soil water, DUL = drained upper limit and SAT = soil water at saturation. Each point refers to the soil hydraulic limit value for the corresponding layer in volumetric water content.

4.2.3.3. Spatial calibration of APSIM

Calibration of soil and plant parameters was performed independently on each considered production year and evaluated on the other year as explained in Section 4.2.3.4. Weather data, soil information, initial soil water, nitrogen content and agronomic management practices were used as input for the calibration. APSIM was spatialized by using a downscaling approach that mainly involved spatial calibration, by defining within-field zones using ancillary data and identify which parameters could be spatially calibrated. The high-resolution soil sensor data, EC_aV and EC_aH data, were used together for delineating within-field zones (minimum of 2 and maximum of 20) using a segmentation algorithm (Pedroso et al., 2010) with the *GeoFIS* R package (Guillaume and Lablée, 2022). Default settings were used with the segmentation algorithm. Both ancillary data were correlated to durum wheat yield so they could potentially explain the yield variability. Following the zone delineation, the different spatial scales which were considered for the calibration and the evaluation in this study were:

- (i) the measurement sites scale ($n = 100$);
- (ii) several within-field zones ($z \in [2;20]$) (Figure 4.5);
- (iii) and the whole field (equivalent to a one zone solution), i.e. the APSIM native spatial footprint.

For scales higher than the site-scale, data were aggregated at the different zonal scales by averaging the observations located within each individual zone. For all performance assessments, the output scale was disaggregated to the measurement site-scale to evaluate the modeling performance (Figure 4.6). However, some parameters were calibrated with the same value regardless of the of spatial modeling units, i.e. these parameters were fixed whatever the modeling scale. For instance, this is the case for the cultivar parameters (related to phenology), because it was the same sown cultivar under the same climatic conditions.

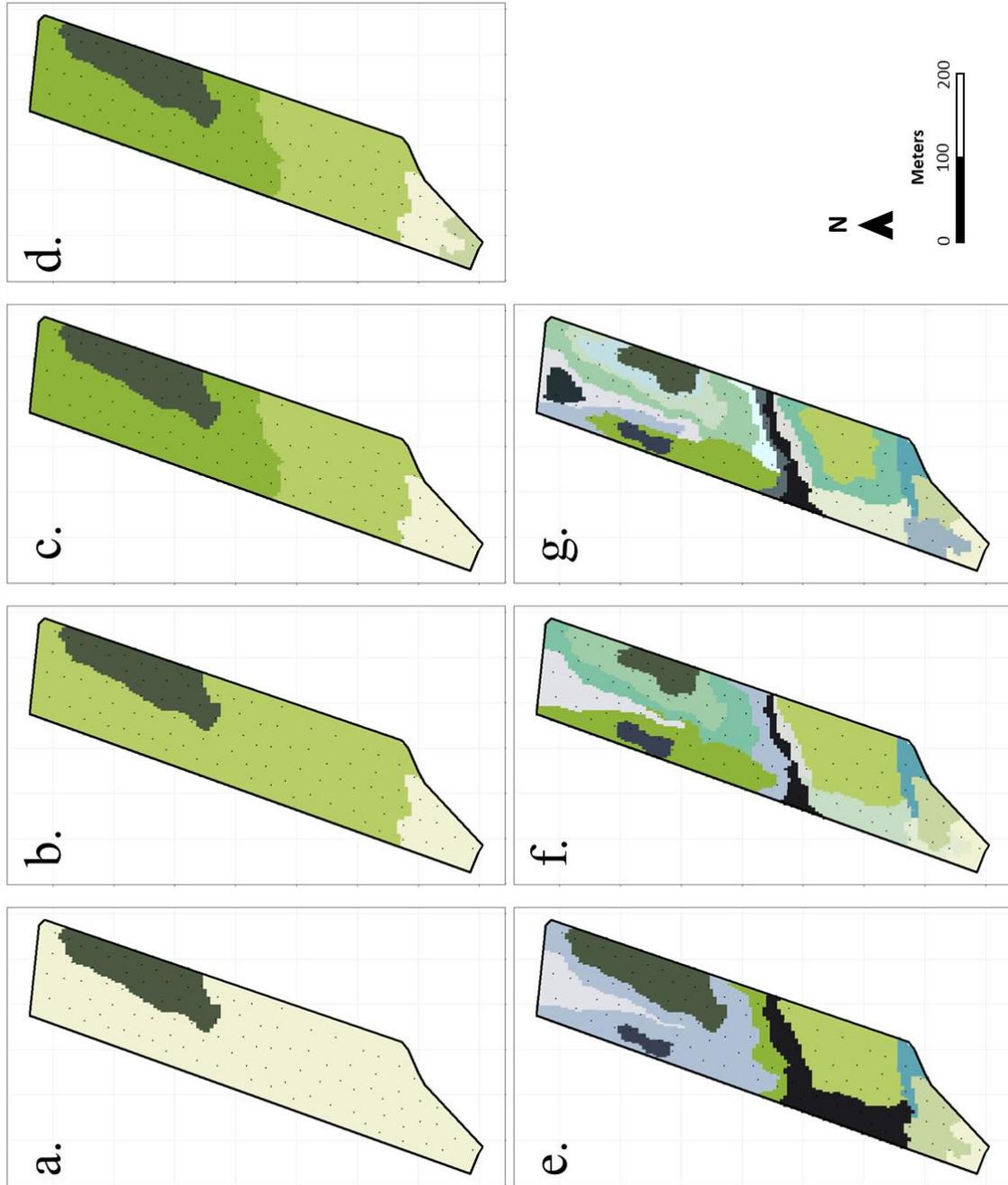


Figure 4.5 Maps of different within-field zones defined using a segmentation algorithm based on soil apparent electrical conductivity data on horizontal and vertical mode together; 2-zone (a.), 3-zone (b.), 4-zone (c.), 5-zone (d.), 10-zone (e.), 15-zone (f.) and 20-zone (g.) for the experimental field. Note that the 1-zone is equivalent to the whole field scale (not shown). Points within the field indicate sampling locations for soil and crop parameters.

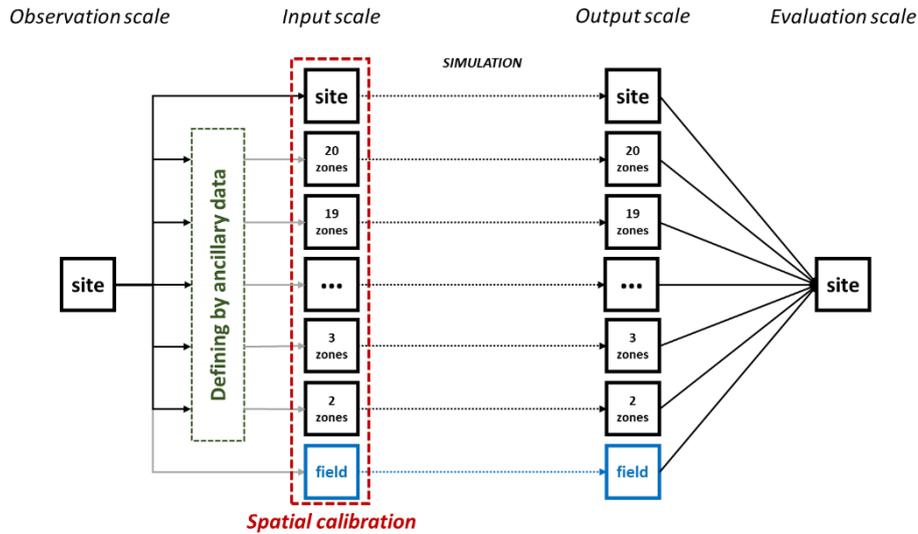


Figure 4.6 Different modeling scales of durum wheat yield using APSIM from site measurement observation scale, intermediate within-field scales (2 to 20 zones) defined with ancillary data related to soil characteristics up to the whole field scale. The APSIM’s native spatial footprint is shown in blue and corresponds to the field scale. Measurement site-scale corresponds to the original observation scale. The grey arrows correspond to the upscaling process associated with aggregations of the observed data to a higher spatial scale as model input. The spatial calibration is performed at this input scale.

Spatial calibration was performed through several spatial and aspatial steps (Figure 4.7) to adjust (1) the observed phenology stages with the modeled phenology stages, then (2) the observed soil water content with the modeled soil water content, then (3) the observed biomass with the modeled biomass and then (4) the observed yield with the modeled yield. These steps were selected to match with de Wit’s concept (van Ittersum et al., 2003) of crop growth modeling, i.e. calibrating first phenology, constraints related to light, temperature and crop genetic and then stresses related to soil water content. This calibration methodology based on expertise was also applied to others studies on crop modeling particularly because it helps complex optimization which can be too difficult with lots of local optima (Seidel et al., 2018; Wallach et al., 2021). Calibration of the aspatial parameters assumed that parameters were constant over the field regardless of the number of within-field zones/sites (e.g. cultivar parameters). Calibration of spatial parameters assumed the parameter values could be different in the within-field zones or at the measurement sites (e.g. soil parameters).

First, cultivar phenology parameters were adjusted to match the observed phenology stages with the modeled phenology stages. As only one cultivar (Gargano) was sown, the phenology was assumed to be homogeneous over the field and corresponded to an aspatial calibration (Figure 4.7a.). Secondly, the soil water content was calibrated spatially by modification of the SHL values. The CLL and DUL values on the first layer (0-0.05 m) were estimated directly from the observed soil texture measurements related to the plant available water capacity. Calibration was used to determine CLL and DUL values for the other layers (2nd to 7th) and also the values of other SHLs (AD, LL15, SW, and SAT) for the whole soil profile (Figure 4.7b.). This spatial calibration tended to conserve the global shape of the local soil profile (example described in Figure 4.4) for each site. Thirdly, biomass was calibrated by adjusting spatial and aspatial model parameters. Aspatial calibration was related to cultivar parameters considered homogeneous over the whole field (Figure 4.7c.). However, the biomass was not considered homogeneous at the within-field scale since it may depend on water and nutrient availability in the field (Mon et al., 2016). Thus, parameters not related to potential biomass but to water stress affecting its expansive growth needed to be calibrated spatially (Figure 4.7d.).

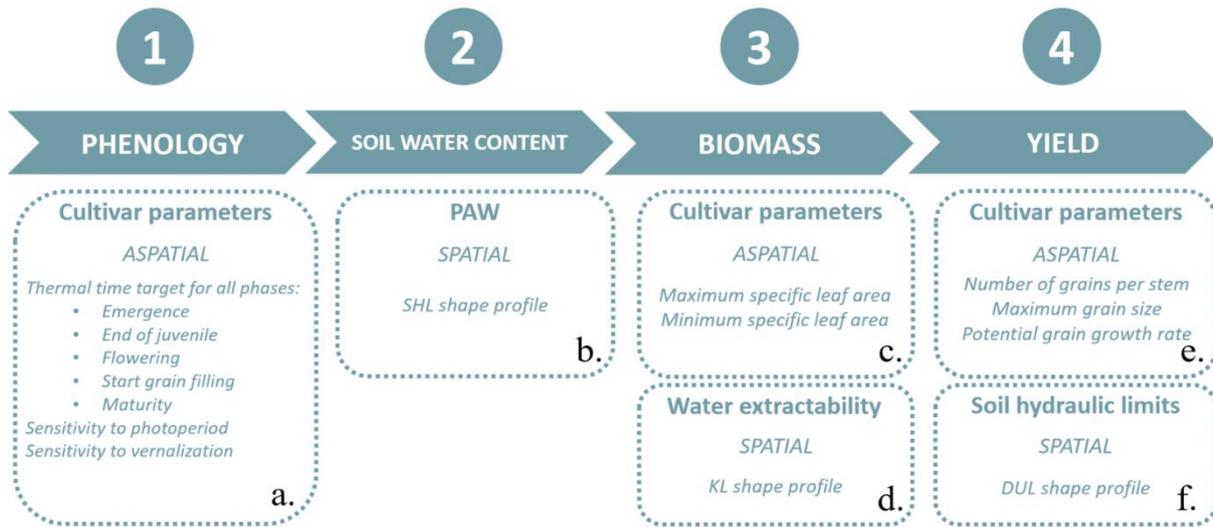


Figure 4.7 Calibration steps followed for the spatial calibration of APSIM. Cultivar parameters were not spatially calibrated, whereas the other parameters were spatially calibrated. Italic parameters correspond to APSIM input parameters. SHL: soil hydraulic limits, PAW: plant available water, KL: fraction of plant available water able to be extracted, DUL: drained upper limit.

As a first approximation to account for biomass variability, the fraction of plant available water able to be extracted (KL) in the model was considered to be the main driver of the biomass. Before this calibration step, four classes of soil KL (A, B, C and D) were defined that were considered representative of the soil profiles with increasing extractable water from A to D (see Supplementary Figure S.B1). The KL was spatially calibrated by assigning the different soil profiles (in the case of the site-specific modeling) or a zonal mean soil water profile to one of these four classes. Calibrating KL this way was a first approach to account for spatial soil moisture variability in the field. Finally, yield was calibrated by again adjusting the aspatial cultivar parameters in terms of grain number size (Figure 4.7e.) and then, in a similar way as to the KL calibration, three soil DUL classes (1, 2 and 3) were specified and spatially adjusted (Figure 4.7f.) to correspond to different available soil water regimes (see Supplementary Figure S.B2). KL and DUL were chosen to be spatially calibrated because a previous study (Basso et al., 2009) showed that subsoil constraints (especially soil water retention properties) were the main factors impacting spatial structure in this field. All calibrated and estimated parameters needed as APSIM inputs are shown in Table 4.2. Note that in order to ensure a consistency of the calibrated parameters, calibration was carried by an optimization of the value of each parameters previously cited. The optimization was made as objective as possible by finding the best combination of values that resulted in the best calibration results from an expert-defined value domain range for each parameters, i.e. the optimum values were found by exhaustive search in a grid of values with physically consistent bounds. The possible domain range used for each parameter was consistent regardless of the modeling scale. This procedure aimed to be a relevant approach to following and testing the calibration method whilst maintaining logical values for the parameters that were describing the underlying biophysical processes.

For each of these calibration steps, the root mean square error of calibration (RMSE_C) (Eq. 4.2) was used to determine the optimal value of the target model parameter by comparing the observed and modeled parameter values.

$$RMSE_C = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_{C,i})^2} \quad (4.2)$$

where y_i is the observed value, $\hat{y}_{C,i}$ is the corresponding modeled value for the calibration and n is the number of observations.

4.2.3.4. Model output evaluation

To test the temporal stability of the spatial calibration, APSIM performance was tested for the different spatial scales in two distinct cases: (i) calibration on 2005-2006 data and evaluation on 2007-2008 data and, (ii) the inverse, with calibration on 2007-2008 data and evaluation on 2005-2006 data. Note that both seasons were dissimilar in weather conditions and grain yield production. The cultivar parameters were calibrated differently for both cases because having only 2 years of data was not enough to estimate the general parameters of this cultivar under these climatic conditions. Thus, cultivar parameters were calibrated individually for both cases to better match the predicted yield. To evaluate APSIM performance, durum wheat biomass and yield were qualitatively evaluated from the maps and quantitatively evaluated using two metrics: root mean square error of prediction ($RMSE_p$) (Eq. 4.3) to evaluate prediction performance and spatial balanced accuracy (SBA) (Eq. 4.4) to evaluate simulation performance accounting for spatial relevance.

$$RMSE_p = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_{E,i})^2} \quad (4.3)$$

where y_i is the observed value, $\hat{y}_{E,i}$ is the corresponding modeled value and n is the number of observations.

The SBA is a specific metric for spatialized crop models (Pasquel et al., 2023b) calculated by assessing both aspatial and spatial pattern errors. Thus, SBA is able to identify which simulation scale is the most relevant for modeling an agronomic variable (durum wheat yield here) using a given model (APSIM) and a given downscaling process (spatial calibration of selected model parameters).

$$SBA = \frac{1}{100} \sum_{q=1}^{100} [1 - BA(O_{t(O,M,q)}, M_{t(O,M,q)})] \quad (4.4)$$

where O and M are respectively the observed and modeled maps, $O_{t(O,M,q)}$ and $M_{t(O,M,q)}$ are respectively the observed and modeled maps at threshold level $t(O,M,q)$ that is defined relative to percentile q on the merging data distributions of O and M .

The closer SBA is to 0, the better the agreement between the observed data and output from the spatialized model. Note that in the spatial calibration step this metric was not used to determine the spatialized model parameters (only $RMSE_C$ was used). However, the SBA scores for biomass and yield maps were calculated during the calibration process to provide a greater understanding on how the calibration was affecting the spatialized APSIM model outputs at different simulation spatial scales.

Table 4.2 Summary of calibrated and estimated parameters necessary as APSIM inputs to simulate durum wheat yield. Further details on the values of the aspatially calibrated parameters are available in Supplementary Table S.B1 and S.B2.

Soil property	Abbreviation	Acquisition	Calibration
Lower limit -1.5 MPa	LL15	Measured on the 1 st layer Estimated from Dalglish et al. (2016) for other layers	
Drained upper limit	DUL	Implement different profiles (1, 2 and 3) to match with the yield of the calibration year (see SHL profiles), variation of DUL profile to reduce wheat available soil water	
Bulk density	BD		
Air dry	AD		
Saturation	SAT	Estimated from Dalglish et al. (2016) based on LL15 and DUL measurement on 1 st layer	Spatial
Crop lower limit	CLL		
Soil water	SW	Equation 1	
Fraction of plant available water able to be extracted	KL	Estimated from Dalglish et al. (2016) Implement different profiles (A, B, C and D) to match with the biomass of	
Organic carbon	OC	Measured on the 1 st layer Estimated from Dalglish et al. (2016) for other layers	
Root exploration factor	XF		
Proportion of non-inert C in the	FBiom	Estimated from Dalglish et al. (2016)	
Proportion of initial organic C	FInert		
Target thermal time (cultivar parameter)	-		Aspatial
Leaf area growth (cultivar parameter)	-	Calibrated by an exhaustive search on a discredited range	
Seed growth (cultivar parameter)	-		

Results

4.2.3.5. Spatial calibration maps for the different simulation scales

Spatial calibration of the KL and DUL profiles differed with simulation scale when performed on the 2005-2006 data (Figures 4.8 and 4.9). For the within-field spatial calibration (zonal approach), several within-field zones had the same parameter values (profiles). For the site-scale, the diversity of calibrated KL and DUL profiles was more important than for higher simulation scales (more details on KL and DUL profiles are shown in Supplementary Figures S.B1 and S.B2; note that KL and DUL profiles will be hereafter designed respectively from A to D profiles and from 1 to 3 profiles as explained in Section 4.2.3.3). For the site-scale spatial calibration, the KL and DUL profiles represented the full diversity of considered profiles. However, for simulation at larger zones, this diversity was reduced, e.g. there was an absence of the KL profile D for simulation scales below 10 within-field zones (Figures 4.8a to 4.8e) and an absence of the DUL profile 3 for simulation scales below 20 within-field zones (Figures 4.9a to 4.9h). Spatial calibration for scales higher than the site-scale simulation tended to homogenize calibrated profiles. Moreover, the merging of different within-field zones did not necessarily lead to an intermediate profile of these different within-field zones. For example, in 2005-2006, the western side of the field was mainly calibrated to DUL profile 2 with some 1 and 3 profile zones at the site scale, but calibrated to profile 1 especially for lower order zoning (Figure 4.9).

Spatial calibration of KL and DUL profiles performed with the 2007-2008 data showed the same trends as the 2005-2006 data described above. They are given in Supplementary Figure S.B3 and S.B4 and. Maps of the other spatial soil inputs to APSIM, OC and difference between CLL and DUL, are also shown in Supplementary Figure S.B5. Only the topsoil maps (directly observed data) are shown as the values in the subsoil layers were estimated from these topsoil data (Supplementary Figure S.B5).

Minimization of $RMSE_C$ for the spatial calibration of KL and DUL profiles are shown in Supplementary Figures S.B6, S.B7, S.B8 and S.B9. The values for parameters aspatially calibrated are available in Supplementary Table S.B1 and S.B2.

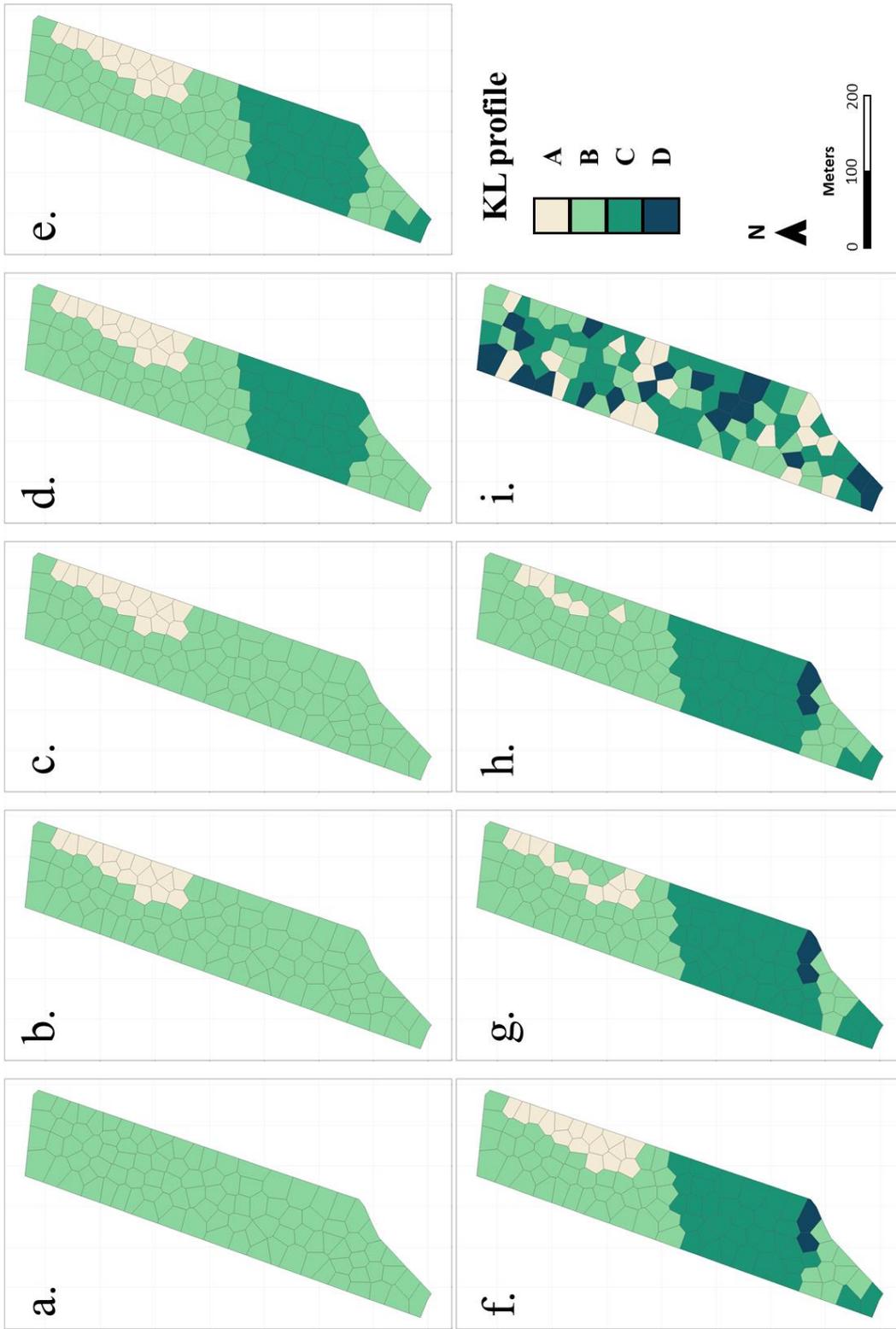


Figure 4.8 Maps of spatially calibrated fraction of plant available water able to be extracted (KL) profiles for each simulation scale on 2005-2006 data: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Voronoi tessellation was used to convert each data point into polygons. Note that A and D KL profiles are respectively the soil profiles with the least and the most extractable water (more details on KL profiles are shown in Supplementary Figure S.B1).

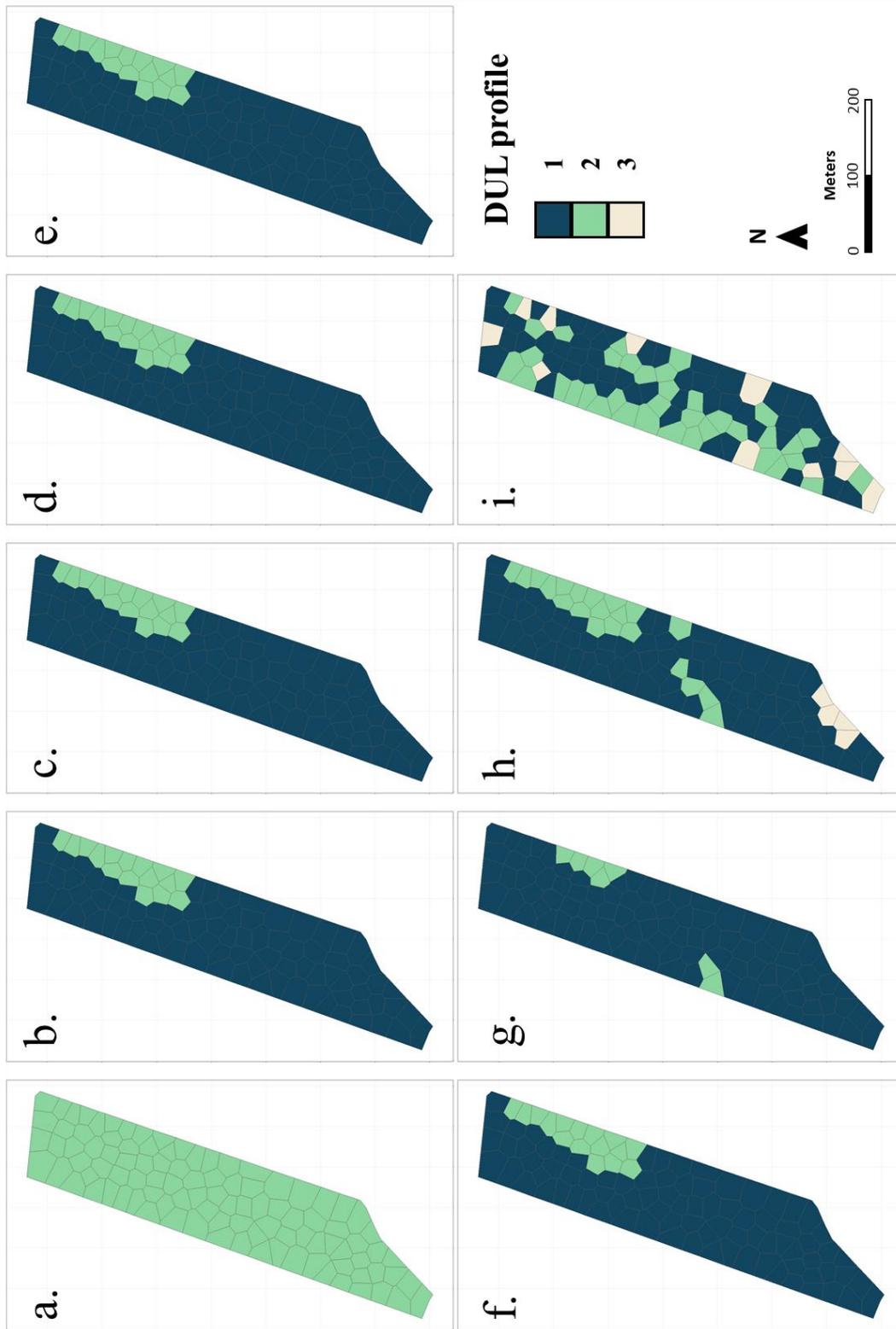


Figure 4.9 Maps of spatially calibrated drained upper limit (DUL) profiles for each simulation scale on 2005-2006 data: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Note that 1 and 3 DUL profiles are respectively the soil profiles with the most and the least available soil water (more details on DUL profiles are shown in Supplementary Figure S.B2).

4.2.3.6. Calibration performances using SBA

The SBA scores (Table 4.3) were also calculated during the model calibration process to see if the best performing modeling scale could be identified before (and was consistent with) the evaluation step. SBA scores were computed for the spatial calibration steps 3 and 4, i.e. to assess the coherence between observed and modeled biomass and yield data, respectively. Compared with the SBA scores computed for the model evaluation (Section 4.2.4.3), there was less variation in the biomass and yield SBA scores during the calibration step (Table 4.3). For the yield SBA scores for calibration, the variations were similar to the yield prediction SBA scores, in particular there was a relative stability in the SBA scores for simulation scales from 4-zone to site scale modeling. The exception to this was the calibrated site-specific yield SBA score in 2007-2008 data that was considerably lower than the 20-zone SBA score. Thus, SBA scores on calibration steps did not match with SBA scores computed for the evaluation step (Section 4.2.4.3) on durum wheat yield.

Table 4.3 Spatial balanced accuracy (SBA) to assess calibration of APSIM spatialized version for simulating durum wheat biomass and yield at the field level and at different within-field spatial scales for the two different years.

Year	Variable	Scale (zones)								
		Field	2	3	4	5	10	15	20	Site
2005- 2006	Biomass	0.13	0.19	0.19	0.18	0.18	0.18	0.19	0.21	0.16
	Yield	0.14	0.28	0.26	0.27	0.24	0.23	0.24	0.24	0.23
2007- 2008	Biomass	0.16	0.23	0.24	0.23	0.22	0.20	0.19	0.21	0.15
	Yield	0.15	0.26	0.23	0.27	0.27	0.27	0.25	0.26	0.19

4.2.3.7. Spatialized APSIM performance to simulate durum wheat yield

When the model was calibrated on the 2005-2006 data and applied to the 2007-2008 data, there was a difference of 20% in the mean aspatial error between the best performed modeling scale (field scale) and the worst performed modeling scale (site-scale), with respectively a RMSE_P of 0.94 t/ha and 1.17 t/ha (Table 4.4). There was no strong visual linear 1:1 relationship between the observed and modeled zonal or site-specific yield (Supplementary Figure S.B10). In most cases, the zone with the highest modeled yield tended towards an overestimation of yield. In this situation, it is unsurprising that the yield prediction at field-scale modeling generated the lowest RMSE_P. When the data sets were inverted (calibration on 2007-2008, prediction on 2005-2006), there was a difference of 9% of aspatial error between the best performed modeling scale (2-zone scale) and the worst performed modeling scale (site-scale), with respectively 1.32 t/ha and 1.45 t/ha (Table 4.4). Again, there was no clear linear 1:1 relationship between the observed and modeled zonal or site-specific yield (Supplementary Figure S.B11). However, modeled yields at within-field scales had less overestimation with this combination and the 2-zone scale was identified as the best performing scale.

Regarding the spatial error between the observed and modeled data, it was difficult to identify which modeling scale had the best performance by only using the 1:1 plots (Supplementary Figures S.B10 and S.B11) and the simulated yield maps (Figures 4.10 and 4.11). Indeed, concerning yield maps, it was difficult to identify a real spatial pattern among the site-scale observed data (Figures 4.10 and 4.11), whereas modeling at within-field scales exhibit clear delineated zones which did not clearly match visually with the observed data (Figures 4.10 and 4.11). None of the within-field scale for both years of calibration/modeling seemed to be the more relevant.

Table 4.4 Root mean square error of prediction ($RMSE_p$) to assess the ability of the spatialized APSIM version to simulate durum wheat yield at field level and different within-field spatial scales for two different years of calibration and evaluation. Values indicated are in t/ha.

Calibration year	Evaluation year	Scale (zones)								
		Field	2	3	4	5	10	15	20	Site
2005-2006	2007-2008	0.94	1.04	1.02	1.13	1.13	1.09	1.04	1.04	1.17
2007-2008	2005-2006	1.44	1.32	1.45	1.42	1.42	1.40	1.41	1.38	1.45

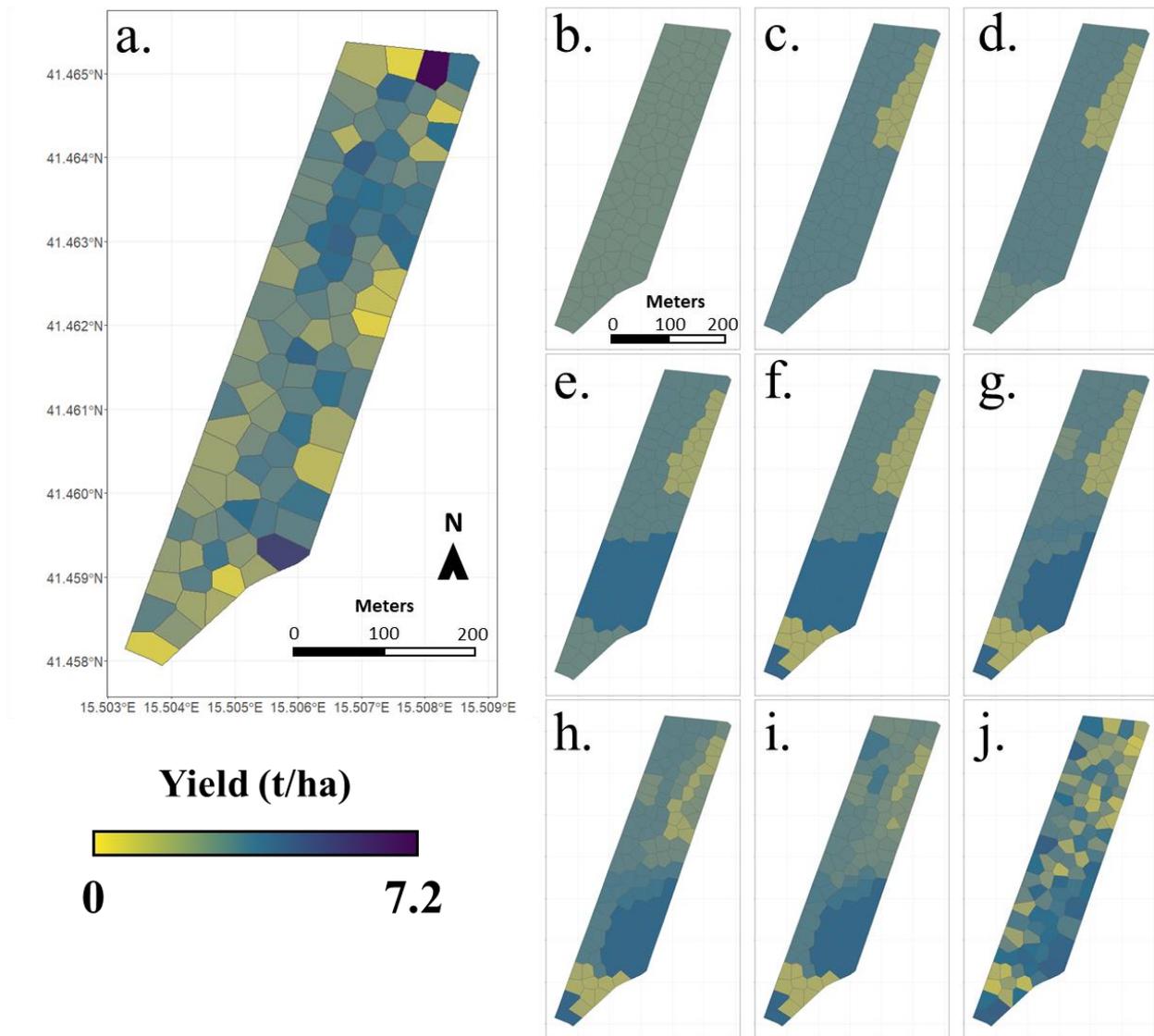


Figure 4.10 Spatial patterns of durum wheat yield for (a.) observed data from 2007-2008 and modeled data from APSIM calibrated on 2005-2006 data for different modeling scales: (b.) field scale, (c.) 2-zone scale, (d.) 3-zone scale, (e.) 4-zone scale, (f.) 5-zone scale, (g.) 10-zone scale, (h.) 15-zone scale, (i.) 20-zone scale and (j.) site-scale.

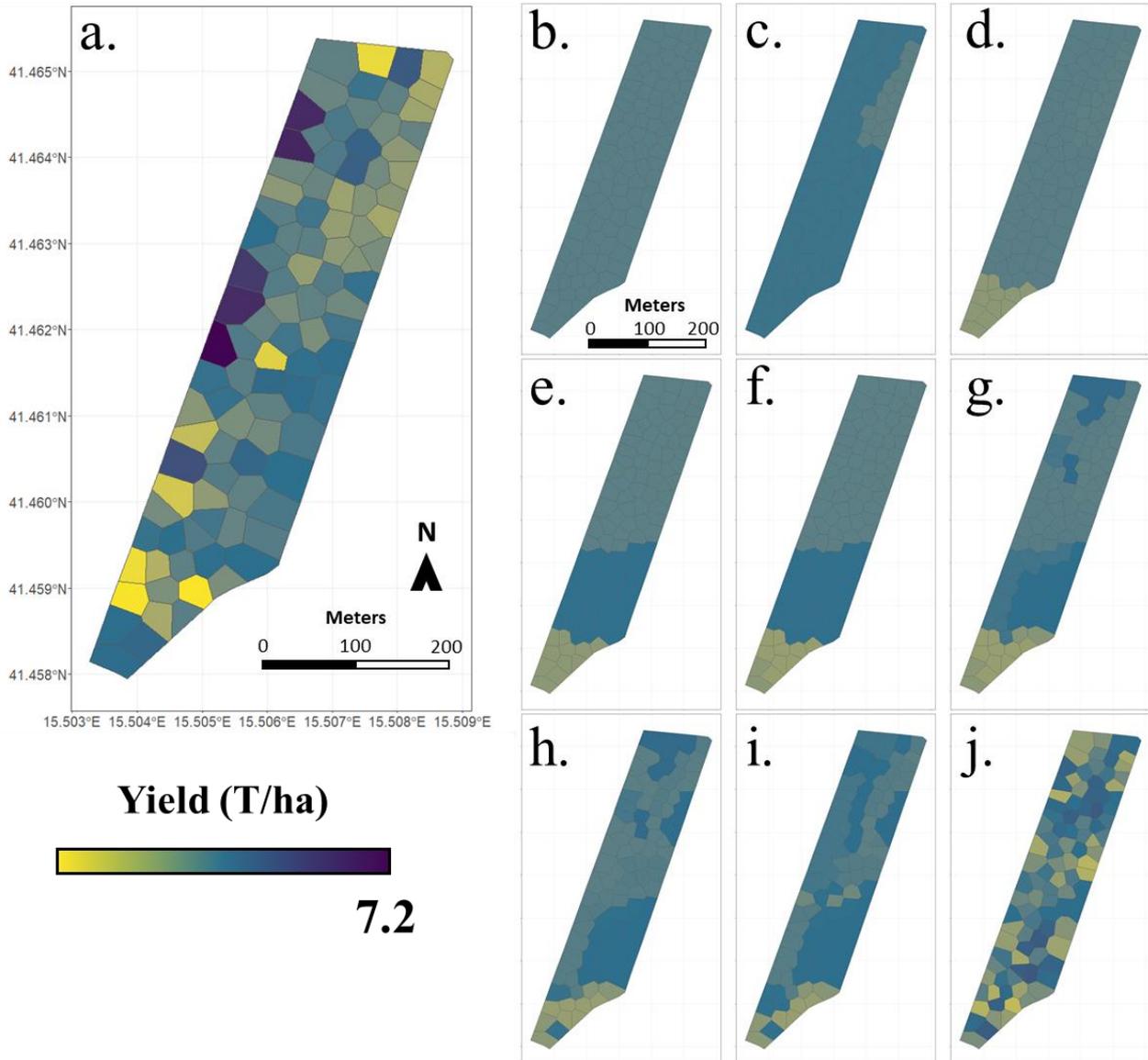


Figure 4.11 Spatial patterns of durum wheat yield for (a.) observed data from 2005-2006 and modeled data from APSIM calibrated on 2007-2008 data for different modeling scales: (b.) field scale, (c.) 2-zone scale, (d.) 3-zone scale, (e.) 4-zone scale, (f.) 5-zone scale, (g.) 10-zone scale, (h.) 15-zone scale, (i.) 20-zone scale and (j.) site-scale.

The SBA scores (Table 4.5) on the predicted 2005-2006 and 2007-2008 yields respectively showed that field scale modeling and 5-zone modeling were identified as the best performing, with SBA scores of 0.15 and 0.17 respectively. Compared with results shown in Figure 4.10 and 4.11 and Supplementary Figure S.B10 and S.B11, the SBA scores showed additional information of the model performance that could not be identified from the $RMSE_P$ values, the observed vs modeled plots or the visual yield map comparisons. With only a $RMSE_P$ interpretation, evaluation on the predicted 2005-2006 yield identified the 2-zone modeling scale as the best performing scale. In contrast, the SBA scores identified the 5-zone modeling as the best performed (Table 4.5 and Supplementary Figure S.B11). The SBA scores gave a relevant spatial evaluation of the APSIM performance as defined for spatialized crop models with estimation of aspatial and spatial error (Pasquel et al., 2023b). There was a stabilization of SBA scores between the 10-zone to site-scale modeling. The biggest deviations in the SBA scores tended to be located between the field-scale and 5-zone scale modeling in both years, although the SBA scores and their evolution with the number of zones was very different depending on the calibration/evaluation year. It is clear that a calibration in 2007-2008 resulted in lower SBA scores.

Table 4.5 Spatial balanced accuracy (SBA) scores to assess the ability of the spatialized APSIM version to simulate durum wheat yield at field scale and at different within-field spatial scales for two different years of calibration and evaluation.

Calibration year	Evaluation year	Scale (zones)								
		Field	2	3	4	5	10	15	20	Site
2005-2006	2007-2008	0.15	0.33	0.30	0.32	0.30	0.30	0.29	0.29	0.31
2007-2008	2005-2006	0.19	0.25	0.18	0.18	0.17	0.19	0.19	0.19	0.18

4.2.4. Discussion

4.2.4.1. Different modeling performances depending on the calibration/prediction year

Depending on the years used for calibration and prediction, the results showed different spatialized APSIM performances for yield prediction at the within-field scale. It is hypothesized that this was affected by the differences in spatial structure of the yield data along with difference in rainfall patterns, even though the mean and numerical distribution of the yield data were similar. The 2005-2006 yield data was spatially structured while the 2007-2008 yield data was not. When the spatially structured 2005-2006 yield data is used for the spatial calibration of some model parameters and then applied to the poorly spatially structured 2007-2008 yield data, the field scale was identified as the most relevant modeling scale. The lower level of spatial structure (and patterning) in the 2007-2008 yield data suited a mean yield response that fits to the native spatial footprint. In effect, each part of the field could be calibrated with the same values for the APSIM input parameters, i.e. the whole field has the same modeled yield value (even if yield actually varied over the field). Therefore, under these conditions, the SBA score identified field-scale modeling as the best performing scale even if there were some within-field yield variations ignored by this modeling scale. Field-scale modeling allowed the best trade-off for a better modeling of the yield based on aspatial and spatial coherence with the observed data. In contrast, when the spatial calibration was done on the poorly spatially structured 2007-2008 yield data, and evaluated on the spatially structured 2005-2006 yield data, the 5-zone scale was identified as the most relevant modeling scale according to the SBA score. In this situation, the spatial calibration was carried out on within-field zones segmented from spatially structured ancillary data that were correlated with durum wheat yield. Even with a relatively poor spatial structure in the 2007-2008 yield data, this spatial calibration succeeded in defining a distinction between the zone/site-specific APSIM input parameter values to reproduce the spatial patterns in the observed data. Thus, spatial calibration was relevant in this case. However, the relevance of the spatial calibration was based on the data from the end of the season that does raise questions regarding the management decisions that should have been made during the season. This is an important question to make this method applicable in a real world situation. However, in the present study the aim was to understand how APSIM is working at the within-field scale, i.e. to understand if it is relevant to use APSIM at finer spatial scales. Based on these results, considerations on how this method applicable for farmers to advise them in differentially adapting their management for in-season production are better known. The spatial calibration aims to constrain the calibration process to ensure the reproduction of the spatial pattern of the agronomic variable. However, more work is still necessary to make a relevant spatial calibration in a truly operational context, i.e. there needs to be a trade-off between the calibration zones and the management zones, which reflect the real within-field management practices. The optimal scale of spatial calibration for the model will not necessarily align with the optimal scale of management possible by the grower. Other complementary methods could be applied to the output maps to take into account the operational constraint linked to used machines for instance (Leroux and Tisseyre, 2018). However, the aim of this study was to investigate the modeling of the durum wheat yield at the within-field scale (i.e. theoretical objective) and evaluation was made regarding this objective. Modeling related to an operational context was

considered here as an perspective to this work, the first step was to investigate if a use of existing crop models was relevant at within-field scale.

There was not a lot of variation among the SBA scores, with spatial calibration on the 2007-2008 data being similar between the 3-zone scale and site-scale even if the 5-zone scale modeling was identified as the best performing scale. Given these similarities, consideration could also be given to the principle of parsimony and the need to have zones and decisions that can be enacted from an agronomic perspective. Calibrating APSIM at 3 within-field zones is less time consuming and likely more relevant for management (avoids calibration on outliers and conserves spatial pattern consistency) compared to finer resolution calibration (site-scale). However, it could also ‘miss’ punctual specific patterns if not representative of the within-field segmentation.

The spatial structures and the resulting spatial patterns of the durum wheat yield were not the same between the two years because of varying precipitation and temperature profiles between years (Figure 4.1). Weather conditions heavily affect rain-fed grain yield determination, especially rainfall amount and distribution over the crop season (Buttafuoco et al., 2017). Previous studies have attributed spatial variation of yield components mainly to different levels of available soil water between production years (Diacono et al., 2012; Guastaferro et al., 2010) and on this same field, Basso et al. (2012) showed that growing season rainfall and fallow rainfall were correlated with grain yield over a five year period as a result of the complex dynamic interactions between spatial static properties (e.g. soil texture) and dynamic properties (e.g. soil water content). In this study, soil information (ECa) was used to determine management (modeling) zones that were constant for both years. Annual modeling will be influenced by how well these soil-based zones reflect production potential in a given year. The differences in the calibration/prediction outcomes by inverting the role of the two years is indicative of this limitation. Thus, while spatial calibration has shown some benefit in this study, research is still needed to better understand how ancillary data, especially soil data, can be used to generate seasonal-specific zoning for modeling, i.e. the local climate (predicted and/or observed weather conditions) should influence how downscaling methods are applied to account for known or expected local soil-plant-environment interactions.

The choice of calibration and evaluation data had a significant impact on the spatialized APSIM outputs. Thus, regarding the preliminary results of this study and the choice of static soil zones for downscaling, it was more relevant to perform spatial calibration when the target agronomic variable exhibited good spatial structures/patterns. This was expected as calibrating a model with (yield) data that is atypical of the expected response or exhibits a large amount of stochastic variation is not expected to be effective. This result could be generalized to other large and complex crop models.

In other words, the main interests in the spatial calibration are twice. First, using the spatial calibration approach as a spatialization process allows to constrain the spatial pattern of the agronomic variable. allowing to maintain a spatial consistency in the within-field modeling. Second, the spatial calibration is able to manage a trade-off between the accuracy of prediction at changing scales against the ‘noise’ in the available input data at different spatial resolutions. A crop model calibrated individually for each modeling unit (i.e. at the site-scale) may have a significant stochastic error through the calibration process.

4.2.4.2. Sources of uncertainty linked with parametrization

The aim of this study was to investigate if using complex spatialized crop models at the within-field scale was relevant, i.e. if the uncertainty in the modeling process was acceptable to support agronomic decision-making. With the SBA score, the most relevant spatial scale for modeling could be identified for each combination of calibration/prediction years. However, there was a huge source of uncertainty using the spatialized APSIM model, especially with the assumptions made for input parameterization. APSIM is a complex crop model requiring several input parameters to work correctly to reproduce the

biophysical processes implemented. Within-field downscaled modeling with spatial calibration involves an increase in input parameterization, which may not be useful to enhance model performances (Adam et al., 2011; Soltani and Sinclair, 2015; Zhen et al., 2023). Moreover, uncertainties in the downscaled crop model use could mostly be attributed to input data and to the downscaling methods used (Porwollik et al., 2017). Even with a fairly comprehensive data set, many parameters, especially parameters related to soil characterization, needed to be estimated and were not directly measured (e.g. SHL). Furthermore, as the resolution of the spatialization increased (from the whole field to site-specific), the available data also decreased for a given spatial area of prediction, which may introduce higher stochastic variance effects in the model calibration and evaluation. Therefore, higher input parameterization may explain why more uncertainties have been observed.

Cultivar parameters, especially growth parameters, are important parameters for crop models because they drive yield production, but they are also heavily influenced by punctual changes in water and nutrients (Archontoulis et al., 2014; Rötter et al., 2012). For this study, cultivar parameters were considered homogeneous. The assumption of homogeneity for the weather data, especially precipitation, is highly questionable. Precipitation can be variable over even small areas (Krajewski et al., 2003), altering site-specific plant available water and even small changes in slope can impact subsoil water movement and accumulation (Subedi and Fullen, 2009). Although considered flat, there was a 10 m drop in altitude in the field and a trend to higher soil water content in the southern tip has already been noted (Basso et al., 2009). Directly accounting for these variations will improve the soil water representation, soil-cultivar interactions and the APSIM calibration (Hao et al., 2021; Huth et al., 2012). For this study, soil parameters were considered the key spatial parameters to calibrate APSIM, because these parameters are known to be the main local drivers of durum wheat yield. To achieve this, many subsoil parameters were estimated from topsoil information to have a modeled soil profile up to 2 m depth. This introduced uncertainties from the pedotransfer models. Crop models are known to operate more effectively when all required input parameters are measured (Cammarano et al., 2021), thus measuring the needed input soil properties for each soil layer would clearly improve the spatial APSIM calibration. However, in commercial production systems, these subsoil data are difficult to obtain due to cost and time constraints, and pedoclimatic approaches for subsoil information are likely to commonly used in the future to generate these data when needed. Therefore, further improvements in subsoil pedotransfer functions in an obvious starting point for improving crop model spatialization.

Spatial calibration for steps 3 and 4 was made through KL and DUL profile assignments, rather than detailed soil water observations. Again, an accurate calibration of the KL and DUL values for each soil layer at each measurement site was time and cost prohibitive, even for this research study. Profile assignment was considered here to be a good trade-off between computation time and improving modeling crop model performance. A limitation to this approach is that the assigning of a profile does not necessarily assigned the optimum value to minimize the calibration error.

4.2.4.3. Calibration performance using the SBA score

The calibration error of each simulation scale using the SBA scores (Table 4.3) would be interpreted differently compared to the SBA scores for evaluation (Table 4.5). If the relevance of the simulation scale was only based on the calibration error, then field-scale simulation (native footprint) would have been selected for both scenarios and both target variables (biomass and yield) (Table 4.3). In terms of practical use, these results suggest that SBA scores are highly dependent of the spatial structure/pattern of the considered agronomic variable and SBA score interpretation should be taken carefully. Given this is the first attempt to use the SBA to assess spatial calibration of a complex crop model, no strong conclusions can be drawn from the identification of the native footprint as the preferred modeling scale during calibration; however, this is an area where further study is needed. Metrics to best calibrate the model are as important as prediction metrics.

4.2.4.4. Within-field segmentation dependency and uncertainties

Results of this study were very dependent on the zoning performed, i.e. the ancillary data and segmentation algorithm. Ancillary data that measures soil EC_a were chosen to delineate zones because these data are known to be related to soil texture, which in turn is temporally stable and indirectly related to soil moisture holding capacity. The assumption in this pedoclimatic region, and this particular production system, was that soil water is the dominant driver of yield. Nawar et al. (2017) highlighted that using EC_a by itself could be insufficient to quantify the spatial variability in production at the within-field scale and suggested coupling soil spatial properties with data related to crop productivity to have more reliable zones. However, this was not feasible in this study as the available yield and biomass data were part of the calibration/prediction data set and could not be considered as independent data for zone delineation. Within-field zone segmentation was also assumed to be temporally stable in this study because the segmentation was based on temporally stable soil properties. If data related to crop productivity is involved in zone delineation, it may give stable and unstable zoning overtime (multiple years). This is especially true in systems subject to very variable weather conditions. Indeed, data from multiple years will be needed for a reliable spatial calibration using crop productivity data, especially with rain-fed crops. Ideally, some historical data, such as production from previous years with similar evolving seasonal weather, and/or near real-time production data, such as within-season remote sensing images, could help to adapt segmentation during the production year and delineate more relevant zones for in-season crop model simulations (Maestrini and Basso, 2018).

The strength of the spatial structure/pattern of the agronomic variable of interest was identified in this study as a driver of the success of a spatialized downscaling approach. Similarly, the strength and spatial structure/patterning of the ancillary data used in zone delineation, as well as the strength of its correlation with the modeled agronomic variable of interest, will affect the quality of the spatialized model outputs. For example, in this study, soil zones that had the highest predicted yield tended to be overestimating yield. The reason for this was unclear, but it is clear that the model was not accounting for some effect associated with production loss in these zones. In an extreme example, zoning with uncorrelated, poorly structured ancillary data will not generate sensible and relevant solutions beyond chance. As with any other application of zoning in PA, issues in the number and spatial distribution of measured sites and on the zoning methods applied should be carefully tackled (Xu et al., 2020). In studies concerning the use of crop models on a large scale, e.g. regional scale, when downscaling processes are necessary to match input scale models, correction methods are often applied (Ji et al., 2018). This kind of correction for downscaling processes at within-field scale using crop models is still rarely applied. Another limitation of current crop models is that they seldom consider lateral water movement that can greatly affect water stress as experienced by the crop, and this could explain uncertainties in the APSIM spatial calibration in this study. Huth et al. (2012) for APSIM, and Xiang et al. (2020) and Shelia et al. (2018) for DSSAT have respectively improved these crop models to account for better water movements within the soil. These improvement modules could be used in further investigations when using, for instance, APSIM at within-field scale to improve spatial calibration and spatialized model performances.

4.2.5. Conclusion

Results showed that using APSIM at a within-field scale generated more relevant yield predictions than simulating yield at the field-scale when the target variable (durum wheat yield) was spatially structured in the predicted year. Spatial calibration of selected key model parameters allowed APSIM to approximate the spatial pattern of the durum wheat yield. When the target yield was more randomly distributed, APSIM's native spatial footprint, i.e. field-scale, was identified as the most relevant. The spatial calibration of complex crop models, such as APSIM, requires many inputs and assumptions of parameter values that could lead to uncertainties in the simulations. Finer spatial calibration has a need for more observed and sensed data to be collected to drive the calibration/evaluation process, which

could be limiting in commercial situations. The delineation of within-field zones was identified as an area which could also be improved to improve the spatial model calibration. Delineation should take into account various ancillary data types, including crop parameters, such as biomass or yield maps, over multiple, climatically varying seasons so that ancillary data choices can be better targeted to predicted in-season conditions. However, when production variables exhibit a strong spatial patterning, the use of a spatial calibration approach to spatialized a crop models shows promise for within-field simulation at a scale that can support decision support tools to optimize their efficiency and their field management at within-field scale. Further work is certainly needed to validate these preliminary findings in other systems and other pedoclimatic regions.

4.2.6. Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

4.3. General conclusion of the chapter

The aim of this chapter was to investigate if using a complex crop model at the within-field level was relevant to model durum wheat yield. This chapter represented an opportunity to apply the methods and metrics investigated in the previous chapter of this manuscript on a real data set using APSIM. This later is considered here as a complex crop model. The spatial calibration approach for crop models, whose the relevance of use was investigated in Chapter 3, was used to downscale and spatialize the crop model. Additionally, the SBA score as proposed in Chapter 2 was used to evaluate the spatialized crop model performances. Although the SBA score is still new, this metric was used to account for both the aspatial and spatial error from model predictions. Furthermore, the use of this metric was considered relevant as it was employed to obtain a relative comparison of prediction error across modeling scales, which is the primary purpose for which SBAs score was designed. Using this metric in a real case study also permitted the potential usefulness of such a metric to be assessed and to potentially identify improvement needed with the SBA score. Limitations highlighted in Chapter 2 are still valid for the SBA score in this chapter, and the limits and expectations of this metric will be further discussed in the following chapter.

Depending on the calibration/evaluation years used, spatial calibration produced differing results. When the yield in the evaluation year was spatially structured, the spatial calibration approach allowed the crop model to be optimized to produce improved spatialized crop model predictions, i.e. the spatial calibration allowed a reproduction of the spatial pattern of the durum wheat yield. However, when the yield map in the evaluation (predicted) year was relatively homogeneous, a classical calibration approach at the field scale showed to be the best performing model. Thus, when the target variable being modeled produces a relative homogeneous case study, spatial calibration lead to an increase in the calibration uncertainty and to a decrease in the spatialized APSIM performance for modeling, compared to its native spatial footprint, i.e. the field scale. These results corroborated results found in Chapter 3. When an agronomic variable has a strong spatial structure and is correlated with the ancillary data used to spatialize the model, spatial calibration improves modeling performances, i.e. it is relevant to use that approach. However, when the spatial structure of the agronomic variable was weakly structured and/or correlation with the ancillary data is low, a classical calibration approach should be preferred because the spatialized crop model performances is not better than the field-scale, ‘average’ modeling results.

In this presented work, calibration zones were delineated at the within-field scale for spatial modeling purposes. These within-field zones were not defined as management zones, as it is usually seen in the literature of precision agriculture scientific works, they are defined as calibration zones for the modeling. On an operational purpose, management zones are directly linked to the cultural practices

potentially applied at the within-field scale, whereas calibration zones are dedicated to find the most relevant modeling scale to best reproduce the spatial patterns of the considered agronomic variable. Thus, on a technical view, calibration zones are not directly delineated to improve advice regarding cultural practices, rather they are delineated to have spatial pattern that reflects the target agronomic variable (on which potentially a precision agriculture management could be advised in a following step).

The relevance of using spatial calibration in this case study was based on predicting yield at the end of the season using input data, excepting meteorological and management data, obtained at the start of the season. Nevertheless, this spatial calibration approach is the only way of spatializing these models before the season starts and field-specific growth data becomes available to allow for data assimilation approaches to be used. Thus, although this chapter is less theoretical than the previous ones, this work is still relatively theoretical regarding the limit of its capacity to be applied on an operational case study. The main goal of the spatial calibration was to constrain the spatial pattern to enable the calibration process of the crop model. However, this delineation of calibration zones was fixed at the beginning of the production season. It may be that the calibration zones, and the spatialization of the model, may benefit from updating the calibration zones as in-season production data becomes available, i.e. adapting data assimilation into this process by incorporate data during the modeling phase.

Chapter 5

General discussion and perspectives

5.1. Retrospective on the results obtained during the PhD project

In a precision agriculture context, within-field variability of an agronomic variable is of high interest as variability may be seen as an opportunity for site-specific crop management towards more sustainable cultural practices. The spatial pattern of the agronomic variable has to be considered in order to set up differential cultural practices in the most efficient way. Crop models are considered as relevant decision support tools for farmers and agronomists in achieving this objective. However, currently existing crop models are typically designed to model agronomic variables as a point-based model at the field scale. The model equations and parameters were developed from, and designed to be used with, average field-scale data. Thus, within-field variability is not inherently considered within the model construction (i.e. model equations) so sub-field scales cannot be considered as a native spatial footprint for these existing crop models. So, even if applied at spatial scales smaller than the field, there is no guarantee that the biophysical processes will be properly simulated and that the spatial pattern of the agronomic variable will be reproduced.

Regarding these considerations, this PhD project has focused on addressing some scientific questions raised for the effective spatialization of existing crop models to within-field scales using downscaling methods. In particular, it has addressed issues associated with the spatial calibration of crop models for both simple and complex models. It also addressed the question of the evaluation of outputs derived from such spatialized crop models. The specific scientific questions, already formulated in Section 1.3.1.2, are recalled hereafter:

(I) How to perform a relevant evaluation and comparison of spatialized crop model performances across different spatial scales?

- (i) Are the currently used methods to evaluate spatialized crop models performances efficient?*
- (ii) Does taking into account the spatial component of the model outputs improve model evaluation?*

(II) Is the spatial calibration of selected crop model parameters an effective method of downscaling existing crop models to permit modeling at within-field scales?

- (i) Can the delineation of potential management units from high-resolution ancillary data assist with a spatial calibration approach? How does the relationship between the ancillary data and the agronomic variable of interest influence this?*
- (ii) Is there a trade-off between the spatial scale of modeling at within-field scale and the noise brought by the errors of measurement of the data used for the calibration and evaluation step of the spatialized crop models?*
- (iii) Does the complexity of a crop model affect the ability of the model to be downscaled for precision agriculture purposes?*

A review on the spatialization of crop model processes and on the methods to evaluate their performances was first performed. This state of the art analysis allowed an identification of existing methods but also knowledge gaps (Chapter 1). The reasons and the ways that crop models are currently spatialized, either for upscaling or downscaling purposes, were analyzed. This led to a clarification of the existing frameworks used for spatialization, i.e. a clarification on the methods used to change the spatial footprint of existing crop models without changing their internal model structure. It also allowed the identification of metrics, both aspatial and spatial, that have been used in the scientific literature to date to evaluate the performances of such spatialized crop models. The review also helped to identify some limitations in current metrics and potential improvements for the evaluation of spatialized crop model performances.

Then, the main limitations of common metrics used for evaluating spatialized crop model performances were highlighted (Chapter 2). In particular, it was shown that no metric was able to assess simultaneously the aspatial and the spatial components of the outputs from a spatialized crop model. To address these issues, the spatial balanced accuracy (SBA) was proposed, tested and validated to evaluate both aspatial and spatial errors resulting from modeling with such crop models.

In this PhD research, spatialization was performed for the purpose of precision agriculture. In this context, downscaling methods were considered of high interest. Investigations therefore focused on this specific aspect when it deals with a change to the spatial resolution (footprint) of a crop model. A spatial calibration approach was used in the case studies (simulated and real data sets) as a novel approach to downscaling and spatializing two contrasting crop models. By using ancillary data for delineating within-field calibrations zones, the calibration process was performed at various scales so that the most relevant modeling spatial scale, which minimized the spatial prediction uncertainties, could be identified. This is the first attempts to use this spatial calibration approach for downscaling crop models. The first results obtained from this work have demonstrated that this spatial calibration approach, based on within-field calibration zones derived from high spatial ancillary data assumed to be more or less correlated to the agronomic variable of interest, is relevant for downscaling crop models. These results showed that the output of the spatialized crop models remained relevant (spatial structure/pattern and variability). The results also showed that, using spatial calibration improved the (spatialized) modeling performances of the crop models in some cases, but not in all cases. Spatial calibration was most effective, compared to a classical calibration approach, when the target variable exhibited a strong spatial structure. When the variable of interest only had a moderate or weak spatial structure, even if it a high magnitude of variation, a classical calibration approach generated equivalent or better model predictions (Chapter 3).

Both spatial calibration and the newly proposed SBA score were applied to a real case study to investigate the scope and relevance of using such methods on real world data (Chapter 4). The results showed that spatial calibration and the SBA score were able to firstly downscale the model predictions and to identify which spatial modeling scale was the most relevant to reproduce both the aspatial and spatial characteristics of the considered agronomic variable. In this context, these approaches were useful to optimize the performance of the spatialized crop models and evaluate their modeling performances. However, at this stage, a generalization of these results should not be made beyond the case studies and additional research is certainly necessary in this field to test and validate these ideas with other crop models and in other production systems.

Regarding the contribution of this PhD project, the following discussion is structured around how the project results have answered the specific scientific questions raised previously. In a second step, some perspectives to the PhD are proposed for potential future investigations.

5.2. Discussion and contribution for a relevant evaluation and comparison of spatialized crop model performances

5.2.1. The needs for assessing spatialized crop models with a metric based on spatial pattern instead of spatial structure

There are many available and used metrics to evaluate crop models. These metrics could be either aspatial, i.e. evaluating only the numerical error of model outputs, or spatial, i.e. evaluating the spatial structure of model outputs (even if this kind of metric is often used for data description and spatial estimations, such as in a kriging process). However, a ‘combined’ metric that is capable of assessing simultaneously both spatial and aspatial errors was not found in the performed literature review. Concerning the spatial error, in most precision agriculture studies, the focus is made on assessing the spatial structure of the considered agronomic variable, and this has some limitations.

To adapt their within-field management, farmers use the spatial patterns observable in their fields. These spatial patterns allow a relevant and reliable decision-making that reflects real crop needs. Describing the spatial structure, generally using geostatistics, remains an interesting approach to describe an agronomic variable and is widely understood and used in the scientific precision agriculture community. However, any given spatial structure could arise from multiple and varying spatial patterns, which would not need the same level of differential management. Therefore, the use of geostatistical metrics to assess the spatial structure of model predictions as a means for evaluating spatialized crop models could lead to misinterpretations. Instead, the use of spatial pattern metrics (in place of spatial structure metrics) should provide more definitive information to support differential management and provide a suitable evaluation of spatialized crop model performances for precision agriculture purposes. Therefore, the need for a metric able to evaluate simultaneously aspatial and spatial (i.e. based on spatial pattern) prediction errors was identified as a necessary first step, especially for the assessment of the spatialized application of crop models in this PhD project. In terms of spatialization, it was also deemed necessary to have a metric that could identify which spatial scale is the most relevant, i.e. to have a metric that is applicable at several spatial modeling scales and capable of identifying which spatial scale is the most relevant. As identified in the review, a synthetic metric that accounts for both kinds of error and allows for a quick assessment and ranking of model performance among a pool of scale-variable simulations did not exist. The SBA score, as proposed in this work, answers these issues. It is a first attempt of a spatial/spatialized model metric based on connectivity theory. However, this remains a new metric that has only been tested and validated on a limited number of case studies, although simulations have allowed the robustness of the SBA score to be tested in very different conditions. As a new metric, testing and validation on other case studies may highlight some limitations requiring some further improvements. Despite this, it was effective in its, albeit limited, use within this PhD project.

5.2.2. The spatial balanced accuracy: expected limitations and benefits

The SBA score was designed to identify which modeling scale was the most relevant by taking into account both aspatial and spatial errors in model predictions. More specifically, it was designed to compare several spatial modeling scales, using the same crop model, in order to find the most relevant one to carry out the modeling. Thus, the SBA allows a ranking of the spatialized crop model outputs to find the most suitable spatial modeling scale for a considered agronomic variable over a considered area. Nevertheless, at the current stage of development, it is preferable to only use this metric to compare modeling within a considered field. In other words, the SBA is relevant if the purpose is to seek a first evaluation on the modeling. In the case studies considered in this manuscript, the use of the SBA score was possible because it was carried out with modeling from the same crop model, for the same crop and for the same agronomic variable over the same field, but considering different spatial scales. Until now, the SBA has not been tested for other purposes and other potential applications. For instance, its

relevance in assessing different modeling approaches on a same field or different variables was never tested nor validated. As a result, the use of the SBA for such application cannot be promoted since it could lead to a biased or even wrong outcome.

The SBA is, for now, proposed to be used as a relative metric, i.e. it should be used to assess the spatialization of a crop model in a given situation, and not between models or between different scenarios. Thus, depending on the purpose of the study, using complementary quantitative metrics (e.g. RMSE) is still considered to be necessary to further quantify the degree of uncertainty present in the spatialized crop model outputs. This composite metric approach would allow for a finer evaluation of the spatialized crop model performances, e.g. to understand if model outputs are under-performing on the spatial or aspatial aspects. Therefore, the SBA interpretation should be complemented by existing metrics that account for the aspatial or spatial modeling errors. In other words, this proposed metric is not a replacement of existing metrics but rather an addition to the suite of metrics available to researchers and modelers to assess several modeling approaches in the first instance. The results presented in this manuscript (Chapters 3 and 4) have shown that the use of the SBA score brings additional value and information to model evaluation. For example, the distinction that could be made between modeling with similar results of RMSE when only a spatial error is added to refine the evaluation. Thus, its use has to be recommended as a complement to other existing metrics, i.e. the RMSE (or other classical metrics) should not be used without consideration of the SBA score (or some other similar metric).

With this in mind, the new SBA score is, for now, mostly intended to be used by agricultural scientists and researchers. Indeed, the comparison of different spatial modeling scales remains a relatively theoretical aspect of crop modeling. The SBA score is mostly based on a relative adequacy between observed and modeled data. It is probably too soon to consider using this metric for practical and technical uses and for it to be applied to a reliable decision-making process. Broadening the use of this metric to farmers and growers could be interesting to match crop modeling behavior with the footprint of real cultural practices on-farm. For instance, the technical aspects related to agronomic machine footprint is of primary interest for farmers to know if the cultural practices decision is truly applicable with their tools (Lajili et al., 2021; Tisseyre and McBratney, 2008). This could potentially accelerate the use of these spatialized crop models as decision support tools and the adaptation of their management to a new spatial scale that is more relevant. The short investigation into irrigation ‘decision’ zoning using Balanced Accuracy approach (Section 2.2), which facilitate the development of the SBA score, is an example of how this could be operationally achieved. Having more study cases to test the use of the SBA score is paramount and should constitute the continuation of the research and development of metrics to support spatial and spatialized crop model use.

5.3. Discussion and contribution of the results for downscaling existing crop models at within-field scale

5.3.1. Consistency of the spatial calibration approach for the spatialization of crop models at the within-field scale (compared to existing methods of downscaling)

5.3.1.1. Contribution of the spatial calibration approach

A spatial calibration approach has been identified as having a high potential to be used as a downscaling method to spatialize crop models at the within-field scale. This approach is considered as a spatialization method because the internal structure of the crop models remains unchanged during the process of downscaling. Therefore, the fundamental model remains a point-based model. In the workflow of the spatialization process, spatial calibration is used upstream of the crop modeling (i.e. before the formal use of the model) to calibrate the model ‘locally’ based on available input data. As a result, the crop

model can be applied as spatial scales that are finer than the original intent. The spatial calibration can also be performed at any time, pre-season or in-season, using historical field information and/or in-season information. It is the former potential to have spatialized pre-season predictions that has been the driving motivation behind much of the applied work in this PhD project.

The spatial calibration approach was expected to be able to manage a trade-off between the accuracy of prediction at changing scales against the ‘noise’ in the available input data at different spatial resolutions. The field-scale uses smoothed, average field data for model calibration, which the model is explicitly designed for. The site-scale relies on individual point data, which may have a significant stochastic error, to calibrate the model. Thus, there was a general hypothesis that the optimum scale of modeling will be where there is a gain from downscaling by spatial calibration but where the data (point or aggregated) for a given, local spatial calibration is an accurate reflection of the modeling point environment. More specifically, the hypothesis here was that the best modeling scale could be achieved (identified) by generating calibration units, using a general management unit methodology common in precision agriculture, and using aggregated data within the calibration units spatially calibrate the model within-field. However, results presented in this manuscript did not necessarily show a systematic improvement when using management units compared to field or site-specific modeling scales. The size of the considered field and the number of measurement sites available to describe the spatial pattern of the agronomic variable appeared to have an impact on the relevance of spatial calibration approach, especially for the vineyard water stress modeling (49 measured sites in 1.2 ha). The higher the number of measurements, the more the spatial pattern can be represented and, to a large degree, this is dependent on the amount of production variability, rather than the field size. Limitations with available point data may be partially ameliorate when measurement sites are cleverly positioned in regards to the considered spatial pattern. Note that in this study, for the vine water stress modeling, this was not the case and a regular grid was used to obtain the site-scale data.

5.3.1.2. Spatial parameter selection

A main point of the spatial calibration approach is the selection of the model parameters that are most likely to have a significant impact on the spatialized crop model performance. In this thesis, most of the parameters considered as spatial were related to soil characteristics, consequently only a few parameters were assumed to drive the process spatially. These soil parameters were known to be spatially variable even at the within-field scale and could have a real impact on the modeling, thus, this seemed to be a relevant working assumption. Additionally, historical, high resolution spatial soil information was available for many soil properties that are temporally invariant in the short to med-term (e.g. soil texture, soil water holding capacity, etc.). Other model parameters related to microclimate or nutrient absorption (interaction between soil and plant) could also have an impact on the spatialized crop models. However, obtaining high-resolution data to use in the calibration process for precision agriculture purposes for some of these model-driving factors is currently difficult. The actual trend in precision agriculture is the research and the development of various sensors and probes at within-field scale (Mouazen et al., 2014; Tardaguila et al., 2021), which may provide new information sources in the future. When such information does become routinely available, it will be possible to enlarge the number of model parameters to be spatially calibrated. Having more input data and more parameters to calibrate should be really beneficial as it could avoid assumptions on variables that are not usually measured and it should increase reliability of crop models. However, depending on the quality of the measured data, there is always a risk that parametrization could be adversely affected by the measurement characteristics of the sensor (i.e. high stochastic noise). A strategy that could be implemented is to use several sensors at the within-field scale, even if individual sensors are not of ‘perfect’ quality, to better capture the spatial pattern of the considered ancillary data, under the assumption that the higher number of available data could offset the higher uncertainty from any single sensor.

If the homogeneity of crop model parameters could be an acceptable working assumption in a context of homogenous species, the heterogeneity of parameters is likely to increase in the context of

agroecological transition. Indeed, agroecological systems highlight the use of multispecies over a same field leading to interactions between species that may affect crop model performances. For example, interactions between trees and crops in agroforestry systems and subsequent changes in crop behavior have already been highlighted in the literature (Luedeling et al., 2016). This complexity could represent a limitation of spatialized crop models compared to the building of ‘true’ spatial crop models (i.e. crop models taking into account neighboring spatial modeling units to compute the value on a considered spatial modeling unit). Since this kind of multi-species interaction has not yet been considered in most existing crop models, the spatialization of such interactions could be very challenging in future agroecological applications.

5.3.1.3. Spatial calibration approach and data assimilation methods

The proposed spatial calibration approach is a method with a close objective to other, existing data assimilation processes that are often advocated for downscaling crop models (Jin et al., 2018). Data assimilation is a general term that encompasses a wide variety of methods for adjusting crop model parameters based on observed in-season spatial data (usually remotely-sensed biomass information) and the simulated model state variables. It may be done via a data assimilation calibration (DAC), a forcing or an updating methodology (Jin et al., 2018). The DAC is analogous to the spatial calibration approach being advocated in this work, i.e. it seeks to adjust the initial parameters of the model to achieve better spatialized modeling. The difference between the spatial calibration proposed in this PhD project and the DAC is important. The spatial calibration proposed here did not influence any state variables in the model, the calibration is only based on the considered agronomic variable (and potentially other parameters related to phenology or soil parameters) (Launay and Guerif, 2005; Shuai and Basso, 2022). Moreover, DAC is iterative and retrospective in fixing parameters using in-season information, i.e. it cannot be done at the start of the season. The forcing and updating approaches are primarily concerned with adjusting the model’s state variable(s) to reflect in-season observed data. All these methods allow ancillary data of high spatiotemporal resolution to be used to ‘correct’ crop models and consequently increase their prediction accuracy. Data assimilation is often used at large scale (regional, international scales) with low-resolution remote sensing data. However, for precision agriculture purposes, i.e. focusing on within-field scale, proximal and remote data with a higher resolution should be used.

A strength of data assimilation is the accurate temporality allowed for the estimation of crop growth status. Having observed data that is related to a model state variable at multiple times during the season allows modelers to adjust the model multiple times during the season, with the intent of improving the end-of-season prediction. However, this temporal adjustment is not a guarantee to have a consistent spatial pattern of the agronomic variable, because there is no direct spatial constraint on the crop model adjustments that optimize the spatialized crop model outputs. Thus, there remains an open question when using data assimilation regarding the scale of modeling, i.e. the DAC/forcing/updating scale that should be employed. Should the finest scale be selected for modeling the agronomic variable? Or an intermediate spatial scale between the finest scale and the field scale? By default, data assimilation processes are usually applied at the spatial scale of the acquired ancillary data, but there is little justification for this in the literature at the moment. Undeniably, spatial calibration and data assimilation are not mutually exclusive and can obviously complement each other regarding the temporal characteristics of data assimilation and the spatial characteristics of spatial calibration. Compared to the data assimilation method, spatial calibration aims to identify the most suitable spatial scale to model the agronomic variable. Thus, an in-season recalibration could be allowed by coupling spatial calibration (pre-season) and data assimilation (in-season).

The relevance of using a spatial calibration approach is also related to the spatial pattern of the agronomic variable at the desired modeling date. However, it is not always easy or even possible to have spatial information (or spatial variability) of agronomic variables early in the production season and spatial patterns are known to be temporally dynamic for many agronomic traits. In the presented work of this manuscript (Chapter 4), the relevance of using, or not, a spatial calibration approach was based

on the already measured data from a previous production year. Thus, estimating spatial pattern for in-season production would have been very challenging and this point needs more investigation. Moreover, in this case-study, between the two considered production years, durum wheat yield spatial patterns were not the same, leading to a different preferable approach to better modeling the considered agronomic variable. Therefore, the temporality of the spatial pattern is an important part of the understanding needed to really use spatial calibration for prediction in an operational context. For an operational and technical use, the spatial calibration approach should be applied and preferably followed by a continuous in-season recalibration to better adapt the decision-making based on this spatial pattern temporality.

5.3.2. Dependencies of the spatial calibration performances

The proposed zonal approach to spatial model calibration approach was found to be highly dependent on the segmentation and zone delineation that was performed at the within-field scale. The criterion on which this segmentation was based was selected to try to faithfully reproduce the spatial pattern of the agronomic variable. Two issues with zone delineation that affected crop model spatialization were clearly highlighted by the studies presented in Chapter 3 and 4. These issues are (i) the choice of algorithm used to define management units and (ii) the choice of ancillary data used to generate the calibration zones.

5.3.2.1. Dependence of the spatial calibration regarding the ancillary data characteristics

When using ancillary data, questions regarding its temporal characteristics need to be considered. The choice of ancillary data that is representative of the spatial pattern of the considered agronomic variable is not a trivial process. The effectiveness of the spatial calibration is based on the spatial pattern at the target date of modeling, which represents a spatiotemporal issue. That is, the relevance of using a spatial calibration approach is based on the calibration zones delineated at the beginning of the growing season but ultimately dependent on the spatial pattern of the considered agronomic variable at the end of the season (which often is the target date of modeling). However, this prediction is not always easy or even possible and it could constitute a drawback to using spatial calibration in an operational context. Ancillary data will have different temporalities along which their spatial patterns can be considered stable. Ancillary data with a relatively stable spatial pattern (e.g. soil data) could be used for several years. In contrast, other ancillary data have changing spatial patterns depending on the production year or even the timing within a production season (e.g. plant-related variables). Segmentation should reflect the objective of the spatialized crop model use, especially depending on whether tactical or strategic uses are being considered for the output. The temporality problem is mainly related to the need to obtain an accurate (and relevant) segmentation that reflects the spatial pattern of target agronomic variable at the target prediction date. Data assimilation is identified as an interesting methodology to tackle this issue with in-season recalibration (via forcing or updating) which can better adapt the potentially changing spatial pattern of the agronomic variable.

The correlation between the selected ancillary data and the modeled agronomic variable is also an impacting factor to consider. In Chapter 3, the effect of the level of correlation between the ancillary data available for calibration zone delineation and the agronomic target was investigated. It revealed that ancillary data should be highly correlated to obtain a relevant spatial calibration approach (the agronomic variable should also present a significant spatial structure). In reality, high levels of correlation between ancillary and production data are not usually achieved in agriculture. For this preliminary study, only three levels of correlation were tested, representing a very high, medium and very low correlations. Further work is needed to more specifically test the level of correlation needed to achieve an effective level of calibration zoning for spatial calibration approaches, and if the level identified in simulation studies corresponds to expected correlations in real-world situations.

5.3.2.2. Dependence of the spatial calibration regarding the number of ancillary data

In the results of this PhD project, ancillary data have been shown to have an impact on the spatialized crop model performances. High-resolution data remain proxies that do not capture the full complexity of agricultural systems. Consequently, a single high-resolution dataset can only account for one component as the plant (e.g. remote sensing), the soil (e.g. soil conductivity) or water circulation (e.g. topometry). Any combination that takes better account of the complexity of pedoclimatic contexts should also be considered. In order to make the right choice of variables adapted to the problem and the context, agronomic expertise should also be taken into account at this level. For that reason, several ancillary data, relatively correlated with the considered agronomic variable, should be used at the same time to have a more accurate segmentation of the field. This may correct the potential spatial pattern bias with only one ancillary data and even increase the resulting correlation between the merging of several ancillary data and the agronomic variable. Thus, depending on the purpose of the modeling, a better estimation of the spatial pattern of the agronomic variable will likely be achieved after collecting a high amount of different data (even if in some specific cases only one ancillary data layer may be relevant).

In Chapter 2 and 4, several sources of ancillary data were used for calibration zone delineation, although no form of sensitivity analysis on these data layers was performed (only the number of calibration zones). Using multiple, well-chosen data sources (e.g. vine size using trunk circumference and soil texture using EC_a) appeared to be relevant for calibration zone delineation and supported the points made in the previous paragraph.

5.3.2.3. Dependence of the spatial calibration regarding zoning method

The within-field zoning method also had an impact on the performance of the spatial calibration approach. Although, for all the presented works here, only one segmentation algorithm was used, the prevalence between the impact of the selected ancillary data and the selected segmentation algorithm is an important point to highlight. Regarding the results for the experimental vineyard used in the previous chapter, the algorithm was not able to delineate more than five calibration zones. However, a more optimal spatial modeling scale could be potentially found with a higher level of segmentation (or even with a different delineation realized for a segmentation or classification between 2 and 5 zones). This drawback may be attributed to the characteristics of ancillary data available. The choice of the algorithm to delineate within-field calibration zones is therefore decisive and must certainly be chosen according to the characteristics of the available data and the crop model itself. Note that only a few algorithms are available to delineate zones using irregular data (i.e. data which are not collected or interpreted onto a defined regular grid). Moreover, using raw data can raise questions about the robustness and reliability of the segmentation if the data themselves are noisy. Various studies have shown that data classically used in precision agriculture can lead to different interpretations and spatial structures and consequently potentially different delineations of within-field zones (Zhang, 2023). This type of analysis was outside the scope of this work, but this aspect deserves attention and, for now, the role of different types of algorithms for zoning and/or classifying ancillary data sets into calibration zones remains an area for further investigation.

Related to the increasing number of sensors at the within-field scale, data fusion was identified as being important to delineate relevant calibration zones (Mouazen et al., 2014). Some studies have also questioned the static characteristics of calibration (management) units in precision agriculture and have highlighted the possibility to delineated dynamic zones based on remote sensing data and high-resolution sensors (Cao et al., 2017; Fontanet et al., 2020). At this point of research, in this manuscript, it is difficult to give a crisp answer on the relative importance of zoning (relative to the choice of data used) on spatial calibration performances, except to say that it does need to be done properly if spatial calibration is to be successful implemented.

Delineation of within-field zones using for the spatial calibration are defined as calibration zones and not as management zones, usually defined for precision agriculture purposes. Thus, there does not necessarily need to be agreement between the delineated calibration zones and the potential management zones already defined or potentially applicable for within-field differential management. However, the delineation of management units for site-specific crop management in precision agriculture is still currently under research even 20+ years after the first algorithms were introduced (Bell et al., 1995; Corwin and Lesch, 2005; Zhang and Taylor, 2000). Thus, realizing relevant management units is still an open question in precision agriculture. Similarly, realizing relevant calibration zones will raise many of the same issues and is far from a trivial problem with universal way to proceed at the moment. However, the number of management units can only be truly validated after the production season, which is often too late to know if the choice was correct. An advantage of calibration zone delineation is that their relevance can be more easily estimated. Indeed, the choice of the number of calibration zones can be addressed by modeling, if the right metrics are used to evaluate the scaled model prediction. Basing the segmentation on both the operational constraints (e.g. swathing widths) and the representative spatial pattern of the agronomic variable should improve the use on this methodology in decision support tools for practical use by farmers. The spatialized model outputs, that are generated based on the calibration zones, could be used as a data set for the delineation of management zones (together with other ancillary data).

5.4. Perspectives for further research

5.4.1. How to improve spatialized crop models evaluation?

The proposed SBA score in this manuscript was used for the purpose for which it was designed, namely to compare modeling from the same crop model at different spatial modeling scales. The SBA score also allows for a relevant evaluation of the aspatial and spatial pattern error. However, it is a first generation metric to achieve this. Improving the SBA score is of primary interest to be able to better assess the performances of the spatialized crop models at the within-field scale.

A first working perspective would be to improve the interpretation of the assessments using the SBA. These works will need to first investigate the suitable definition range of the metric allowing for a better interpretation. Knowing if the interpretation is relevant as an absolute metric between different modeling from different case studies or if the SBA score should be interpreted only as a relative metric for a specific, defined situation is of primary interest. This improvement of the interpretation of this metric will need more research by using simulated data. Simulation of a large number of fields with various characteristics will be needed to be able to properly understand how to interpret all the properties of the metric. These simulations should also be useful to provide modifications of this metric as necessary. Another potential work perspective will be to transform the relatively qualitative SBA score into a quantitative metric (i.e. like a RMSE). Aspatial metrics are interesting because they give a direct indication of the uncertainty involved over predicted values. Thus, from an operational point of view, this type of evaluation is able to conclude if the crop model can meet the specifications required for a considered decision (e.g. quantity of water or nitrogen application). However, the current SBA score is not able to give this information. This improvement will allow a more detailed interpretation to directly quantify the error in the spatialized crop model output(s) and eventually the error regarding the proposed decision.

5.4.2. How the spatial calibration approach could be improved to increase the performances of the spatialized crop model?

Throughout the work realized during this PhD project, the spatial calibration approach has been applied as it was designed, i.e. the same method to delineate calibration zones was used as well as the same frequentist approach to calibrate crop models. For the purpose of this work, the spatial calibration

approach was used as a downscaling method for modeling agronomic variables at the within-field scale. Thus, the objective was not directly to obtain the best modeling but to show that using a spatial calibration approach could be beneficial to use existing crop models at the within-field scale. For that reason, a segmentation algorithm that was easy to use (and applicable to the considered ancillary data) and a basic frequentist approach were chosen. However, it has been shown in this work that segmentation algorithm and ancillary data may have a considerable impact on the spatial calibration performances. Thus, the performances of the spatial calibration approach are still to be perfected and perspective future work should also focus further on finding methods that increase its performances. Regarding the delineation of calibration zones, as highlighted previously, this is still an important question within the precision agriculture community. That is why, other segmentation/classification algorithms should be tested in perspective works. In addition, delineation based on other constraints (e.g. operational constraints, farmer knowledge) should also be considered and tested. For the calibration approach to estimates crop model parameter values, several studies have applied machine-learning algorithms instead of frequentist approaches. Often, these machine-learning approaches are more time consuming than frequentist methods. However, some studies have shown that Markov chain Monte Carlo simulations could generate better parameter estimations for crop model calibration. Perspective future work will focus on the improvement of the performance of the spatial calibration approach by applying different within-field calibration zones delineation algorithm and using other methods of parameter estimation.

In this thesis, all the data sets, that were not simulated data, were measured for other studies for other, different purposes. On one hand, it was a great opportunity to realize this PhD project by profiting from this valuable data and having real data sets that are representative of data that could be usually obtained in a precision agriculture context. However, it is recognized that these data were not collected with the objective of model calibration and that this will have introduced some limitations. For example, some measurements could also have been made with the aim to make the spatial calibration approach more efficient or allow to have more feedback on its performance. A lot of ancillary data used in these studies were related to soil variables. These variables were really useful to spatialize crop models because soil characteristics are known to present a spatial variability even at within-field scale (Hao et al., 2021; Huth et al., 2012). However, increasing the amount of measured ancillary data could allow more diversity in the choice at the type of ancillary data used to delineate the calibration zones. Thus, another future perspective could be to investigate how to design an acquisition campaign using several sensors at the within-field scale to obtain more ancillary data that can potentially improve the spatial calibration approach (or give more detailed feedback on its application). Additionally, increasing the number of measurements at the within-field scale could allow a better representation of the spatial patterns of the considered agronomic variable. Further testing of the theory of spatial calibration will likely need even richer, denser data sets than those used here but beyond this, there also lies a question of how the spatial calibration approach can be pared back to be relevant to typical, on-farm data sets that may not be as rich or as dense as those used for research purposes.

An underlying question of the PhD project was to investigate if spatialized crop models could be used to model agronomic variable at the within-field scale instead of using 'true' spatial crop models, i.e. more detailed models that incorporate spatial processes. In the results presented here, spatialized crop models showed relatively good performances once the best spatial modeling scale was selected. However, it was far beyond the scope of this thesis to directly determine the performance of a spatialized vs a spatial crop model. Some studies have already used crop models coupled with hydrologic models to have a better spatial representation of the modeling, especially related to soil water, and to get closer to the concept of a 'true' spatial crop models. These coupling were used at rather large scales (watershed or territorial scales) by using upscaling process. There is a remaining question and the potential for perspective work to compare modeling performances of both spatialized and spatial crop models at the within-field scale. This is predicated on spatial crop models being available. These future works could show if existing crop models via spatialization processes are still great candidates to be used in a

precision agriculture context. Future works could also be related to the agroecology transition, with the development of multispecies systems, to investigate if crop models are still relevant in this context. Perspective works could also investigate if the use of corrective methods (often used for climate model estimations) for downscaled crop model prediction are preferable to a spatial calibration approach to obtain reliable agronomic variable modeling at the within-field scale.

5.4.3. How the spatialized crop models could be involved in more operational decision-making in a precision agriculture context?

The developed methodology for spatial calibration presented in this manuscript was mainly designed and applied in a theoretical context. This calibration approach was, however, validated on real case studies. The aim of this work was not to directly propose an operational decision-making system but rather to address scientific questions related to the application of crop models at a different spatial scale to the one that they were originally designed for. Developing an operational methodology to be incorporated into a decision support tool dedicated to farmers to help them in their cultural practice scheduling remains completely in the scope of precision agriculture purposes. Thus, making spatial calibration a routine and robust method in the use of spatialized crop models for agricultural decision support is also an envisaged future objective.

Most crop models take into account a temporal dimension to model the evolution of crop systems over time. This is an important aspect for the farmer's tactical decisions (e.g. irrigation, fertilization, sanitary management). However, this temporal dimension still presents some questions when it comes to spatialize crop models, especially regarding the spatial calibration approach. As it was implemented, the spatial calibration approach cannot be used directly in a fully operational context dedicated mainly to farmer support. Indeed, the temporality of this method has raised several questions regarding its relevance on the date of prediction of the considered agronomic variable. Accurate information about spatial patterns at a prediction date is, in most cases, complicated or even impossible to obtain for a prediction method that is applied at the beginning of the season. A perspective area of work is to improve the temporality of the spatial calibration approach. Data assimilation approaches have been identified as having a real potential to improve the spatial calibration approach due to their ability to 'force' or 'update' the modeling process, in regards to some identified state variables, to improve the crop model performance. Thus, developing a method based on both spatial calibration and data assimilation could potentially allow an in-season recalibration to better define the most relevant within-field calibration zones for the considered agronomic variable and related tactical decision. Such 'dynamic' calibration zones could be used to propose cultural managements that best meet crop needs at the right time, allowing for a more efficient and sustainable crop management. Thus, coupling spatial calibration with the temporal aspects of data assimilation can open up research directions towards the design of spatialized crop models fully consistent with the issues raised by precision agriculture.

However, to be fully used in an operational context, spatial calibration needs to closer from to decision-making that is spatially applied at the within-field scale. For that, spatial calibration should be based not only on calibration zones, but also on management zones, characteristic of the real within-field management practices that can be applied by the farmers. Another work perspective is to improve the operational use of the spatial calibration to find an optimum between calibration and management zones that allows for a reliable and feasible crop management schedule at the within-field scale.

Chapter 6

Résumé étendu de la thèse

6.1. Introduction, contexte, positionnement et questions scientifiques du projet de thèse

6.1.1. Etat de l'art sur la spatialisation des modèles de culture et sur l'évaluation de leur performance prédictive

Les agrosystèmes sont des systèmes complexes dans lesquels plusieurs interactions biotiques et abiotiques se combinent pour influencer la production agricole. Les modèles de culture jouent un rôle clé dans la simplification et la compréhension des systèmes agronomiques complexes. Ces modèles sont couramment utilisés pour conseiller les agriculteurs, les décideurs politiques et les chercheurs en agronomie dans leur processus décisionnel (Jones et al., 2017a; Silva and Giller, 2020). Cependant, tous ces acteurs ne sont pas nécessairement intéressés par la modélisation aux mêmes échelles spatiales car leurs décisions n'ont pas la même empreinte spatiale, ce qui soulève la question de changement d'échelle spatiale dans l'utilisation de ces modèles (Ewert et al., 2011).

La spatialisation, dans ce projet de thèse, est définie comme l'application d'un modèle de culture à une échelle spatiale différente de son échelle native. Les modèles résultant de cette spatialisation sont donc considérés comme des modèles spatialisés. Les modèles de culture spatialisés utilisés sont couramment utilisés à des échelles spatiales supérieures que l'échelle de leur empreinte spatiale native notamment pour prédire des scénarios dans des contextes de changement climatique ou pour estimer l'impact d'un changement de pratiques culturales sur les productions agricoles (Asseng et al., 2014; Hoffmann et al., 2016; Zhao et al., 2015). Plus particulièrement, les processus de spatialisation de réduction d'échelle sont identifiés pour ce projet comme une opportunité d'utiliser les modèles de culture existants qui ont été conçus pour une utilisation à l'échelle "moyenne" de la parcelle à d'autres échelles de simulation plus fines (intra-parcellaire) sans modification de la structure interne du modèle. Les modèles de culture sont donc considérés dans ce projet comme des outils « boîte noire ». Dans ce projet de thèse, une attention particulière a été accordée aux modèles de culture mécanistes car ils permettent une meilleure compréhension des processus biologiques, physiologiques et physiques associés aux variables agronomiques modélisées par rapport aux autres types de modèles. En effet, la question du changement de résolution spatiale est particulièrement pertinente pour les modèles de culture mécanistes, car les équations décrivant les processus biophysiques des cultures sont généralement conçues à l'échelle parcellaire.

Les modèles de culture spatialisés par réduction d'échelle doivent être capables de maintenir une cohérence spatiale de la variable agronomique modélisée à l'échelle intra-parcellaire. La performance de ces modèles repose non seulement sur la précision de prédiction des valeurs numériques de la variable agronomique considérée mais aussi sur la capacité à reproduire le motif spatial de cette variable agronomique. Les métriques classiquement utilisées pour évaluer leur performance ne tiennent pas compte explicitement des caractéristiques spatiales des données modélisées, ce qui peut conduire à des biais dans l'interprétation de la qualité des modèles de culture spatialisés. Par conséquent, il devrait être nécessaire de développer des métriques qui tiennent compte à la fois des erreurs aspatiales et spatiales des sorties des modèles spatialisés pour améliorer la pertinence de leur évaluation.

6.1.2. Positionnement et questions scientifiques du projet de thèse

Même si les erreurs potentielles liées à la réduction d'échelle sont encore incertaines, la communauté agronomique reconnaît qu'il existe une réelle opportunité d'adapter les modèles de cultures existants dans des outils d'aide à la décision (OAD) pour aider les agriculteurs dans la mise en place de pratiques en vue de l'agriculture de précision, c'est-à-dire à cette échelle intra-parcellaire (Cheng et al., 2023).

Les modèles de culture existants sont bien documentés, bien reconnus par la communauté agronomique et sont connus pour fonctionner correctement s'ils sont utilisés de manière pertinente. Ainsi, repenser la manière d'utiliser ces modèles devrait permettre de tirer avantage des performances reconnues de ces modèles ponctuels à l'échelle de la parcelle pour des modélisations à d'autres échelles spatiales plus fines. Cela permettrait ne de pas avoir recours à la construction de nouveaux modèles. En effet, l'alternative serait de concevoir des modèles de culture « réellement » spatiaux où les unités de modélisation spatiales interagiraient et auraient un impact sur la modélisation des unités de modélisation spatiale voisines. Cette approche assurerait la reproduction des interactions spatiales potentielles à l'échelle intra-parcellaire, mais nécessiterait un changement fondamental dans la structure et les équations sous-jacentes du modèle de culture, ce qui représente un lourd processus. Pour cette raison, la spatialisation des modèles de cultures est considérée dans ces travaux comme une approche à considérer pour les changements d'échelle spatiaux. L'influence des sources de variation stochastique dans les entrées des modèles de culture est supposée devenir plus importante lors de la réduction d'échelle. Néanmoins, l'effet et les limites de cette variabilité stochastique sur les sorties des modèles spatialisés sont encore mal connus. Un compromis entre la finesse de la modélisation et le « bruit » dans les données d'entrée à différents niveaux spatiaux est supposé exister. De plus, pour être en mesure de modéliser à des échelles intra-parcellaire, les modèles de cultures spatialisés devront généralement s'appuyer sur des données auxiliaires à haute résolution pour améliorer leurs performances, en particulier dans l'étape de calibration.

L'évaluation des performances de ces modèles de culture spatialisés doit pouvoir être menée en comparant des résultats de modélisation à différents niveaux d'échelle spatiale. Cependant, des métriques pertinentes pour pouvoir évaluer ces performances sont nécessaires et n'ont pas été identifiées dans la littérature scientifique. Ces potentielles métriques permettraient notamment d'évaluer la pertinence d'utiliser les modèles de culture existants à des échelles intra-parcellaires et d'identifier l'échelle de simulation la plus performante. Sur ce constat, la problématique de recherche générale de ce projet de thèse est donc appuyée sur :

La spatialisation des modèles de cultures existants, par l'utilisation de processus de réduction d'échelle spatiale, est-elle envisageable et pertinente pour leur utilisation à des échelles intra-parcellaires ?

L'objectif de ce projet de thèse était de mener une première analyse et de développer des outils pour la spatialisation des modèles de culture existants à l'échelle intra-parcellaire à des fins d'agriculture de précision. Une première étape théorique a été jugée nécessaire pour établir les bases de l'évaluation des performances des modèles de culture spatialisés (mécanistes). Les façons d'évaluer ces modèles de cultures à différentes échelles devaient être investiguées, et éventuellement repensées, pour tenir compte à la fois des erreurs aspatiales et spatiales. Une seconde étape a été de s'interroger sur la pertinence de la réduction d'échelle des modèles de culture par l'utilisation d'une approche de calibration spatiale. Plusieurs études de cas ont été menées dans ce sens pour tester différentes hypothèses associées à la spatialisation de modèles de culture. L'intention n'était pas de tirer des conclusions générales sur la spatialisation des modèles de culture, mais de formaliser ce concept dans un contexte d'agriculture de précision et de construire une base pour de futures recherches pour une utilisation tactique et à l'échelle intra-parcellaire de ces modèles plutôt que sur leur utilisation stratégique telle qu'actuellement pratiquée. Ainsi, par rapport à la problématique de recherche générale et les points soulignés ci-dessus,

deux questions scientifiques principales ont été identifiées et détaillées en plusieurs sous-questions scientifiques plus spécifiques :

(I) Comment effectuer une évaluation et une comparaison pertinentes des performances des modèles de culture spatialisés à différentes échelles spatiales ?

(i) Les méthodes actuellement utilisées pour évaluer les performances des modèles de culture spatialisés sont-elles efficaces ?

(ii) La prise en compte de la composante spatiale des sorties du modèle améliore-t-elle son évaluation ?

(II) La calibration spatiale de paramètres spécifiquement sélectionnés du modèle de culture est-elle une méthode efficace de réduction d'échelle des modèles existants pour permettre la modélisation à l'échelle intra-parcellaire ?

(i) La délimitation des unités de gestion potentielles à partir de données auxiliaires à haute résolution peut-elle aider à une approche de calibration spatiale ? Comment la relation entre les données auxiliaires et la variable agronomique d'intérêt influence-t-elle cela ?

(ii) Existe-t-il un compromis entre l'échelle spatiale de modélisation à l'échelle intra-parcellaire et le bruit apporté par les erreurs de mesure des données utilisées pour l'étape de calibration et d'évaluation des modèles de culture spatialisés ?

(iii) La complexité des modèles de culture augmente-t-elle la difficulté de réduction d'échelle de leurs prédictions ?

Les jeux de données utilisés au cours de cette thèse avaient déjà été acquis en amont pour des études antérieures. Le premier jeu de données utilisé est lié à des observations spatio-temporelles de l'état hydrique de la vigne (par mesure de potentiels hydriques foliaires : Ψ_{PD}) au sein d'un vignoble expérimental de Syrah (*Vitis vinifera* cv. Syrah) de 1,2 ha à Pech Rouge (43.144°N, 3.131°E, Gruissan, Aude, France) (Acevedo-Opazo, 2003). Le second jeu de données utilisé concerne une parcelle de blé dur (*Triticum durum* Desf. cv. Gargano) expérimentale de 12 ha de la CREA (Centre de recherche sur les cultures céréalières et industrielles) près de Foggia (41.462°N, 15.506°E, Italie). Les travaux réalisés ont été conduits avec deux modèles de culture. Leur capacité à modéliser les variables agronomiques mesurées dans les jeux de données et leur niveau de complexité différent ont été un choix déterminant dans leur sélection (la complexité étant définie par le nombre d'équations internes et le nombre de données d'entrée nécessaires). Les modèles retenus ont été WaLIS (Water balance for Intercropped System) (Celette et al., 2010) comme modèle de culture simple pour modéliser Ψ_{PD} et APSIM (Agricultural Production Systems sIMulator) (Holzworth et al., 2014) comme modèle de culture relativement plus complexe pour modéliser le rendement du blé dur.

Le premier chapitre décrit le contexte général (Article 1) et les questions scientifiques spécifiques abordées dans le projet de thèse. Les recherches présentées dans la suite du manuscrit sont divisées en deux thématiques principales. La première thématique aborde les concepts théoriques liés à une évaluation pertinente des performances des modèles de culture spatialisés. La pertinence et la fiabilité de conduire cette évaluation par l'utilisation de métriques aspatiales et spatiales classiquement utilisées ont été investiguées (Article 2). Une nouvelle métrique conçue pour évaluer les modèles de culture spatialisés par l'estimation simultanée de l'erreur aspatiale et de l'erreur basée sur le pattern spatial a aussi été proposée (Article 3). La seconde thématique aborde des concepts plus pratiques. Un chapitre est consacré à l'utilisation d'une méthode de calibration spatiale afin d'étudier et de mieux comprendre la pertinence d'utiliser une telle approche sur des modèles de cultures (Article 4). Le second chapitre de cette partie est une étude transversale qui aborde la spatialisation et l'utilisation d'un modèle de culture complexe à des échelles intra-parcellaires (Article 5). Le dernier chapitre est consacré à une

discussion et aux perspectives de l'utilisation future de modèles de cultures spatialisés à des échelles intra-parcellaires à des fins d'agriculture de précision.

6.2. Parvenir à une évaluation pertinente de la performance des modèles de culture spatialisés lors de la simulation de variables agronomiques à des échelles variables

6.2.1. Objectif du chapitre

Le but de ces travaux de recherche était double. Le premier objectif a été de montrer que les métriques actuellement utilisées pour l'évaluation, qu'elles soient aspatiales ou spatiales, ne sont pas nécessairement les plus pertinentes pour l'évaluation des performances des modèles de culture spatialisés. Ces travaux s'appuient sur des données réelles d'un vignoble expérimental dans un contexte opérationnel de conduite de la viticulture de précision. Sur la base des résultats de ces travaux préliminaires, une nouvelle métrique, le Spatial Balanced Accuracy (SBA) a été proposée. Cette métrique est capable de prendre en compte l'erreur aspatiale et l'erreur basée sur le pattern spatial dans les sorties de modèle de culture spatialisés (et plus globalement des modèles spatiaux). La pertinence de cette métrique a été testée sur des données simulées.

6.2.2. Comparaison de différentes métriques pour évaluer la performance des modèles de culture spatialisés à différentes échelles intra-parcellaires

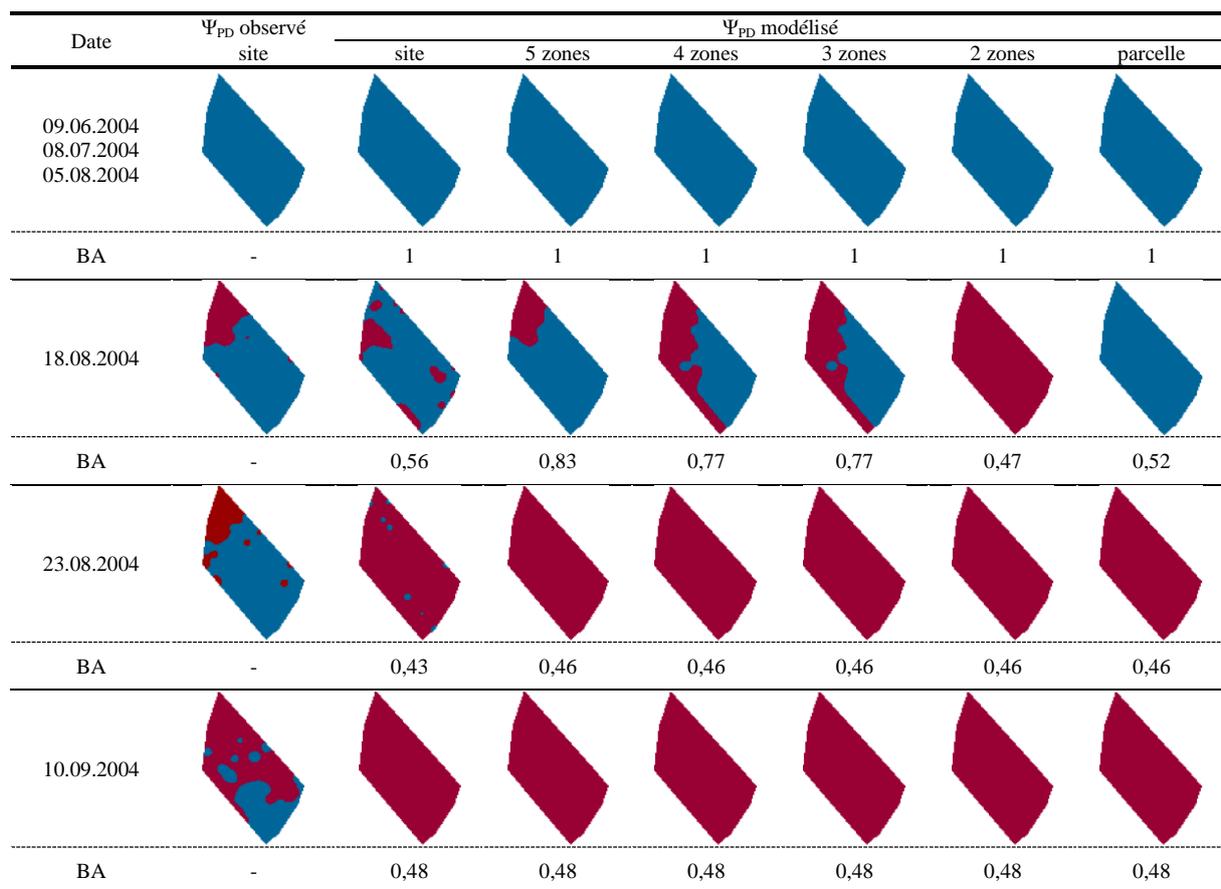


Figure 6.1 Evaluation en contexte opérationnel de cartes d'irrigation intra-parcellaire en fonction de la date et de la valeur du potentiel hydrique foliaire (Ψ_{PD}) résultant des données observées et des données modélisées à différentes échelles spatiales. BA = Balanced Accuracy. Résultats extraits de Pasquel et al (2022b).

Les métriques aspatiales (RRMSE et D-index) et spatiales (Cambardella index et Z-score) testées ne sont pas révélées être les plus pertinentes pour l'évaluation des modèles de culture spatialisés. En effet, l'interprétation de chacune de ces métriques a indiqué des échelles de modélisation spatiales différentes comme étant la plus performante. Néanmoins, l'évaluation en contexte opérationnel, via l'utilisation de la Balanced Accuracy s'est révélée être un bon compromis pour une évaluation plus pertinente des modèles de culture spatialisés (Figure 6.1).

6.2.3. Présentation de la métrique Spatial Balanced Accuracy (SBA)

La construction d'une métrique permettant d'évaluer à la fois l'erreur aspatiale et l'erreur basée sur les motifs spatiaux de la variable considérée s'est avérée être primordiale. Le Spatial Balanced Accuracy (SBA) a été proposé. Cette métrique repose principalement sur une adaptation de l'analyse de connectivité (Figure 6.2), méthode habituellement utilisée dans le domaine de l'hydrologie (Koch et al., 2017). Cette métrique (i) permet une évaluation pertinente des modèles de culture spatialisés par la prise en compte des erreurs aspatiales et spatiales, (ii) est basée sur une approche automatisable et robuste, (iii) permet l'identification de la modélisation la plus performante parmi un ensemble de modélisation à différentes échelles spatiales et (iv) peut être utilisée quel que soit le contexte agronomique ou l'échelle spatiale de modélisation considérés. L'évaluation des motifs spatiaux plutôt que de la structure spatiale des variables considérées a été jugé pertinente car l'application des pratiques agricoles pour l'agriculture de précision se réalise sur les motifs spatiaux observables. La structure spatiale des variables étant principalement réservée à une caractérisation et une synthèse de l'information apportée par les motifs spatiaux (ainsi une même structure spatiale peut donner plusieurs motifs spatiaux différents).

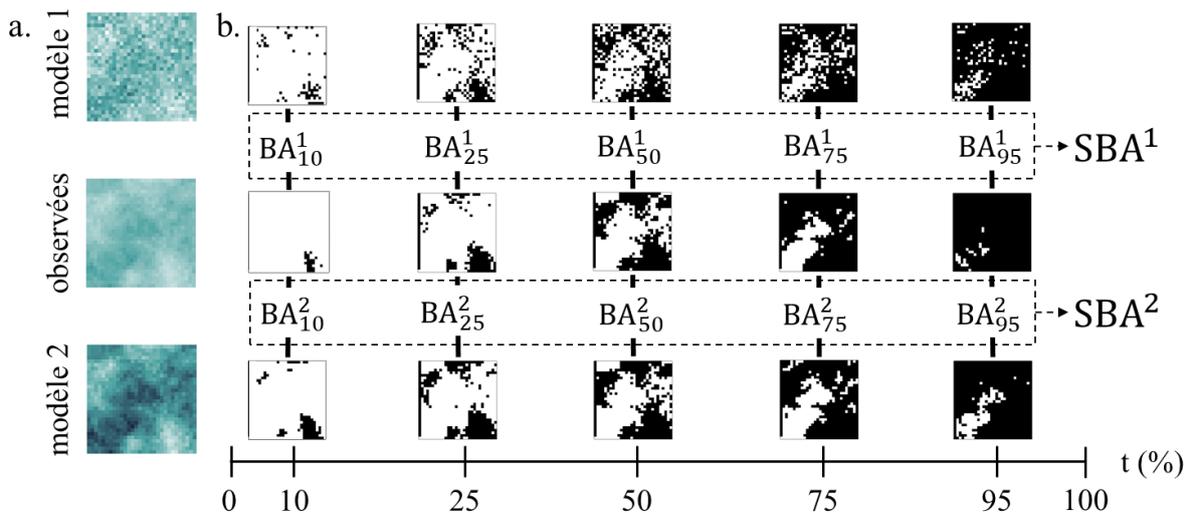


Figure 6.2 Représentation du calcul de la Spatial Balanced Accuracy (SBA). (a.) Cartes d'une variable agronomique observée et modélisée résultant de différents modèles de culture spatialisés. (b.) Cartes binaires générées pour les cartes observées et modélisées calculées pour les seuils des 10e, 25e, 50e, 75e et 95e centiles sur toute la plage de données observées et modélisées. BA est calculée pour chaque seuil de centile considéré (valeurs en indice) et SBA est calculée pour chaque modélisation (valeurs en exposant). Schéma extrait de Pasquel et al (2023).

6.2.4. Validation de la métrique SBA

La comparaison avec d'autres métriques a été faite afin de montrer la pertinence du SBA pour l'évaluation des modèles de culture spatialisés (Tableau 6.1). Cette comparaison a été menée sur des données simulées. Le SBA a été la seule métrique capable d'identifier de manière claire (une seule modélisation identifiée) la modélisation la plus performante. Ainsi, ces travaux ont montré que le SBA était capable de prendre en compte les motifs spatiaux, contrairement au RMSE et au RMSE_{vario},

respectivement basés sur l'évaluation de l'erreur aspatiale et sur l'évaluation de la structure spatiale. Mais le SBA peut également prendre en considération l'erreur aspatiale si les motifs spatiaux des modélisations sont identiques, contrairement au $RMSE_{con}$.

Tableau 6.1 Comparaison de différentes métriques pour identifier la modélisation théorique la plus performante. Modification par rapport au données observées : (A) ajout d'un bruit aléatoire, (B) ajout du même bruit de manière positive par rapport aux données observées, (C) ajout du même bruit de manière positive par rapport aux données observées avec ajout d'un biais numérique, (D) translation du motif spatial de 10 pixels et (E) données modélisées ayant la même structure spatiale mais des motifs spatiaux différents. Chaque valeur indiquée a été calculée entre les données observées et modélisées. Résultats extraits de Pasquel et al (2023).

Donnée	Modèle A	Modèle B	Modèle C	Modèle D	Modèle E
RMSE	9.90	9.90	26.63	14.10	17.90
$RMSE_{vario}$	106.49	317.87	317.87	21.67	9.14
$RMSE_{con}$	0.16	0.05	0.05	0.17	0.10
SBA	0.18	0.04	0.17	0.37	0.50

6.3. Focus une approche de calibration spatiale comme processus de réduction d'échelle pour spatialiser les modèles de culture

6.3.1. Objectif du chapitre

Les travaux de recherche présentés dans ce chapitre visaient à étudier la pertinence d'utiliser une approche de calibration spatiale en fonction des caractéristiques d'une variable agronomique considérée. L'objectif était ici de comprendre si une calibration spatiale pouvait améliorer et optimiser les performances des modèles de culture spatialisés. Cette pertinence a été étudiée en fonction des caractéristiques de la donnée agronomique : (i) sa structure spatiale, (ii) sa variabilité et (iii) sa corrélation avec une donnée auxiliaire associée qui a servi à spatialiser (zoner) la parcelle à différentes échelles spatiales. Du fait de la contrainte technique de ces travaux nécessitant beaucoup de parcelles différentes avec des caractéristiques propres identifiées, l'étude a été en partie menée sur des données simulées.

6.3.2. Présentation de la méthode de calibration spatiale comme processus de réduction d'échelle des modèles de culture spatialisés

La méthode de calibration spatiale considérée dans ces travaux vise, en plus de l'ajustement des paramètres du modèle de culture afin de réduire l'écart entre les données observées et modélisées, à réduire l'échelle de prédiction du modèle à l'échelle intra-parcellaire. Cette méthode nécessite différentes étapes (Figure 6.3). Premièrement, des données auxiliaires, dont les motifs spatiaux sont représentatifs de ceux de la variable agronomique considérée, sont utilisées pour réaliser des zones de calibration intra-parcellaires. La délinéation de ces zones de calibration repose sur l'utilisation d'un algorithme de segmentation. Au cours des travaux menés, afin de montrer la pertinence de la méthode de calibration spatiale considérée, plusieurs niveaux de zonage intra-parcellaire ont été réalisés afin d'identifier le plus performant pour la modélisation de la variable agronomique. A la différence d'un processus de calibration classique (c'est-à-dire à l'échelle des sites de mesure ou de la parcelle), qui ne prendrait pas en considération les caractéristiques spatiales des données, la calibration spatiale proposée ici contraint spatialement le processus de calibration des paramètres du modèle afin d'assurer la reproduction des motifs spatiaux de la variable agronomique.

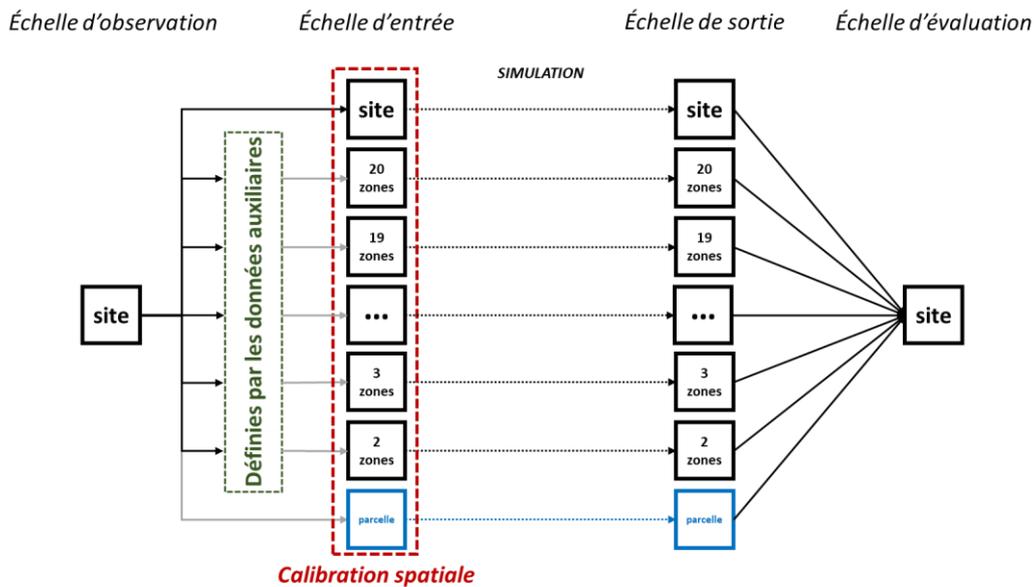


Figure 6.3 Approche de calibration spatiale pour la réduction d'échelle des modèles de culture pour la prédiction intra-parcellaire d'une variable agronomique. Différentes échelles de modélisation sont considérées : l'échelle d'observation des sites de mesure, des échelles intermédiaires intra-parcellaire (ici ce 2 à 20 zones) définies avec des données auxiliaires. L'empreinte spatiale native du modèle de culture est représentée en bleu et correspond à l'échelle de la parcelle. L'échelle du site de mesure correspond à l'échelle d'observation d'origine. Les flèches grises correspondent au processus de changement d'échelle associé à l'agrégation des données observées à des échelles spatiales supérieures en tant que données d'entrée du modèle. La calibration spatiale est effectuée à cette échelle d'entrée. Schéma extrait de Pasquel et al (2022b).

6.3.3. Quels sont les déterminants de la performance d'un modèle de culture spatialisé à l'échelle intra-parcellaire par processus de réduction d'échelle ?

L'approche de calibration spatiale améliore les performances du modèle de culture spatialisé par rapport à une approche de calibration classique (à l'échelle de la parcelle ou des sites de mesure), en particulier lorsque la variable agronomique présente une forte structure spatiale et une corrélation élevée avec les données auxiliaires (Figure 6.4). L'utilisation d'une approche de calibration classique ou d'une approche de calibration spatiale étaient dans la plupart des cas équivalentes (en particulier pour les variables agronomiques à faible variabilité). Cependant, une attention particulière doit être portée lorsque la variable agronomique est très variable et que la structure spatiale est modérée ou faible. Dans ces cas, la calibration spatiale pourrait détériorer de manière significative les performances du modèle de culture par rapport à une approche classique, de sorte qu'une approche de calibration classique devrait être préférée dans ces situations.

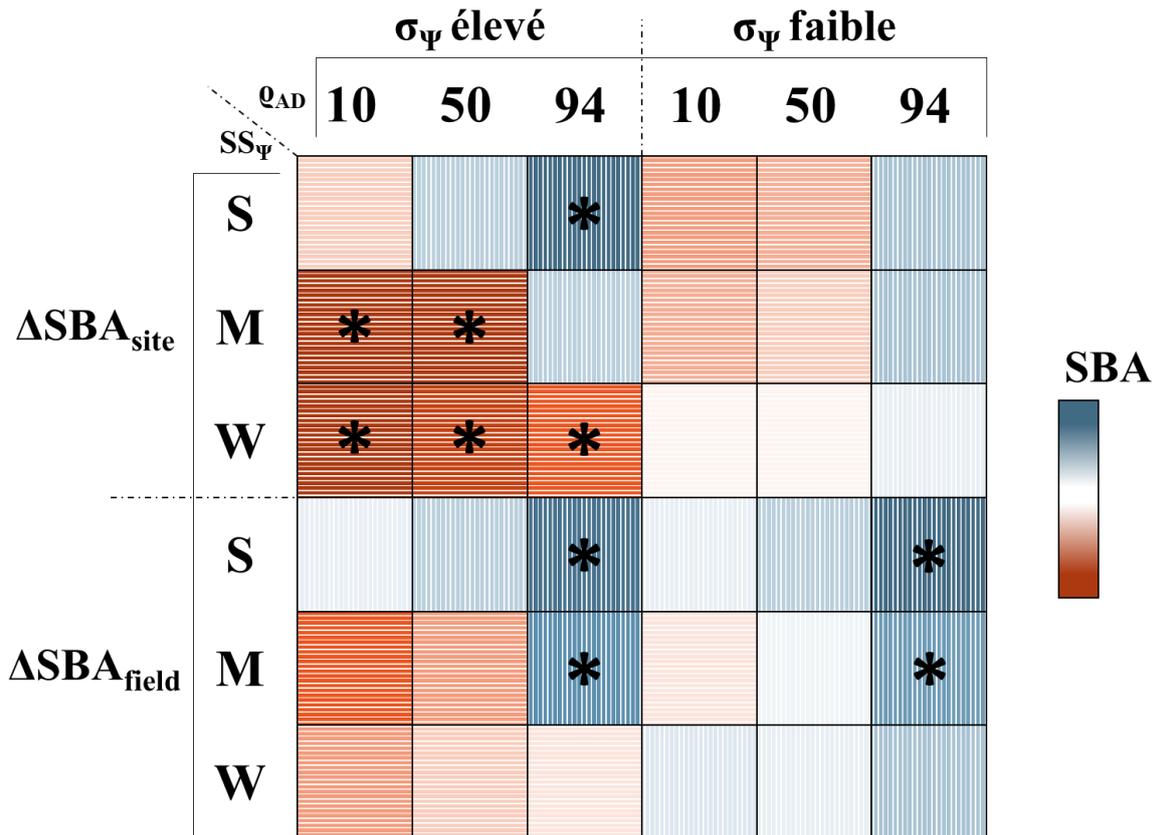


Figure 6.4 Pertinence de l'utilisation d'une calibration spatiale par rapport à une calibration classique par rapport à la différence de SBA pour différentes échelles spatiales de modélisation avec WaLIS. La significativité des différentes modalités a été estimée avec un test de Tukey (modalités avec des lettres différentes) et correspond à une amélioration ou une détérioration significative du modèle de culture spatialisé (représenté par *). σ_{Ψ} correspond à la variabilité de Ψ_{PD} , S, M et W correspondent aux modalités de structure spatiale SS_{Ψ} (respectivement forte, modérée et faible), 10, 50 et 94 correspondent à la corrélation des données auxiliaires ρ_{AD} (respectivement 10%, 50% et 94%).

6.4. Effet de l'échelle de simulation sur la pertinence d'utilisation des modèles de culture spatialisés par réduction d'échelle à un niveau intra-parcellaire

6.4.1. Objectif du chapitre

L'objectif était d'agrèger tous les travaux antérieurs réalisés dans le cadre du projet de thèse et de les appliquer à une étude de cas réelle. Ce chapitre vise à présenter un travail de recherche sur l'étude du comportement d'un modèle de culture complexe utilisé avec une méthode de réduction d'échelle pour simuler une variable agronomique à l'échelle intra-parcellaire. Ceci dans le but de voir si l'utilisation d'un tel modèle de culture (ici APSIM) représente une opportunité pour la prédiction dans un contexte d'agriculture de précision (dans la cadre de la modélisation de rendement de blé dur).

6.4.2. Utilisation du SBA et de l'approche de réduction d'échelle par calibration spatiale appliquées à une cas d'étude sur un modèle de culture complexe

Selon l'année de calibration, l'utilisation de l'approche de calibration spatiale s'est avérée plus ou moins pertinente. Lorsque le rendement de blé dur de l'année d'évaluation était spatialement structuré, l'approche de calibration spatiale a permis d'optimiser les performances du modèle de culture spatialisé.

En effet, les contraintes spatiales associées à la calibration spatiale ont permis de reproduire la structure spatiale du rendement du blé dur. Cependant, lorsque le rendement de l'année d'évaluation était relativement homogène, une approche de calibration classique à l'échelle de la parcelle a montré la modélisation la plus performante (Figure 6.5). Ainsi, dans un cas relativement homogène, la calibration spatiale a augmenté l'incertitude de calibration et a conduit à une diminution de la performance de APSIM par rapport à une modélisation à son empreinte spatiale native, c'est-à-dire à l'échelle de la parcelle. Ces résultats corroborent les résultats précédemment obtenus dans le manuscrit. En effet, il a été montré qu'avec une variable agronomique ayant une forte structure spatiale et des données auxiliaires corrélées, la calibration spatiale pouvait améliorer les performances de modélisation, c'est-à-dire qu'il était pertinent d'utiliser cette approche. Cependant, lorsque la structure spatiale de la variable agronomique est faible, une approche de calibration classique doit être préférée car les performances du modèle de culture spatialisé sont meilleures.

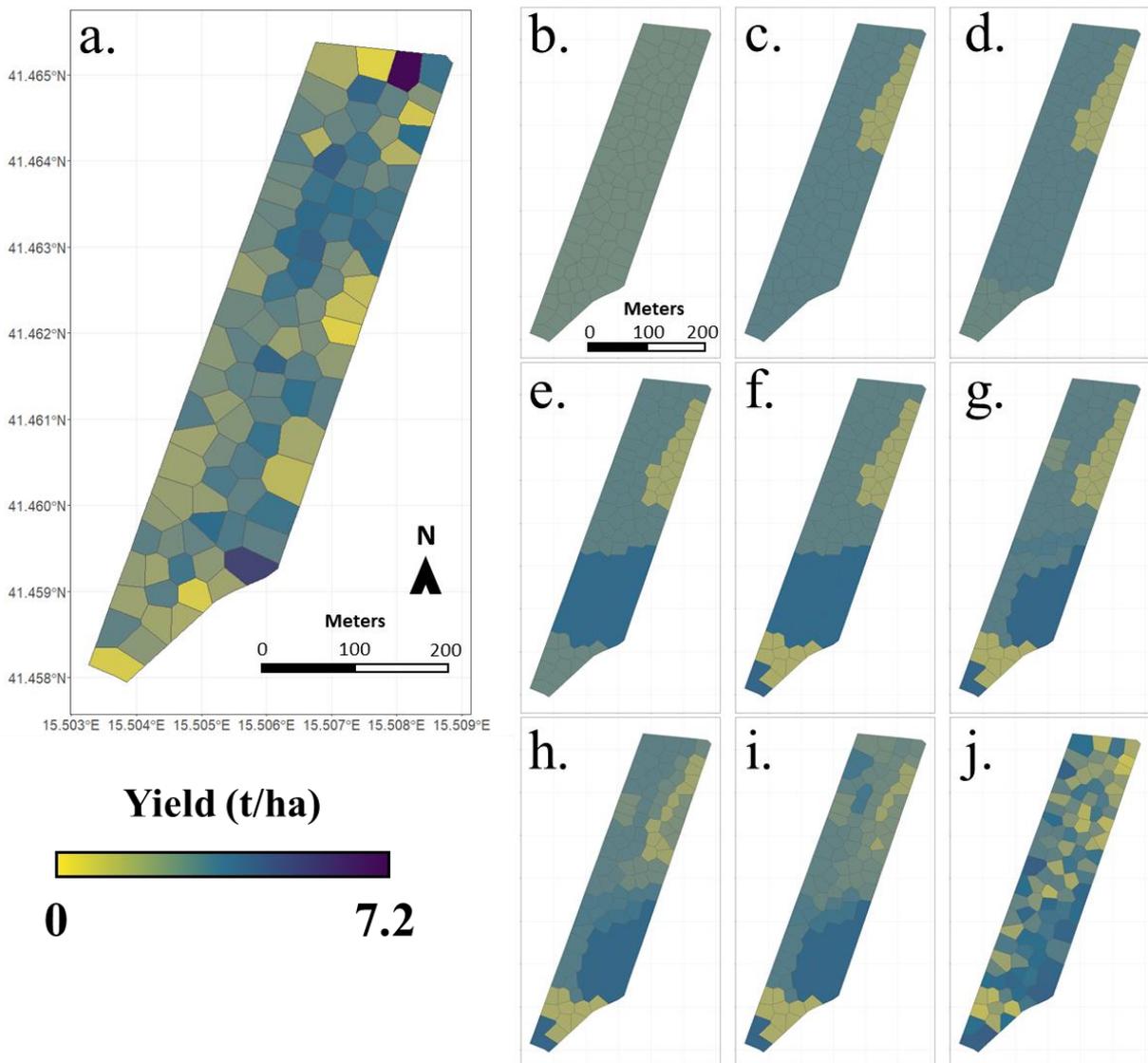


Figure 6.5 Exemple de résultats obtenus avec la version spatialisées d'APSIM. Motifs spatiaux du rendement du blé dur pour (a.) les données observées de 2007-2008 et les données modélisées par APSIM calibré sur les données de 2005-2006 pour différentes échelles de modélisation : (b.) parcelle, (c.) 2 zones, (d.) 3 zones, (e.) 4 zones, (f.) 5 zones, (g.) 10 zones, (h.) 15 zones, (i.) 20 zones et (j.) sites de mesure.

6.5. Discussion et perspectives

Une première approche a permis d'établir les connaissances existantes et les lacunes concernant les processus de spatialisation des modèles de culture et des méthodes d'évaluation des performances des modèles spatialisés. Une clarification des méthodes utilisées pour réduire l'échelle de prédiction de ces modèles a donc été apportée. Cette approche a également permis d'identifier les métriques, à la fois aspatiales et spatiales, qui sont utilisées dans la littérature scientifique à ce jour pour évaluer les performances de tels modèles.

Les métriques identifiées dans la littérature se sont révélées imparfaites pour évaluer les performances des modèles de culture spatialisés. Dans ce but, le Spatial Balanced Accuracy (SBA) a été développé. Il s'agit d'une métrique plus pertinente et robuste pour l'évaluation des modèles de culture spatialisés. Il se base sur les erreurs aspatiales définies à partir des motifs spatiaux. Plus précisément, le SBA a été conçu pour comparer plusieurs échelles spatiales de modélisation en classant les sorties du modèle spatialisé pour trouver l'échelle spatiale la plus appropriée pour estimer une variable agronomique avec la meilleure résolution possible. Néanmoins, au stade actuel de développement, il est préférable de n'utiliser cette métrique que pour comparer des modélisations provenant d'une même parcelle. Ainsi, le SBA est, pour l'instant, proposé comme une métrique relative, c'est-à-dire qu'il doit être utilisé pour évaluer la spatialisation d'un modèle de culture dans une situation donnée, et non entre différents modèles ou différents scénarios. L'amélioration du SBA est d'un intérêt primordial pour pouvoir juger des performances des modèles de culture spatialisés à l'échelle intra-parcellaire. Une première perspective de travail serait d'améliorer l'interprétabilité du SBA (comme l'étude de la plage de définition). Une autre perspective consistera à transformer le SBA relativement qualitatif en une métrique quantitative. Cette amélioration permettra d'avoir une interprétation plus détaillée permettant de quantifier directement l'erreur des modèles de culture spatialisés.

Une approche de calibration spatiale a été identifiée comme ayant un fort potentiel pour être utilisée comme méthode de réduction d'échelle pour spatialiser les modèles de culture à l'échelle intra-parcellaire. Les premiers résultats issus de ces travaux ont permis d'identifier la pertinence d'utiliser une telle approche par rapport aux caractéristiques de la variable agronomique d'intérêt (structure/motifs spatiaux et variabilité) et sa corrélation avec les données auxiliaires utilisées pour délimiter les zones de calibration. Les résultats ont montré que dans certains cas, l'utilisation de la calibration spatiale améliore les performances de modélisation des modèles de culture spatialisés. La pertinence d'une approche de calibration spatiale est liée aux motifs spatiaux de la variable agronomique à la date de modélisation souhaitée. Cependant, il n'est pas toujours facile ou même possible d'avoir des informations spatiales de ces variables au début de la saison de production et les motifs spatiaux sont connus pour être changeant dans le temps pour de nombreuses variables agronomiques, ce qui soulève un problème de temporalité de la méthode de calibration proposée. Une méthode fréquemment utilisée pour la réduction d'échelle des modèles de culture est l'assimilation de données. Une force de cette méthode est la temporalité précise autorisée pour l'estimation de l'état de croissance des cultures. Avoir des observations liées à l'évolution d'une variable d'état des cultures à plusieurs reprises au cours de la saison devrait permettre de réajuster le modèle au cours du temps dans le but d'améliorer la prédiction. Indéniablement, la calibration spatiale et l'assimilation de données ne s'excluent pas mutuellement et pourraient même se compléter. L'assimilation de données permettant un réajustement temporel grâce à leur capacité à « forcer » ou « mettre à jour » le processus de calibration pendant le processus de modélisation et la calibration spatiale permettant de contraindre spatialement la modélisation par la définition de zones de calibration propre à la variable agronomique considérée. Une perspective de travail est donc d'améliorer la temporalité de la méthode de calibration spatiale par l'assimilation de données. Ainsi, le couplage de la calibration spatiale avec l'aspect temporel de l'assimilation des données pourrait rendre les modèles spatialisés de cultures pleinement cohérents avec les enjeux soulevés par l'agriculture de précision et redéfinissant des zones de calibration plus proches des besoins de la culture.

L'approche de calibration proposée pour les modèles spatialisés s'est avérée fortement dépendante de la segmentation et de la délimitation des zones à l'échelle intra-parcellaire. Le critère sur lequel repose cette segmentation est d'un intérêt primordial pour obtenir la meilleure modélisation possible en reproduisant le plus fidèlement les motifs spatiaux de la variable agronomique. Cela repose principalement sur le choix de l'algorithme utilisé pour définir les zones de calibration et sur le choix des données auxiliaires utilisées. La délimitation des zones intra-parcellaires, est toujours source de questionnement au sein du domaine de l'agriculture de précision. C'est pourquoi, d'autres algorithmes de segmentation seront utilisés dans les futurs travaux entrepris. Lors de l'utilisation de données auxiliaires, les questions concernant leurs caractéristiques temporelles doivent également être prises en compte. Certaines données auxiliaires peuvent être considérées stables sur plusieurs années après leur mesure (données de sol). En revanche, d'autres données auxiliaires ont des motifs spatiaux changeant en fonction de l'année de production (variables liées aux cultures) voire au cours d'une même saison. Ainsi, la segmentation devrait être effectuée sur des données auxiliaires ayant des temporalités différentes ce qui pourrait être testé dans de prochains travaux.

Un autre point important de l'approche de calibration spatiale est la sélection des paramètres du modèle qui sont les plus susceptibles d'avoir un impact spatial significatif sur les performances du modèle de culture spatialisé. Dans ce manuscrit, la plupart des paramètres considérés comme spatiaux étaient liés aux caractéristiques du sol. Ces paramètres sont connus pour être spatialement variables même à l'échelle intra-parcellaire et peuvent avoir un réel impact sur la modélisation. Néanmoins, dans le but d'améliorer les performances de la calibration spatiale, un travail en perspective pourrait être de concevoir une campagne d'acquisition de données utilisant différents capteurs à l'échelle intra-parcellaire pour obtenir plus de données auxiliaires plus variées (microclimat, disponibilité des nutriments). De plus, augmenter le nombre de mesures à l'échelle intra-parcellaire permettrait une meilleure représentation des motifs spatiaux de la variable agronomique considérée (ou des données auxiliaires avec lesquelles elle est corrélée).

Dans ce manuscrit, la calibration spatiale et le SBA ont permis d'identifier quelle échelle de modélisation spatiale était la plus pertinente pour reproduire à la fois les caractéristiques aspatiales et spatiales de la variable agronomique à modéliser. Ces approches ont été utiles pour optimiser les performances des modèles de culture spatialisés et évaluer leur performance. Cependant, à ce stade, étant donné qu'il s'agit d'une première étude, une généralisation des résultats ne doit pas être faite au-delà des études de cas traitées dans le cadre de ce document de thèse. Des recherches complémentaires restent nécessaires pour être en mesure de préconiser leur utilisation sur un plus grand nombre de cas.

L'utilisation des travaux de spatialisation menés dans ce projet ont pour objectif d'être à terme, utilisés de manière opérationnelle. Cependant, pour satisfaire cet objectif, la calibration spatiale doit se rapprocher de la prise de décision appliquée à l'échelle intra-parcellaire. Pour cela, la calibration spatiale devrait s'appuyer non seulement sur des zones de calibration mais aussi sur des zones de gestion caractérisant la véritable gestion intra-parcellaire appliquée par les agriculteurs. Ainsi, une autre perspective serait d'améliorer l'utilisation opérationnelle de la calibration spatiale afin de trouver un optimum entre les zones de calibration et celles de gestion permettant un plan de gestion des cultures fiable et réalisable à l'échelle de la parcelle.

Appendix A: Supplementary data of the Chapter 3

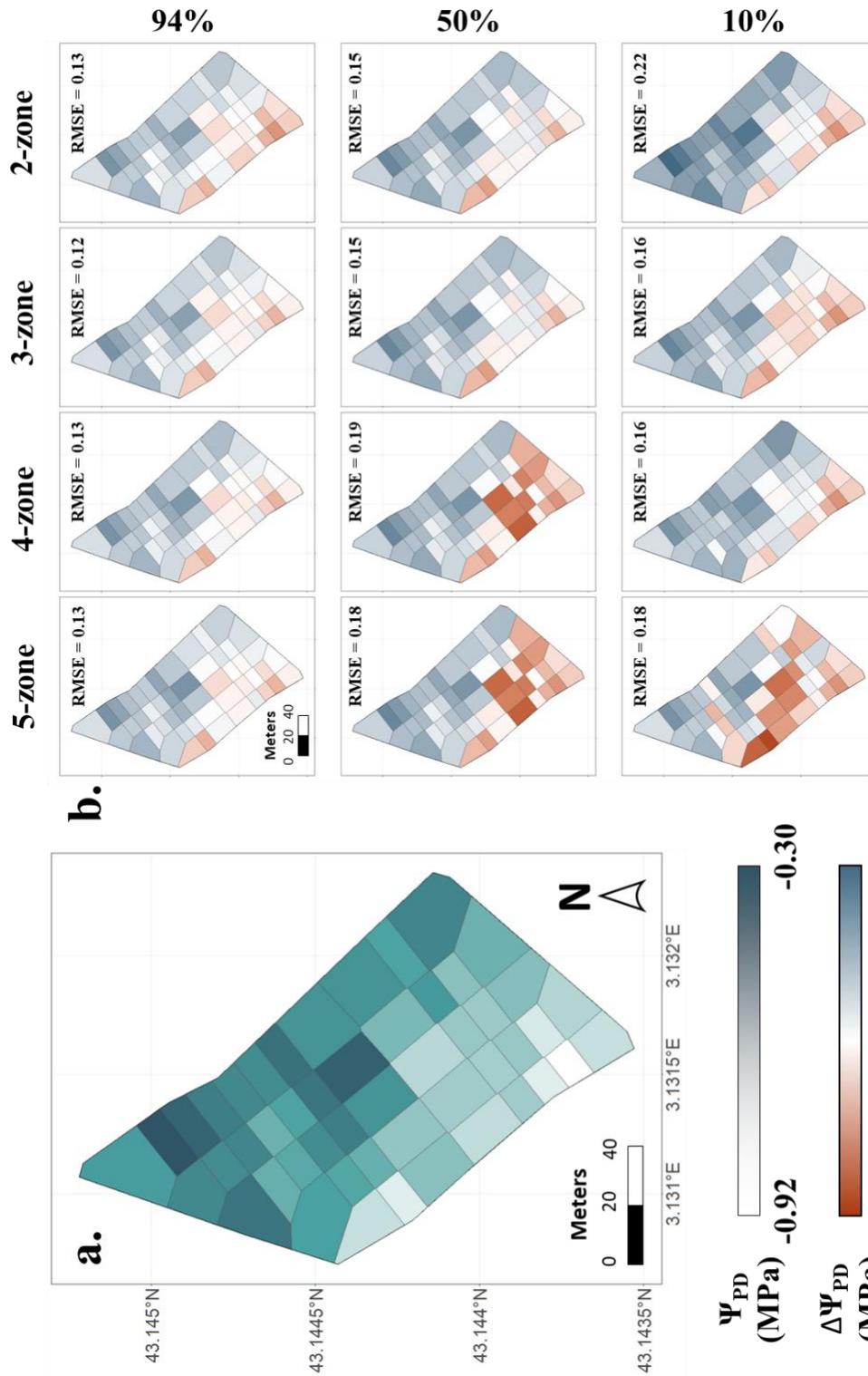


Figure S.A1 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a *strong* spatial structure of Ψ_{PD} between the (a.) observed (synthetic) Ψ_{PD} and (b.) modeling at different within-field spatial scales from 2-zone to 5-zone scales (columns) for each different considered level of correlation between Ψ_{PD} and the synthetic ancillary data (rows; 10 - 94%). The presented modeling date was based on observed data from August 18th with a high Ψ_{PD} variability.

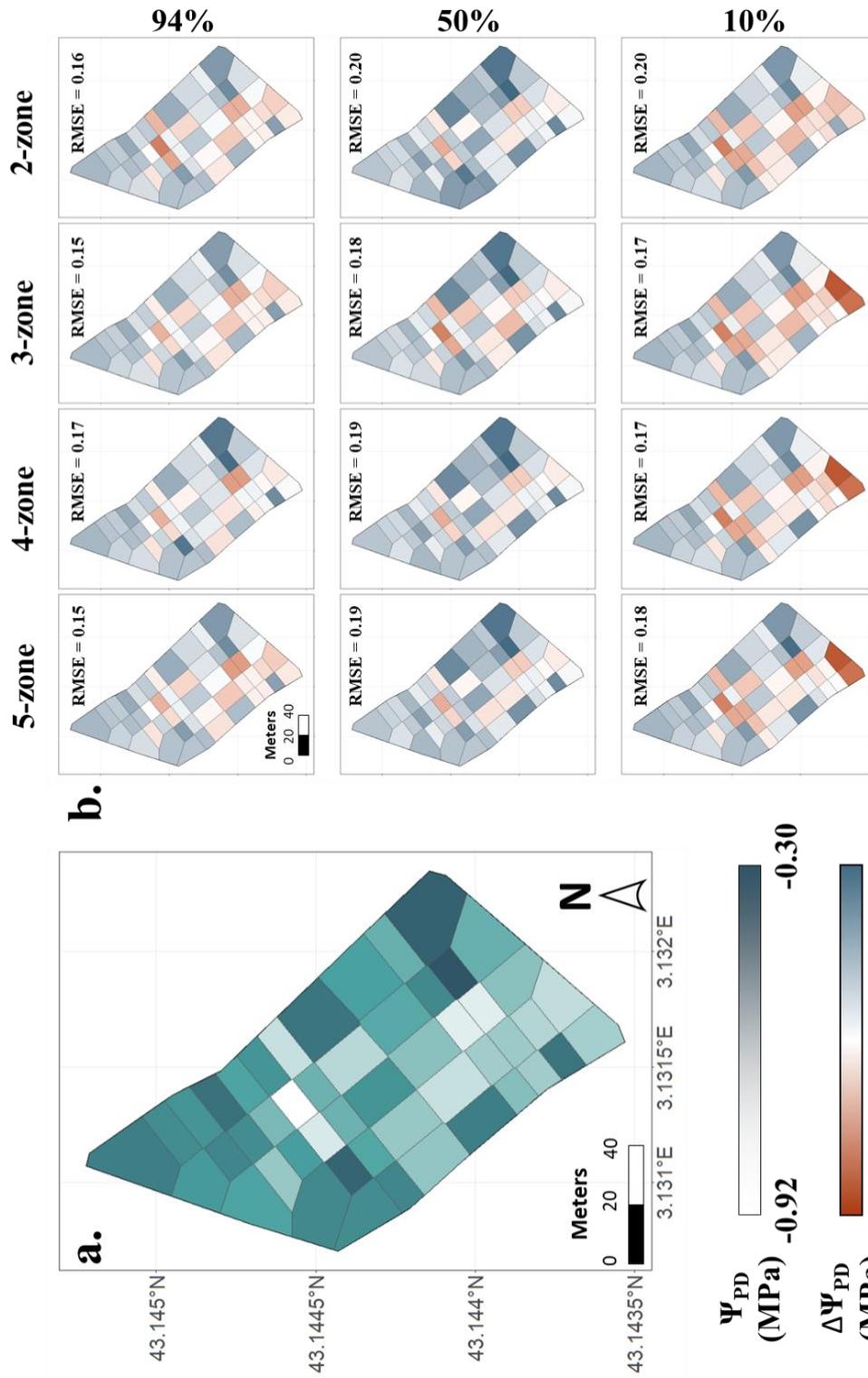


Figure S.A.2 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a moderate spatial structure of Ψ_{PD} between the (a.) observed (synthetic) Ψ_{PD} and (b.) modeling at different within-field spatial scales from 2-zone to 5-zone scales (columns) for each different considered level of correlation between Ψ_{PD} and the synthetic ancillary data (rows; 10 - 94%). The presented modeling date was based on observed data from August 18th with a high Ψ_{PD} variability.

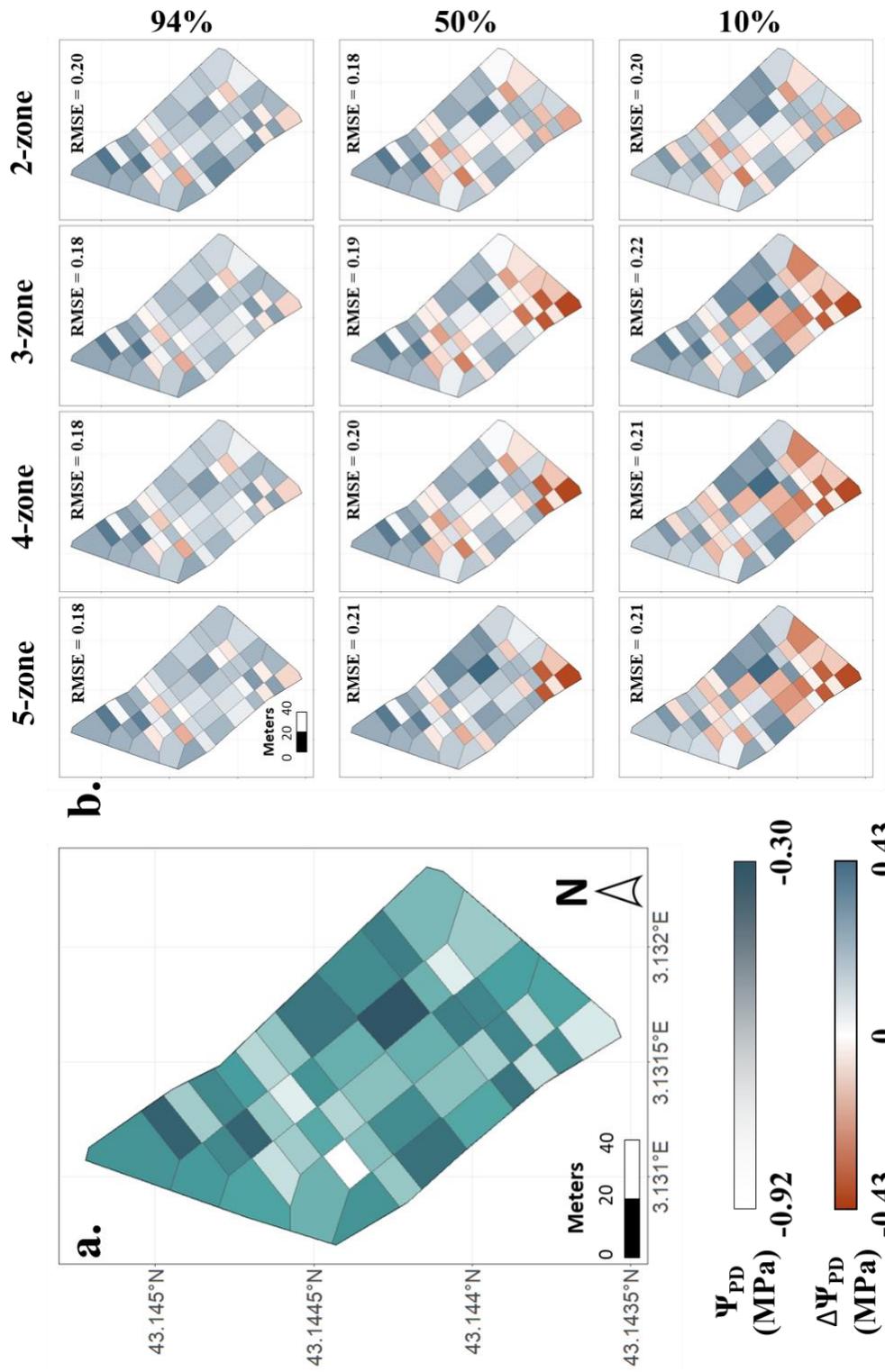


Figure S.A3 Performances of WaLIS in predicting predawn leaf water potential (Ψ_{PD}) for a synthetic field with a *weak* spatial structure of Ψ_{PD} between the (a.) observed (synthetic) Ψ_{PD} and (b.) modeling at different within-field spatial scales from 2-zone to 5-zone scales (columns) for each different considered level of correlation between Ψ_{PD} and the synthetic ancillary data (rows; 10 - 94%). The presented modeling date was based on observed data from August 18th with a high Ψ_{PD} variability.

Appendix B: Supplementary data of the Chapter 4

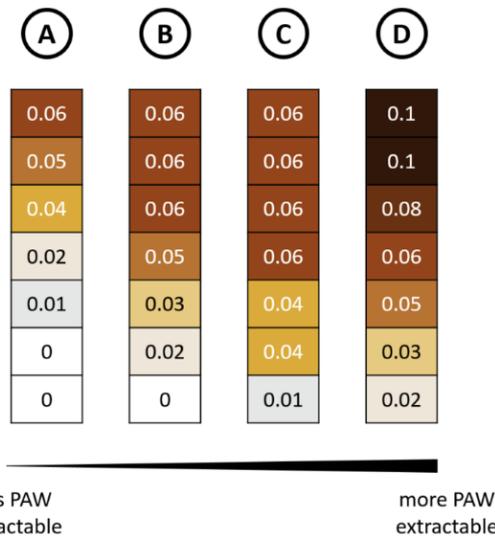


Figure S.B1 Different possibilities of fractions of plant available water able to be extracted (KL) from soil profiles for the durum wheat biomass calibration. KL is a soil characteristic related to extractable water. KL soil profiles are defined from profiles with little extractable water (A) to abundant extractable water (D). PAW = plant available water.

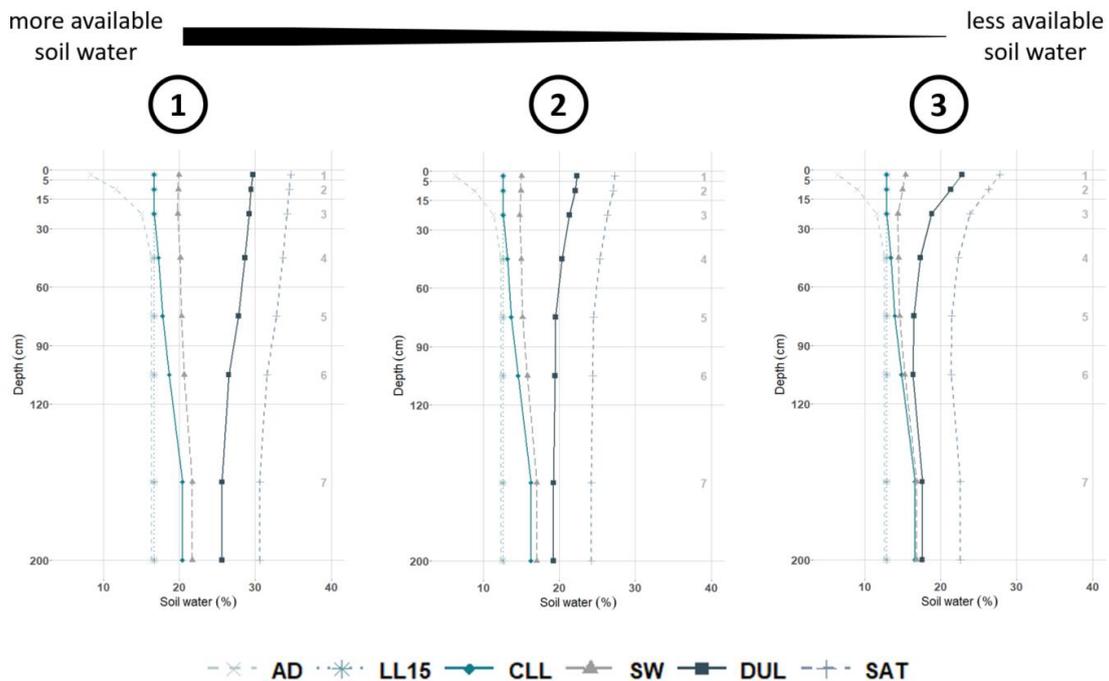


Figure S.B2 Different possibilities of soil hydraulic limits profiles for the durum wheat yield calibration. Difference between the crop lower limit (CLL) and drained upper limit (DUL) define the available soil water. Soil hydraulic limits profiles are defined from profiles with abundant available water (1) where the difference between CLL and DUL is large to profiles with little available water (3) where the difference between CLL and DUL is reduced. PAW = plant available water. Right numbers refer to soil layer number. Lines refer to soil hydraulic limits: AD = air dry moisture content, LL15 = lower limit soil moisture at -1.5 MPa, SW = soil water and SAT = soil water at saturation. Each point refers to the soil hydraulic limit value for the corresponding layer in volumetric water content.

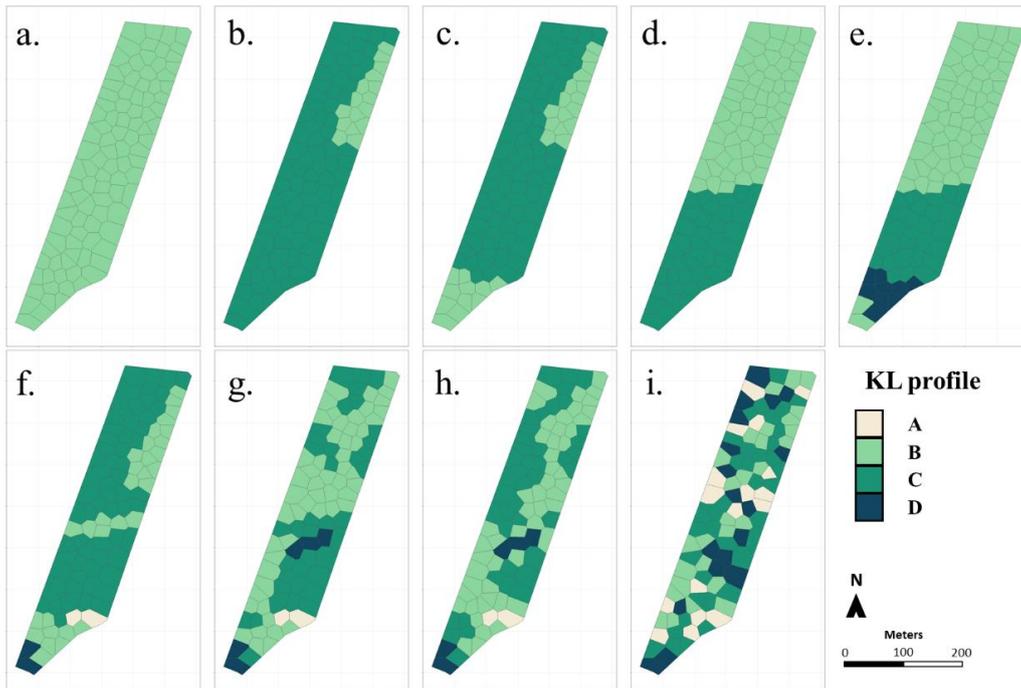


Figure S.B3 Maps of spatially calibrated fraction of extractable plant available water (KL) profiles for each simulation scale on the 2007-2008 data: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Note that the A and D KL profiles are respectively the profiles with the least and most extractable water (more details on KL profiles are shown in Figure S.1).

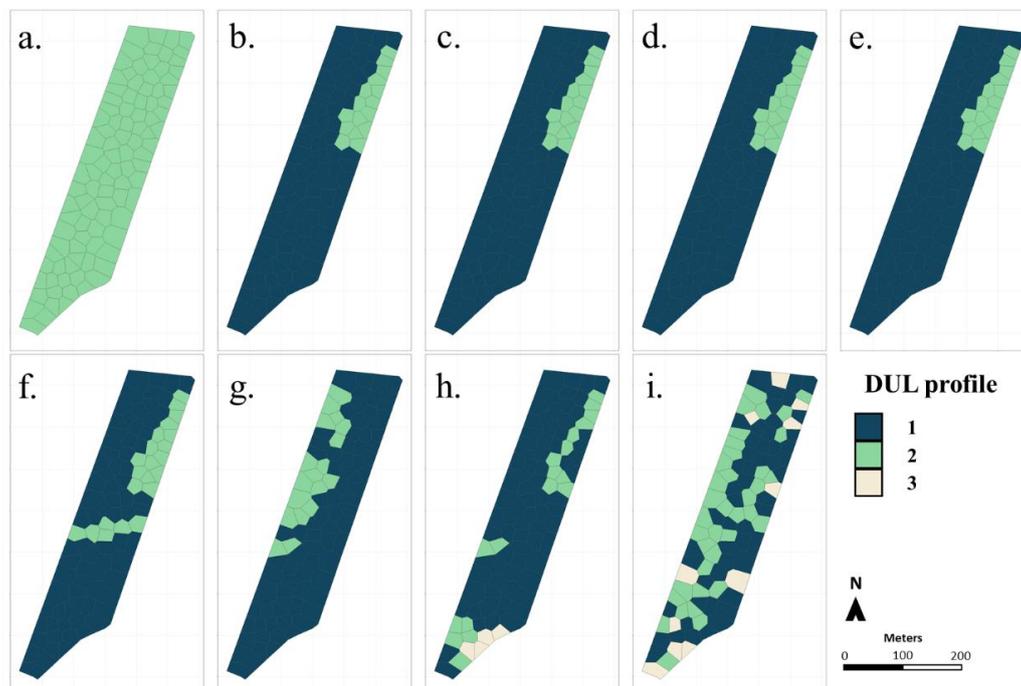


Figure S.B4 Maps of spatially calibrated drained upper limit (DUL) profiles for each simulation scale on the 2007-2008 data: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Note that the 1 and 3 DUL profiles are respectively the soil profiles with the most and least available soil water (more details on DUL profiles are shown in Figure S.2).

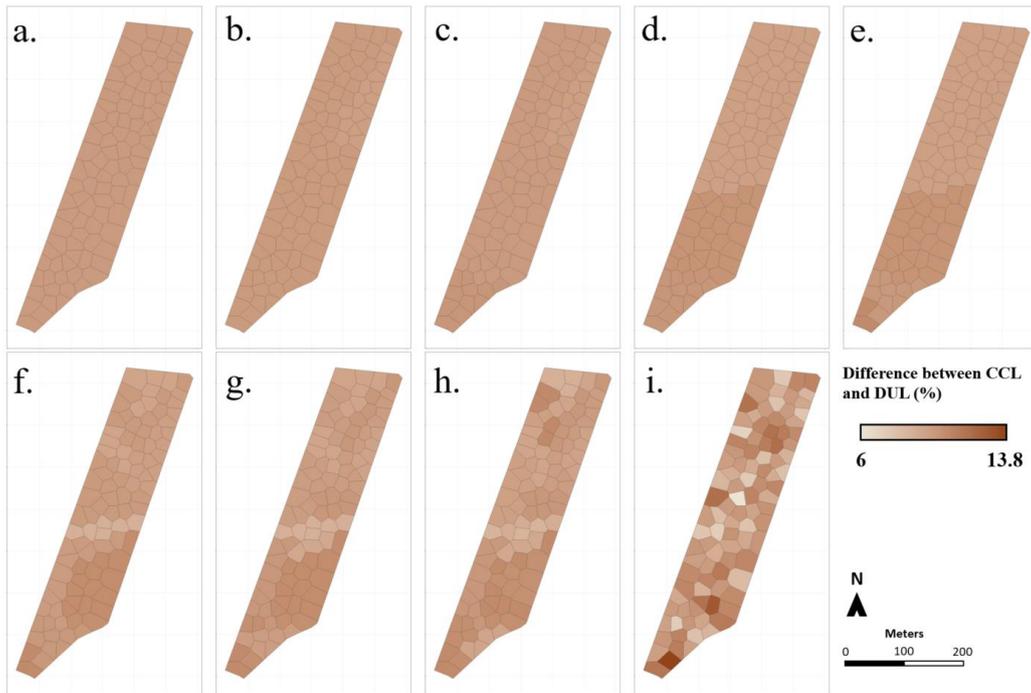


Figure S.B5 Maps of soil hydraulic limits (SHL) profiles represented by difference between crop lower limit (CCL) and drained upper limit (DUL) on the first layer for each simulation scale: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. The organic carbon in % (OC%) showed the same spatial patterning and utilized the same transfer function shape. The OC% maps were identical to the CCL-DUL maps with the legend ranging from 1.05% (minimum) to 1.83% (maximum) OC in the profile. They are not reproduced.

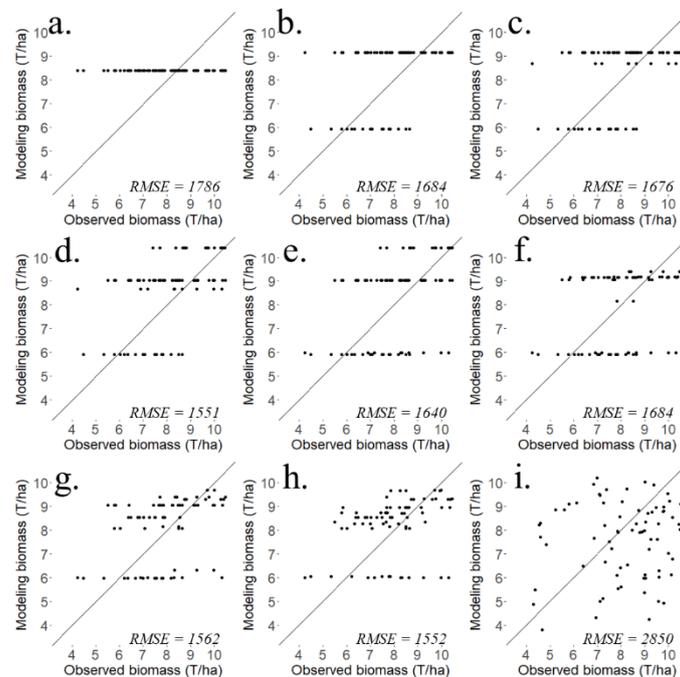


Figure S.B6 Comparison between observed and calibrated durum wheat biomass from APSIM for prediction in 2007-2008 (calibrated on 2005-2006 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Root mean square errors (RMSE) are in $\text{kg}\cdot\text{ha}^{-1}$.

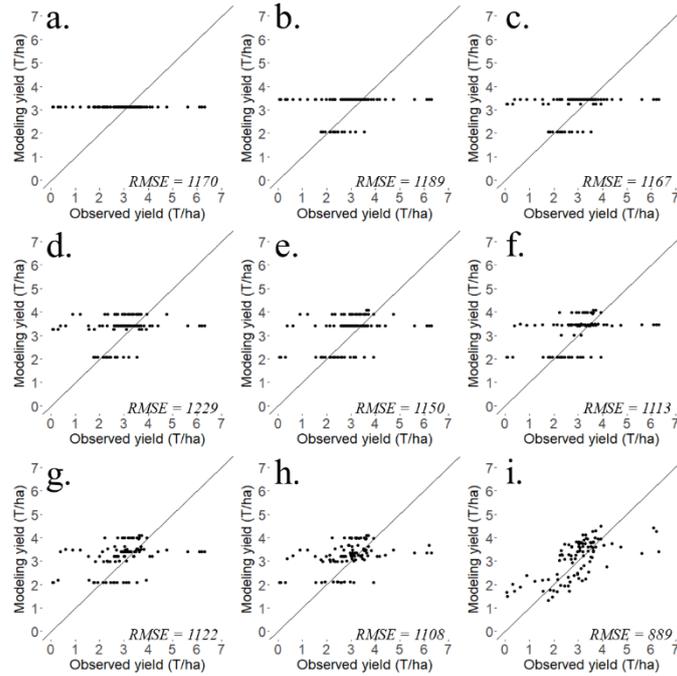


Figure S.B7 Comparison between observed and calibrated durum wheat yield from APSIM for prediction in 2007-2008 (calibrated on 2005-2006 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Root mean square errors (RMSE) are in $\text{kg}\cdot\text{ha}^{-1}$.

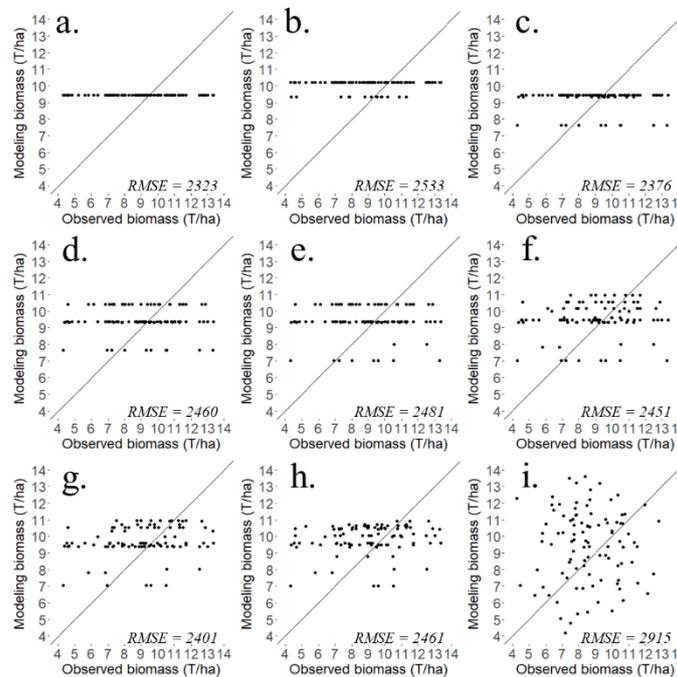


Figure S.B8 Comparison between observed and calibrated durum wheat biomass from APSIM prediction in 2005-2006 (calibrated on 2007-2008 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Root mean square errors (RMSE) are in $\text{kg}\cdot\text{ha}^{-1}$.

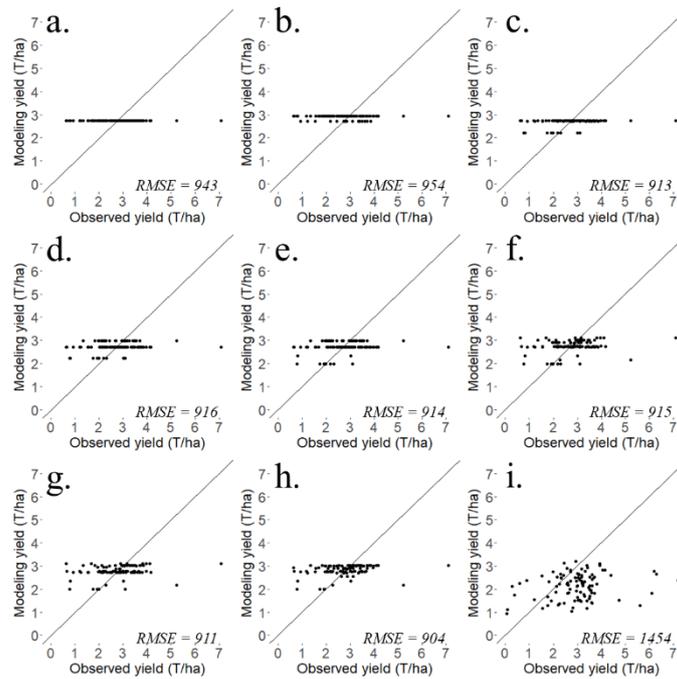


Figure S.B9 Comparison between observed and calibrated durum wheat yield from APSIM prediction in 2005-2006 (calibrated on 2007-2008 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale. Root mean square errors (RMSE) are in $\text{kg}\cdot\text{ha}^{-1}$.

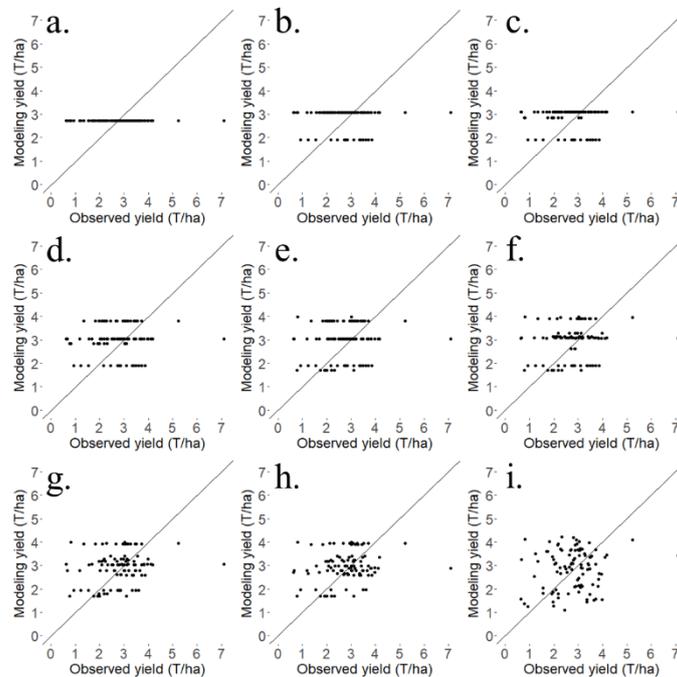


Figure S.B10 Comparison between observed and modeled durum wheat yield from APSIM prediction in 2005-2006 data (calibrated on 2007-2008 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale.

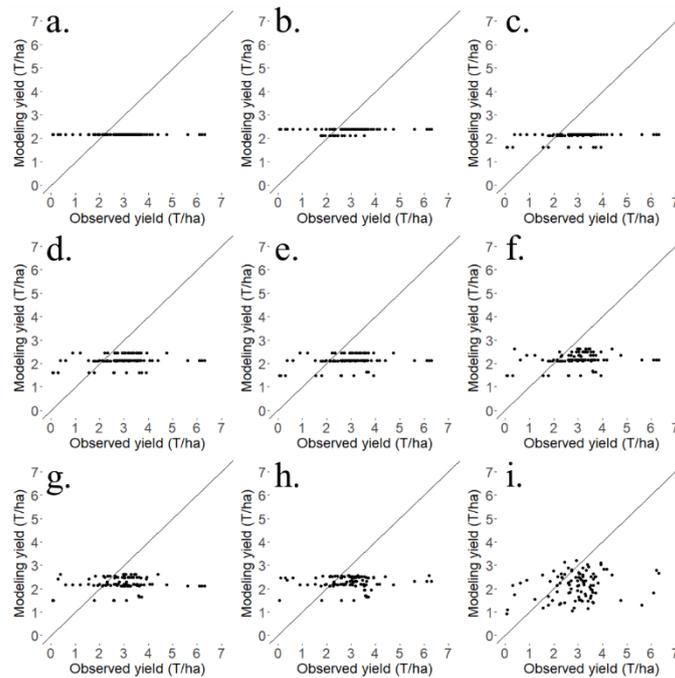


Figure S.B11 Comparison between observed and modeled durum wheat yield from APSIM predictions in 2007-2008 data (calibrated on 2005-2006 data) for different modeling scales: (a.) field scale, (b.) 2-zone scale, (c.) 3-zone scale, (d.) 4-zone scale, (e.) 5-zone scale, (f.) 10-zone scale, (g.) 15-zone scale, (h.) 20-zone scale and (i.) site-scale.

Table S.B1 APSIM cultivar parameters aspatially calibrated for the 2005-2006 and 2007-2008 production years.

Cultivar parameter		Calibration on 2005-2006	Calibration on 2007-2008
Target thermal time	Emergence	180 GDD	200 GDD
	End of juvenile	150 GDD	100 GDD
	Flowering	400 GDD	450 GDD
	Start grain filling	450 GDD	500 GDD
	Maturity	30 GDD	35 GDD
Leaf area growth	Maximum specific leaf area	16000 – 14000	22000 - 20000
	Minimum specific leaf area	13000	19000
Seed growth	Number of grains per stem	39	36
	Maximum grain size	0.035 g	0.02 g
	Potential grain growth rate	0.0005	0.0004

Table S.B2 APSIM soil parameters aspatially calibrated for the both production years 2005-2006 and 2007-2008.

Soil property	Layer	1	2	3	4	5	6	7
XF		1	1	1	1	1	1	1
FBiom		0.03	0.03	0.015	0.015	0.01	0.01	0.01
FInert		0.4	0.4	0.45	0.5	0.7	0.99	0.99

References

- Abrahamsen, P., Hansen, S., 2000. Daisy: an open soil-crop-atmosphere system model. *Environ. Model. Software* 15, 313–330. [https://doi.org/10.1016/S1364-8152\(00\)00003-7](https://doi.org/10.1016/S1364-8152(00)00003-7)
- Acevedo-Opazo, C., 2003. Extrapolation spatialisée d'une mesure locale de l'état hydrique de la vigne à partir de données auxiliaires (Spatialized extrapolation of a local measurement of the vine water status from ancillary data) (Thesis manuscript). Centre International d'Etudes Supérieures en Sciences Agronomiques, Montpellier, France.
- Acevedo-Opazo, C., Tisseyre, B., Ojeda, H., Ortega-Farías, S., Guillaume, S., 2008. Is it possible to assess the spatial variability of vine water status? *OENO One* 42, 203–219. <https://doi.org/10.20870/oeno-one.2008.42.4.811>
- Acevedo-Opazo, C., Tisseyre, B., Taylor, J.A., Ojeda, H., Guillaume, S., 2010. A model for the spatial prediction of water status in vines (*Vitis vinifera* L.) using high resolution ancillary information. *Precis. Agric.* 11, 358–378. <https://doi.org/10.1007/s11119-010-9164-7>
- Adam, M., Van Bussel, L.G.J., Leffelaar, P.A., Van Keulen, H., Ewert, F., 2011. Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions. *Ecol. Model.* 222, 131–143. <https://doi.org/10.1016/j.ecolmodel.2010.09.001>
- Adeyemi, O., Grove, I., Peets, S., Norton, T., 2017. Advanced Monitoring and Management Systems for Improving Sustainability in Precision Irrigation. *Sustainability* 9, 353. <https://doi.org/10.3390/su9030353>
- Allain, S., Plumecocq, G., Leenhardt, D., 2018. Spatial aggregation of indicators in sustainability assessments: Descriptive and normative claims. *Land Use Policy* 76, 577–588. <https://doi.org/10.1016/j.landusepol.2018.02.038>
- Allen, R.G., Jensen, M.E., Wright, J.L., Burman, R.D., 1989. Operational Estimates of Reference Evapotranspiration. *Agron. J.* 81, 650–662. <https://doi.org/10.2134/agronj1989.00021962008100040019x>
- Al-Shammari, D., Whelan, B.M., Wang, C., Bramley, R.G.V., Fajardo, M., Bishop, T.F.A., 2021. Impact of spatial resolution on the quality of crop yield predictions for site-specific crop management. *Agric. For. Meteorol.* 310, 108622. <https://doi.org/10.1016/j.agrformet.2021.108622>
- Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S., Ewert, F., 2013a. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. *Agric. For. Meteorol.* 170, 32–46. <https://doi.org/10.1016/j.agrformet.2012.11.017>
- Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., Ewert, F., 2013b. Characteristic ‘fingerprints’ of crop model responses to weather input data at different spatial resolutions. *Eur. J. Agron.* 49, 104–114. <https://doi.org/10.1016/j.eja.2013.04.003>
- Archontoulis, S.V., Miguez, F.E., Moore, K.J., 2014. A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: Application to soybean. *Environ. Model. Software* 62, 465–477. <https://doi.org/10.1016/j.envsoft.2014.04.009>
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O’Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. *Nat. Clim. Chang.* 3, 827–832. <https://doi.org/10.1038/nclimate1916>
- Asseng, S., Martre, P., Maiorano, A., Rötter, R.P., O’Leary, G.J., Fitzgerald, G.J., Girusse, C., Motzo, R., Giunta, F., Babar, M.A., Reynolds, M.P., Kheir, A.M.S., Thorburn, P.J., Waha, K., Ruane, A.C., Aggarwal, P.K., Ahmed, M., Balković, J., Basso, B., Biernath, C., Bindi, M., Cammarano, D., Challinor, A.J., De Sanctis, G., Dumont, B., Eyshi Rezaei, E., Fereres, E., Ferrise, R.,

- Garcia-Vila, M., Gayler, S., Gao, Y., Horan, H., Hoogenboom, G., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kassie, B.T., Kersebaum, K.-C., Klein, C., Koehler, A., Liu, B., Minoli, S., Montesino San Martin, M., Müller, C., Naresh Kumar, S., Nendel, C., Olesen, J.E., Palosuo, T., Porter, J.R., Priesack, E., Ripoche, D., Semenov, M.A., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Van der Velde, M., Wallach, D., Wang, E., Webber, H., Wolf, J., Xiao, L., Zhang, Z., Zhao, Z., Zhu, Y., Ewert, F., 2018. Climate change impact and adaptation for wheat protein. *Glob. Chang. Biol.* 25, 155–173. <https://doi.org/10.1111/gcb.14481>
- Asseng, S., Zhu, Y., Basso, B., Wilson, T., Cammarano, D., 2014. Simulation Modeling: Applications in Cropping Systems, in: *Encyclopedia of Agriculture and Food Systems*. Elsevier, pp. 102–112. <https://doi.org/10.1016/B978-0-444-52512-3.00233-3>
- Balkovič, J., van der Velde, M., Schmid, E., Skalský, R., Khabarov, N., Obersteiner, M., Stürmer, B., Xiong, W., 2013. Pan-European crop modelling with EPIC: Implementation, up-scaling and regional crop yield validation. *Agr. Syst.* 120, 61–75. <https://doi.org/10.1016/j.agsy.2013.05.008>
- Baralon, K., Payan, J.-C., Salançon, E., Tisseyre, B., 2012. Spider: spatial extrapolation of the vine water status at the whole denomination scale from a reference site. *OENO One* 46, 167–175. <https://doi.org/10.20870/oeno-one.2012.46.3.1517>
- Basso, B., Cammarano, D., Chen, D., Cafiero, G., Amato, M., Bitella, G., Rossi, R., Basso, F., 2009. Landscape Position and Precipitation Effects on Spatial Variability of Wheat Yield and Grain Protein in Southern Italy. *J. Agron. Crop Sci.* 195, 301–312. <https://doi.org/10.1111/j.1439-037X.2008.00351.x>
- Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., Dardanelli, J., 2012. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. *Eur. J. Agron.* 41, 52–65. <https://doi.org/10.1016/j.eja.2012.03.007>
- Basso, B., Ritchie, J.T., Cammarano, D., Sartori, L., 2011. A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. *Eur. J. Agron.* 35, 215–222. <https://doi.org/10.1016/j.eja.2011.06.004>
- Basso, B., Ritchie, J.T., Pierce, F.J., Braga, R.P., Jones, J.W., 2001. Spatial validation of crop models for precision agriculture. *Agric. Syst.* 68, 97–112. [https://doi.org/10.1016/S0308-521X\(00\)00063-9](https://doi.org/10.1016/S0308-521X(00)00063-9)
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., Le Dantec, V., Demarez, V., 2016. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* 184, 668–681. <https://doi.org/10.1016/j.rse.2016.07.030>
- Beaudoin, N., Gallois, N., Viennot, P., Le Bas, C., Puech, T., Schott, C., Buis, S., Mary, B., 2018. Evaluation of a spatialized agronomic model in predicting yield and N leaching at the scale of the Seine-Normandie Basin. *Environ. Sci. Pollut. Res.* 25, 23529–23558. <https://doi.org/10.1007/s11356-016-7478-3>
- Bell, J.C., Butler, C.A., Thompson, J.A., 1995. Soil-terrain modeling for site-specific agricultural management, in: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), *Proceedings of the Second International Conference on Site-Specific Management for Agricultural Systems*. ASA-CSSA-SSSA, Madison, Wisconsin, USA, pp. 209–227. <https://doi.org/10.2134/1995.site-specificmanagement.c16>
- Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. *Environ. Model. Software* 40, 1–20. <https://doi.org/10.1016/j.envsoft.2012.09.011>
- Blanchoud, H., Schott, C., Tallec, G., Queyrel, W., Gallois, N., Habets, F., Viennot, P., Ansart, P., Desportes, A., Tournebize, J., Puech, T., 2020. How Should Agricultural Practices Be Integrated to Understand and Simulate Long-Term Pesticide Contamination in the Seine River Basin?, in: Flipo, N., Labadie, P., Lestel, L. (Eds.), *The Seine River Basin, The Handbook of Environmental Chemistry*. Springer International Publishing, Cham, pp. 141–162. https://doi.org/10.1007/698_2019_385
- Bleiholder, J., Naumann, F., 2009. Data fusion. *ACM Computing Surveys* 41, 1–41. <https://doi.org/10.1145/1456650.1456651>

- Blöschl, G., 2005. Statistical Upscaling and Downscaling in Hydrology, in: Anderson, M.G., McDonnell, J.J. (Eds.), *Encyclopedia of Hydrological Sciences*. John Wiley & Sons, Ltd, Chichester, UK, pp. 135–154. <https://doi.org/10.1002/0470848944.hsa008>
- Bouman, B.A.M., van Keulen, H., van Laar, H.H., Rabbinge, R., 1996. The ‘School of de Wit’ crop growth simulation models: A pedigree and historical overview. *Agric. Syst.* 52, 171–198. [https://doi.org/10.1016/0308-521X\(96\)00011-X](https://doi.org/10.1016/0308-521X(96)00011-X)
- Bramley, R.G.V., Uzman, J., Trought, M.C.T., Neal, S.M., Bennett, J.S., 2019. Spatio-temporal variability in vine vigour and yield in a Marlborough Sauvignon Blanc vineyard. *Aust. J. Grape Wine Res.* 25, 430–438. <https://doi.org/10.1111/ajgw.12408>
- Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussi ere, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudill ere, J.P., H enault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model stics. *Eur. J. Agron.* 18, 309–332. [https://doi.org/10.1016/S1161-0301\(02\)00110-7](https://doi.org/10.1016/S1161-0301(02)00110-7)
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.-H., Ruget, F., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Nicoullaud, B., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.M., Meynard, J.M., Delecolle, R., 1998. STICS: a generic model for simulating crops and their water and nitrogen balances. I. Theory and parametrization applied to wheat and corn. *Agronomie* 18, 311–346. <https://doi.org/10.1051/agro:19980501>
- Brisson, N., Ruget, F., Gate, P., Lorgeou, J., Nicoullaud, B., Tayot, X., Plenet, D., Jeuffroy, M.-H., Bouthier, A., Ripoche, D., Mary, B., Justes, E., 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. *Agronomie* 22, 69–93. <https://doi.org/10.1051/agro:2001005>
- Buttafuoco, G., Castrignan o, A., Colecchia, A.S., Ricca, N., 2010. Delineation of Management Zones Using Soil Properties and a Multivariate Geostatistical Approach. *Ital. J. Agron.* 5, 323. <https://doi.org/10.4081/ija.2010.323>
- Buttafuoco, G., Castrignan o, A., Cucci, G., Lacolla, G., Luc a, F., 2017. Geostatistical modelling of within-field soil and yield variability for management zones delineation: a case study in a durum wheat field. *Precis. Agric.* 18, 37–58. <https://doi.org/10.1007/s11119-016-9462-9>
- Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F., Konopka, A.E., 1994. Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.* 58, 1501–1511. <https://doi.org/10.2136/sssaj1994.03615995005800050033x>
- Cammarano, D., Basso, B., Holland, J., Gianinetti, A., Baronchelli, M., Ronga, D., 2021. Modeling spatial and temporal optimal N fertilizer rates to reduce nitrate leaching while improving grain yield and quality in malting barley. *Comput. Electron. Agric.* 182, 105997. <https://doi.org/10.1016/j.compag.2021.105997>
- Cammarano, D., Holland, J., Basso, B., Fontana, F., Murgia, T., Lange, C., Taylor, J.A., Ronga, D., 2019. Integrating geospatial tools and a crop simulation model to understand spatial and temporal variability of cereals in Scotland, in: *Precision Agriculture ’19*. Presented at the 12th European Conference on Precision Agriculture, Wageningen Academic Publishers, Montpellier, France, pp. 29–35. https://doi.org/10.3920/978-90-8686-888-9_2
- Cammarano, D., Rivington, M., Matthews, K.B., Miller, D.G., Bellocchi, G., 2017. Implications of climate model biases and downscaling on crop model simulated climate change impacts. *Eur. J. Agron.* 88, 63–75. <https://doi.org/10.1016/j.eja.2016.05.012>
- Cammarano, D., Valdivia, R.O., Beletse, Y.G., Durand, W., Crespo, O., Tesfahuney, W.A., Jones, M.R., Walker, S., Mpuisang, T.N., Nhemachena, C., Ruane, A.C., Mutter, C., Rosenzweig, C., Antle, J., 2020. Integrated assessment of climate change impacts on crop productivity and income of commercial maize farms in northeast South Africa. *Food Secur.* 12, 659–678. <https://doi.org/10.1007/s12571-020-01023-0>
- Cao, Q., Miao, Y., Feng, G., Gao, X., Liu, B., Liu, Y., Li, F., Khosla, R., Mulla, D.J., Zhang, F., 2017. Improving nitrogen use efficiency with minimal environmental risks using an active canopy sensor in a wheat-maize cropping system. *Field Crop. Res.* 214, 365–372. <https://doi.org/10.1016/j.fcr.2017.09.033>
- Castrignan o, A., Quarto, R., Venezia, A., Buttafuoco, G., 2019. A comparison between mixed support kriging and block cokriging for modelling and combining spatial data with different support. *Precis. Agric.* 20, 193–213. <https://doi.org/10.1007/s11119-018-09630-w>

- Celette, F., Ripoché, A., Gary, C., 2010. WaLIS—A simple model to simulate water partitioning in a crop association: The example of an intercropped vineyard. *Agric. Water Manag.* 97, 1749–1759. <https://doi.org/10.1016/j.agwat.2010.06.008>
- Challinor, A.J., Ewert, F., Arnold, S., Simelton, E., Fraser, E., 2009. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. *J. Exp. Bot.* 60, 2775–2789. <https://doi.org/10.1093/jxb/erp062>
- Challinor, A.J., Müller, C., Asseng, S., Deva, C., Nicklin, K., Wallach, D., Vanuytrecht, E., Whitfield, S., Ramirez-Villegas, J., Koehler, A.-K., 2017. Improving the use of crop models for risk assessment and climate change adaptation. *Agric. Syst.* 159, 296–306. <https://doi.org/10.1016/j.agsy.2017.07.010>
- Chapagain, R., Remenyi, T.A., Harris, R.M.B., Mohammed, C.L., Huth, N., Wallach, D., Rezaei, E.E., Ojeda, J.J., 2022. Decomposing crop model uncertainty: A systematic review. *Field Crop. Res.* 279, 108448. <https://doi.org/10.1016/j.fcr.2022.108448>
- Chen, H., Leinonen, I., Marshall, B., Taylor, J.A., 2017. Conceptual Spatial Crop Models for Potato Production. *Advances in Animal Biosciences* 8, 678–683. <https://doi.org/10.1017/S2040470017000851>
- Cheng, D., Yao, Y., Liu, R., Li, X., Guan, B., Yu, F., 2023. Precision agriculture management based on a surrogate model assisted multiobjective algorithmic framework. *Sci. Rep.* 13, 1142. <https://doi.org/10.1038/s41598-023-27990-w>
- Choukri, F., Raclot, D., Naimi, M., Chikhaoui, M., Nunes, J.P., Huard, F., Hérivaux, C., Sabir, M., Pépin, Y., 2020. Distinct and combined impacts of climate and land use scenarios on water availability and sediment loads for a water supply reservoir in northern Morocco. *Int. Soil Water Conserv. Res.* 8, 141–153. <https://doi.org/10.1016/j.iswcr.2020.03.003>
- Claverie, M., Demarez, V., Duchemin, B., Hagolle, O., Ducrot, D., Marais-Sicre, C., Dejoux, J.-F., Huc, M., Keravec, P., Béziat, P., Fieuzal, R., Ceschia, E., Dedieu, G., 2012. Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. *Remote Sens. Environ.* 124, 844–857. <https://doi.org/10.1016/j.rse.2012.04.005>
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* 20, 37–46. <https://doi.org/10.1177/001316446002000104>
- Colaço, A.F., Bramley, R.G.V., 2018. Do crop sensors promote improved nitrogen management in grain crops? *Field Crop. Res.* 218, 126–140. <https://doi.org/10.1016/j.fcr.2018.01.007>
- Confalonieri, R., Orlando, F., Palerai, L., Stella, T., Gilardelli, C., Movedi, E., Pagani, V., Cappelli, G., Vertemara, A., Alberti, L., Alberti, P., Atanassiu, S., Bonaiti, M., Cappelletti, G., Ceruti, M., Confalonieri, A., Corgatelli, G., Corti, P., Dell’Oro, M., Ghidoni, A., Lamarta, A., Maghini, A., Mambretti, M., Manchia, A., Massoni, G., Mutti, P., Pariani, S., Pasini, D., Pesenti, A., Pizzamiglio, G., Ravasio, A., Rea, A., Santorsola, D., Serafini, G., Slavazza, M., Acutis, M., 2016. Uncertainty in crop model predictions: What is the role of users? *Environ. Model. Software* 81, 165–173. <https://doi.org/10.1016/j.envsoft.2016.04.009>
- Constantin, J., Raynal, H., Casellas, E., Hoffmann, H., Bindi, M., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.-C., Klatt, S., Kuhnert, M., Lewan, E., Maharjan, G.R., Moriondo, M., Nendel, C., Roggero, P.P., Specka, X., Trombi, G., Villa, A., Wang, E., Weihermüller, L., Yeluripati, J., Zhao, Z., Ewert, F., Bergez, J.-E., 2019. Management and spatial resolution effects on yield and water balance at regional scale in crop models. *Agric. For. Meteorol.* 275, 184–195. <https://doi.org/10.1016/j.agrformet.2019.05.013>
- Cornet, D., Marcos, J., Tournebize, R., Sierra, J., 2022. Observed and modeled response of water yam (*Dioscorea alata* L.) to nitrogen supply: Consequences for nitrogen fertilizer management in the humid tropics. *Eur. J. Agron.* 138, 126536. <https://doi.org/10.1016/j.eja.2022.126536>
- Corwin, D.L., Lesch, S.M., 2005. Apparent soil electrical conductivity measurements in agriculture. *Comput. Electron. Agric.* 46, 11–43. <https://doi.org/10.1016/j.compag.2004.10.005>
- Coucheney, E., Buis, S., Launay, M., Constantin, J., Mary, B., García De Cortázar-Atauri, I., Ripoché, D., Beaudoin, N., Ruget, F., Andrianarisoa, K.S., Le Bas, C., Justes, E., Léonard, J., 2015. Accuracy, robustness and behavior of the STICS soil–crop model for plant, water and nitrogen outputs: Evaluation over a wide range of agro-environmental conditions in France. *Environ. Model. Software* 64, 177–190. <https://doi.org/10.1016/j.envsoft.2014.11.024>

- Crosetto, M., Tarantola, S., Saltelli, A., 2000. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agric. Ecosyst. Environ.* 81, 71–79. [https://doi.org/10.1016/S0167-8809\(00\)00169-9](https://doi.org/10.1016/S0167-8809(00)00169-9)
- Dalglish, N., Hochman, Z., Huth, N., Holzworth, D., 2016. Field Protocol to APSOIL characterisations - Version 4. CSIRO, Australia.
- de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., Supit, I., van der Wijngaart, R., van Diepen, K., 2019. 25 years of the WOFOST cropping systems model. *Agric. Syst.* 168, 154–167. <https://doi.org/10.1016/j.agry.2018.06.018>
- Derby, N.E., Casey, F.X.M., Franzen, D.W., 2007. Comparison of Nitrogen Management Zone Delineation Methods for Corn Grain Yield. *Agron. J.* 99, 405–414. <https://doi.org/10.2134/agronj2006.0027>
- Diacono, M., Castrignanò, A., Troccoli, A., De Benedetto, D., Basso, B., Rubino, P., 2012. Spatial and temporal variability of wheat grain yield and quality in a Mediterranean environment: A multivariate geostatistical approach. *Field Crop. Res.* 131, 49–62. <https://doi.org/10.1016/j.fcr.2012.03.004>
- Domínguez-Álvarez, A., De-Tena-Rey, M.-T., García-Moruno, L., 2021. Modelling global solar radiation to optimise agricultural production. *Span. J. Agric. Res.* 19, e0201. <https://doi.org/10.5424/sjar/2021191-16813>
- Duchemin, B., Maisongrande, P., Boulet, G., Benhadj, I., 2008. A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index. *Environ. Model. Software* 23, 876–892. <https://doi.org/10.1016/j.envsoft.2007.10.003>
- Ehrhardt, F., Soussana, J., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., Sándor, R., Smith, P., Snow, V., de Antoni Migliorati, M., Basso, B., Bhatia, A., Brill, L., Doltra, J., Dorich, C.D., Doro, L., Fitton, N., Giacomini, S.J., Grant, B., Harrison, M.T., Jones, S.K., Kirschbaum, M.U.F., Klumpp, K., Laville, P., Léonard, J., Liebig, M., Lieffering, M., Martin, R., Massad, R.S., Meier, E., Merbold, L., Moore, A.D., Myrghiotis, V., Newton, P., Pattey, E., Rolinski, S., Sharp, J., Smith, W.N., Wu, L., Zhang, Q., 2018. Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N₂O emissions. *Glob. Chang. Biol.* 24, e603–e616. <https://doi.org/10.1111/gcb.13965>
- Ewert, F., Rötter, R.P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K.C., Olesen, J.E., van Ittersum, M.K., Janssen, S., Rivington, M., Semenov, M.A., Wallach, D., Porter, J.R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P.P., Bartošová, L., Asseng, S., 2015. Crop modelling for integrated assessment of risk to food production from climate change. *Environ. Model. Software* 72, 287–303. <https://doi.org/10.1016/j.envsoft.2014.12.003>
- Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepikina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agric. Ecosyst. Environ.* 142, 6–17. <https://doi.org/10.1016/j.agee.2011.05.016>
- Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G., Wallach, D., 2004. Spatialising crop models. *Agronomie* 24, 205–217. <https://doi.org/10.1051/agro:2004016>
- Fontanet, M., Scudiero, E., Skaggs, T.H., Fernández-García, D., Ferrer, F., Rodrigo, G., Bellvert, J., 2020. Dynamic Management Zones for Irrigation Scheduling. *Agric. Water Manag.* 238, 106207. <https://doi.org/10.1016/j.agwat.2020.106207>
- Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K.J., Singh, U., Asseng, S., Kahveci, T., He, J., Zhang, R., Confalonieri, R., Hoogenboom, G., 2020. Comparison of three calibration methods for modeling rice phenology. *Agric. For. Meteorol.* 280, 107785. <https://doi.org/10.1016/j.agrformet.2019.107785>
- García-Vila, M., Fereres, E., 2012. Combining the simulation crop model AquaCrop with an economic model for the optimization of irrigation management at farm level. *Eur. J. Agron.* 36, 21–31. <https://doi.org/10.1016/j.eja.2011.08.003>
- Ginaldi, F., Bajocco, S., Bregaglio, S., Cappelli, G., 2019. Spatializing crop models for sustainable agriculture, in: *Innovations in Sustainable Agriculture*. Springer, pp. 599–619.
- Goffart, J.P., Olivier, M., Frankinet, M., 2008. Potato Crop Nitrogen Status Assessment to Improve N Fertilization Management and Efficiency: Past–Present–Future. *Potato Res.* 51, 355–383. <https://doi.org/10.1007/s11540-008-9118-x>

- Gräler, B., Pebesma, E.J., Heuvelink, H., 2016. Spatio-Temporal Interpolation using gstat. *The R Journal* 8, 204–218.
- Grosz, B., Dechow, R., Gebbert, S., Hoffmann, H., Zhao, G., Constantin, J., Raynal, H., Wallach, D., Coucheney, E., Lewan, E., Eckersten, H., Specka, X., Kersebaum, K.-C., Nendel, C., Kuhnert, M., Yeluripati, J., Haas, E., Teixeira, E., Bindi, M., Trombi, G., Moriondo, M., Doro, L., Roggero, P.P., Zhao, Z., Wang, E., Tao, F., Rötter, R., Kassie, B., Cammarano, D., Asseng, S., Weihermüller, L., Siebert, S., Gaiser, T., Ewert, F., 2017. The implication of input data aggregation on up-scaling soil organic carbon changes. *Environ. Model. Software* 96, 361–377. <https://doi.org/10.1016/j.envsoft.2017.06.046>
- Guastaferro, F., Castrignanò, A., De Benedetto, D., Sollitto, D., Troccoli, A., Cafarelli, B., 2010. A comparison of different algorithms for the delineation of management zones. *Precis. Agric.* 11, 600–620. <https://doi.org/10.1007/s11119-010-9183-4>
- Guillaume, S., Lablée, J.L., 2022. GeoFIS: Spatial Data Processing for Decision Making R package version 1.0.3.
- Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., Cheng, T., Tian, Y., 2018. Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precis. Agric.* 19, 55–78. <https://doi.org/10.1007/s11119-017-9498-5>
- Hall, A., Lamb, D.W., Holzapfel, B.P., Louis, J.P., 2011. Within-season temporal variation in correlations between vineyard canopy and winegrape composition and yield. *Precis. Agric.* 12, 103–117. <https://doi.org/10.1007/s11119-010-9159-4>
- Han, E., Ines, A.V.M., Koo, J., 2019. Development of a 10-km resolution global soil profile dataset for crop modeling applications. *Environ. Model. Software* 119, 70–83. <https://doi.org/10.1016/j.envsoft.2019.05.012>
- Hansen, S., Abrahamsen, P., Petersen, C.T., Styczen, M., 2012. Daisy: Model Use, Calibration, and Validation. *Trans. ASABE* 55, 1315–1333. <https://doi.org/10.13031/2013.42244>
- Hansen, S., Jensen, H.E., Nielsen, N.E., 1991. Simulation of nitrogen dynamics and biomass production in winter wheat using the Danish simulation model Daisy. *Fertil. Res.* 27, 245–259.
- Hao, S., Ryu, D., Western, A., Perry, E., Bogen, H., Franssen, H.J.H., 2021. Performance of a wheat yield prediction model and factors influencing the performance: A review and meta-analysis. *Agric. Syst.* 194, 103278. <https://doi.org/10.1016/j.agsy.2021.103278>
- Heuvelink, G.B.M., 2002. Analysing Uncertainty Propagation in GIS: Why is it not that Simple?, in: *Uncertainty in Remote Sensing and GIS*. John Wiley & Sons, Ltd, Chichester, UK, pp. 155–165. <https://doi.org/10.1002/0470035269.ch10>
- Heuvelink, G.B.M., Brus, D.J., Reinds, G., 2010. Accounting for spatial sampling effects in regional uncertainty propagation analysis. Presented at the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, pp. 85–88.
- Hochman, Z., Horan, H., 2018. Causes of wheat yield gaps and opportunities to advance the water-limited yield frontier in Australia. *Field Crop. Res.* 228, 20–30. <https://doi.org/10.1016/j.fcr.2018.08.023>
- Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T., Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X., Tao, F., Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., Ewert, F., 2016. Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations. *PLoS ONE* 11, e0151782. <https://doi.org/10.1371/journal.pone.0151782>
- Hoffmann, H., Zhao, G.R., van Bussel, L.G.J., Enders, A., Specka, X., Sosa, C., Yeluripati, J., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Zhao, Z., Wang, E., Nendel, C., Kersebaum, K., Haas, E., Kiese, R., Klatt, S., Eckersten, H., Vanuytrecht, E., Kuhnert, M., Lewan, E., Rötter, R., Roggero, P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015. Variability of effects of spatial climate data aggregation on regional yield simulation by crop models. *Clim. Res.* 65, 53–69. <https://doi.org/10.3354/cr01326>

- Hollis, J.M., Hannam, J., Bellamy, P.H., 2012. Empirically-derived pedotransfer functions for predicting bulk density in European soils. *Eur. J. Soil Sci.* 63, 96–109. <https://doi.org/10.1111/j.1365-2389.2011.01412.x>
- Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalglish, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. *Environ. Model. Software* 62, 327–350. <https://doi.org/10.1016/j.envsoft.2014.07.009>
- Hoogenboom, G., 2000. Contribution of agrometeorology to the simulation of crop production and its applications. *Agric. For. Meteorol.* 103, 137–157. [https://doi.org/10.1016/S0168-1923\(00\)00108-8](https://doi.org/10.1016/S0168-1923(00)00108-8)
- Hoogenboom, G., Porter, C.H., Boote, K.J., Shelia, V., Wilkens, P.W., Singh, U., White, J.W., Asseng, S., Lizaso, J.I., Moreno, L.P., Pavan, W., Ogoshi, R., Hunt, L.A., Tsuji, G.Y., Jones, J.W., 2019. The DSSAT crop modeling ecosystem, in: *Advances in Crop Modeling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge, United Kingdom, pp. 173–216. <https://doi.org/10.19103/AS.2019.0061.10>
- Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., White, J.W., Pavan, W., Oliveira, F.A.A., Moreno-Cadena, L.P., Lizaso, J.I., Asseng, S., Pequeno, D.N.L., Kimball, B.A., Alderman, P.D., Thorp, K.R., Jones, M.R., Cuadra, S.V., Vianna, M.S., Villalobos, F.J., Ferreira, T.B., Batchelor, W.D., Koo, J., Hunt, L.A., Jones, J.W., 2021. Decision Support System for Agrotechnology Transfer (DSSAT).
- Hovadik, J.M., Larue, D.K., 2007. Static characterizations of reservoirs: refining the concepts of connectivity and continuity. *PG* 13, 195–211. <https://doi.org/10.1144/1354-079305-697>
- Hsiao, T.C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., Fereres, E., 2009. AquaCrop - The FAO Crop Model to Simulate Yield Response to Water: III. Parameterization and Testing for Maize. *Agron. J.* 101, 448–459. <https://doi.org/10.2134/agronj2008.0218s>
- Hu, S., Shi, L., Huang, K., Zha, Y., Hu, X., Ye, H., Yang, Q., 2019. Improvement of sugarcane crop simulation by SWAP-WOFOST model via data assimilation. *Field Crop. Res.* 232, 49–61. <https://doi.org/10.1016/j.fcr.2018.12.009>
- Huang, J., Gómez-Dans, J.L., Huang, H., Ma, H., Wu, Q., Lewis, P.E., Liang, S., Chen, Z., Xue, J.-H., Wu, Y., Zhao, F., Wang, J., Xie, X., 2019. Assimilation of remote sensing into crop growth models: Current status and perspectives. *Agric. For. Meteorol.* 276–277, 107609. <https://doi.org/10.1016/j.agrformet.2019.06.008>
- Huard, F., Choukri, F., Raclot, D., Pépin, Y., Chikhaoui, M., Naimi, M., Korkot, H., Yassin, M., 2019. Spatialisation fine des projections climatiques dans le bassin versant du Tleta au Nord Maroc (Fine spatialization of climate projections in the Tleta watershed in northern Morocco). *Rev. Maroc. Sci. Agron. Vet.* 7, 294–303.
- Huth, N.I., Bristow, K.L., Verburg, K., 2012. SWIM3: Model Use, Calibration, and Validation. *Trans. ASABE* 55, 1303–1313. <https://doi.org/10.13031/2013.42243>
- Jagtap, S.S., Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agric. Ecosyst. Environ.* 93, 73–85.
- Jankowski, P., Andrienko, N., Andrienko, G., 2001. Map-centred exploratory approach to multiple criteria spatial decision making. *International Journal of Geographical Information Science* 15, 101–127. <https://doi.org/10.1080/13658810010005525>
- Jha, P.K., Ines, A.V.M., Han, E., Cruz, R., Vara Prasad, P.V., 2022. A comparison of multiple calibration and ensembling methods for estimating genetic coefficients of CERES-Rice to simulate phenology and yields. *Field Crop. Res.* 284, 108560. <https://doi.org/10.1016/j.fcr.2022.108560>
- Ji, G., Tian, L., Zhao, J., Yue, Y., Wang, Z., Yang, X., 2018. Error analysis and correction of spatialization of crop yield in China – Different variables scales, partitioning schemes and error correction methods. *Comput. Electron. Agric.* 148, 272–279. <https://doi.org/10.1016/j.compag.2018.03.031>

- Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J., 2018. A review of data assimilation of remote sensing and crop models. *Eur. J. Agron.* 92, 141–152. <https://doi.org/10.1016/j.eja.2017.11.002>
- Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017a. Brief history of agricultural systems modeling. *Agric. Syst.* 155, 240–254. <https://doi.org/10.1016/j.agry.2016.05.014>
- Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017b. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agric. Syst.* 155, 269–288. <https://doi.org/10.1016/j.agry.2016.09.021>
- Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18, 235–265. [https://doi.org/10.1016/S1161-0301\(02\)00107-7](https://doi.org/10.1016/S1161-0301(02)00107-7)
- Kamali, B., Lorite, I.J., Webber, H.A., Rezaei, E.E., Gabaldon-Leal, C., Nendel, C., Siebert, S., Ramirez-Cuesta, J.M., Ewert, F., Ojeda, J.J., 2022. Uncertainty in climate change impact studies for irrigated maize cropping systems in southern Spain. *Sci Rep* 12, 4049. <https://doi.org/10.1038/s41598-022-08056-9>
- Kasampalis, D.A., Alexandridis, T.K., Deva, C., Challinor, A., Moshou, D., Zalidis, G., 2018. Contribution of Remote Sensing on Crop Models: A Review. *J. Imaging* 4, 52. <https://doi.org/10.3390/jimaging4040052>
- Kawakita, S., Takahashi, H., Moriya, K., 2020. Prediction and parameter uncertainty for winter wheat phenology models depend on model and parameterization method differences. *Agric. For. Meteorol.* 290, 107998. <https://doi.org/10.1016/j.agrformet.2020.107998>
- Kazmierski, M., Glémas, P., Rousseau, J., Tisseyre, B., 2011. Temporal stability of within-field patterns of NDVI in non irrigated Mediterranean vineyards. *OENO One* 45, 61. <https://doi.org/10.20870/oeno-one.2011.45.2.1488>
- Kerry, R., Oliver, M.A., 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma* 140, 383–396. <https://doi.org/10.1016/j.geoderma.2007.04.019>
- Khanal, S., Fulton, J., Shearer, S., 2017. An overview of current and potential applications of thermal remote sensing in precision agriculture. *Comput. Electron. Agric.* 139, 22–32. <https://doi.org/10.1016/j.compag.2017.05.001>
- Koch, J., Mendiguren, G., Mariethoz, G., Stisen, S., 2017. Spatial Sensitivity Analysis of Simulated Land Surface Patterns in a Catchment Model Using a Set of Innovative Spatial Performance Metrics. *J. Hydrometeorol.* 18, 1121–1142. <https://doi.org/10.1175/JHM-D-16-0148.1>
- Krajewski, W.F., Ciach, G.J., Habib, E., 2003. An analysis of small-scale rainfall variability in different climatic regimes. *Hydrol. Sci. J.* 48, 151–162. <https://doi.org/10.1623/hysj.48.2.151.44694>
- Lajili, A., Cambouris, A.N., Chokmani, K., Duchemin, M., Perron, I., Zebbarh, B.J., Biswas, A., Adamchuk, V.I., 2021. Analysis of Four Delineation Methods to Identify Potential Management Zones in a Commercial Potato Field in Eastern Canada. *Agron.* 11, 432. <https://doi.org/10.3390/agronomy11030432>
- Launay, M., Guerif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. *Agric. Ecosyst. Environ.* 111, 321–339. <https://doi.org/10.1016/j.agee.2005.06.005>
- Laux, P., Rötter, R.P., Webber, H., Dieng, D., Rahimi, J., Wei, J., Faye, B., Srivastava, A.K., Bliefernicht, J., Adeyeri, O., Arnault, J., Kunstmann, H., 2021. To bias correct or not to bias correct? An agricultural impact modelers' perspective on regional climate model data. *Agric. Water Manag.* 304–305, 108406. <https://doi.org/10.1016/j.agrformet.2021.108406>
- Lebon, E., Dumas, V., Pieri, P., Schultz, H.R., 2003. Modelling the seasonal dynamics of the soil water balance of vineyards. *Funct. Plant Biol.* 30, 699. <https://doi.org/10.1071/FP02222>
- Leo, S., De Antoni Migliorati, M., Nguyen, T.H., Grace, P.R., 2023. Combining remote sensing-derived management zones and an auto-calibrated crop simulation model to determine optimal nitrogen fertilizer rates. *Agr. Syst.* 205, 103559. <https://doi.org/10.1016/j.agry.2022.103559>

- Leroux, C., Jones, H., Pichon, L., Guillaume, S., Lamour, J., Taylor, J., Naud, O., Crestey, T., Lablee, J.-L., Tisseyre, B., 2018. GeoFIS: An Open Source, Decision-Support Tool for Precision Agriculture Data. *Agriculture* 8, 73. <https://doi.org/10.3390/agriculture8060073>
- Leroux, C., Tisseyre, B., 2018. How to measure and report within-field variability: a review of common indicators and their sensitivity. *Precis. Agric.* 20, 562–590. <https://doi.org/10.1007/s11119-018-9598-x>
- Li, J., Zhang, H., Xu, E., 2020. Spatialization of Actual Grain Crop Yield Coupled with Cultivation Systems and Multiple Factors: From Survey Data to Grid. *Agron.* 10, 675. <https://doi.org/10.3390/agronomy10050675>
- Li, Zhenhai, He, J., Xu, X., Jin, X., Huang, W., Clark, B., Yang, G., Li, Zhenhong, 2018. Estimating genetic parameters of DSSAT-CERES model with the GLUE method for winter wheat (*Triticum aestivum* L.) production. *Comput. Electron. Agric.* 154, 213–221. <https://doi.org/10.1016/j.compag.2018.09.009>
- Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P., Naylor, R.L., 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. *Sci.* 319, 607–610. <https://doi.org/10.1126/science.1152339>
- Luedeling, E., Smethurst, P.J., Baudron, F., Bayala, J., Huth, N.I., Van Noordwijk, M., Ong, C.K., Mulia, R., Lusiana, B., Muthuri, C., Sinclair, F.L., 2016. Field-scale modeling of tree–crop interactions: Challenges and development needs. *Agr. Syst.* 142, 51–69. <https://doi.org/10.1016/j.agry.2015.11.005>
- Luo, Q., Griffith, D.A., Wu, H., 2019. Spatial autocorrelation for massive spatial data: verification of efficiency and statistical power asymptotics. *J. Geogr. Syst.* 21, 237–269. <https://doi.org/10.1007/s10109-019-00293-3>
- MacKerron, D.K.L., Marshall, B., McNicol, J.W., 2004. MAPP and the underlying functions that it contains, in: *Decision Support Systems in Potato Production - Bringing Models to Practice*.
- Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., Fritschi, F.B., 2020. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* 237, 111599. <https://doi.org/10.1016/j.rse.2019.111599>
- Maiorano, A., Martre, P., Asseng, S., Ewert, F., Müller, C., Rötter, R.P., Ruane, A.C., Semenov, M.A., Wallach, D., Wang, E., Alderman, P.D., Kassie, B.T., Biernath, C., Basso, B., Cammarano, D., Challinor, A.J., Doltra, J., Dumont, B., Rezaei, E.E., Gayler, S., Kersebaum, K.C., Kimball, B.A., Koehler, A.-K., Liu, B., O’Leary, G.J., Olesen, J.E., Ottman, M.J., Priesack, E., Reynolds, M., Stratonovitch, P., Streck, T., Thorburn, P.J., Waha, K., Wall, G.W., White, J.W., Zhao, Z., Zhu, Y., 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. *Field Crop. Res.* 202, 5–20. <https://doi.org/10.1016/j.fcr.2016.05.001>
- Mairech, H., López-Bernal, Á., Moriondo, M., Dibari, C., Regni, L., Proietti, P., Villalobos, F.J., Testi, L., 2021. Sustainability of olive growing in the Mediterranean area under future climate scenarios: Exploring the effects of intensification and deficit irrigation. *Eur. J. Agron.* 129, 126319. <https://doi.org/10.1016/j.eja.2021.126319>
- Malone, B.P., McBratney, A.B., Minasny, B., 2013. Spatial Scaling for Digital Soil Mapping. *Soil Sci. Soc. Am. J.* 77, 890–902. <https://doi.org/10.2136/sssaj2012.0419>
- Martre, P., Wallach, D., Asseng, S., Ewert, F., 2015. Multimodel ensembles of wheat growth: many models are better than one. *Glob. Chang. Biol.* 21, 911–925. <https://doi.org/10.1111/gcb.12768>
- McClymont, L., Goodwin, I., Whitfield, D.M., O’Connell, M.G., 2019. Effects of within-block canopy cover variability on water use efficiency of grapevines in the Sunraysia irrigation region, Australia. *Agric. Water Manag.* 211, 10–15. <https://doi.org/10.1016/j.agwat.2018.09.028>
- Miguez, F.E., 2022. *apsimx: Inspect, Read, Edit and Run “APSIM” ‘Next Generation’ and “APSIM” Classic R package version 2.3.1.*
- Mon, J., Bronson, K.F., Hunsaker, D.J., Thorp, K.R., White, J.W., French, A.N., 2016. Interactive effects of nitrogen fertilization and irrigation on grain yield, canopy temperature, and nitrogen use efficiency in overhead sprinkler-irrigated durum wheat. *Field Crop. Res.* 191, 54–65. <https://doi.org/10.1016/j.fcr.2016.02.011>
- Moran, P.A.P., 1948. The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society: Series B (Methodological)* 10, 243–251. <https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>

- Mouazen, A.M., Alhwaimel, S.A., Kuang, B., Waive, T., 2014. Multiple on-line soil sensors and data fusion approach for delineation of water holding capacity zones for site specific irrigation. *Soil Tillage Res.* 143, 95–105. <https://doi.org/10.1016/j.still.2014.06.003>
- Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Delineation of Soil Management Zones for Variable-Rate Fertilization, in: *Advances in Agronomy*. Elsevier, pp. 175–245. <https://doi.org/10.1016/bs.agron.2017.01.003>
- Nendel, C., Berg, M., Kersebaum, K.C., Mirschel, W., Specka, X., Wegehenkel, M., Wenkel, K.O., Wieland, R., 2011. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol. Model.* 222, 1614–1625. <https://doi.org/10.1016/j.ecolmodel.2011.02.018>
- Oger, B., Vismara, P., Tisseyre, B., 2021. Combining target sampling with within field route-optimization to optimise on field yield estimation in viticulture. *Precis. Agric.* 22, 432–451. <https://doi.org/10.1007/s11119-020-09744-0>
- Ojeda, H., 2007. Irrigation qualitative de précision de la vigne. *Le Progrès Agricole et Viticole* 124, 133–141.
- Oliveira, D., Martins, L., Mora, A., Damásio, C., Caetano, M., Fonseca, J., Ribeiro, R.A., 2021. Data fusion approach for eucalyptus trees identification. *Int. J. Remote Sens.* 42, 4087–4109. <https://doi.org/10.1080/01431161.2021.1883198>
- Pasley, H., Brown, H., Holzworth, D., Whish, J., Bell, L., Huth, N., 2023. How to build a crop model. A review. *Agron. Sustain. Dev.* 43, 2. <https://doi.org/10.1007/s13593-022-00854-9>
- Pasquel, D., Cammarano, D., Roux, S., Castrignanò, A., Tisseyre, B., Rinaldi, M., Troccoli, A., Taylor, J.A., 2023a. Downscaling the APSIM crop model for simulation at the within-field scale. *Agric. Syst.* 212, 103773. <https://doi.org/10.1016/j.agry.2023.103773>
- Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., Taylor, J.A., 2022a. A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precis. Agric.* 23, 1489–1513. <https://doi.org/10.1007/s11119-022-09885-4>
- Pasquel, D., Roux, S., Tisseyre, B., Taylor, J.A., 2023b. A new metric to evaluate spatial crop model performances. Presented at the 14th European Conference on Precision Agriculture, Bologna, Italy, pp. 603–609. https://doi.org/10.3920/978-90-8686-947-3_76
- Pasquel, D., Roux, S., Tisseyre, B., Taylor, J.A., 2022b. Comparison of different aspatial and spatial indicators to assess performance of spatialized crop models at different within-field scales. Presented at the 15th International Conference on Precision Agriculture, Minneapolis, Minnesota, USA.
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30, 683–691.
- Pedroso, M., Taylor, J., Tisseyre, B., Charnomordic, B., Guillaume, S., 2010. A segmentation algorithm for the delineation of agricultural management zones. *Comput. Electron. Agric.* 70, 199–208. <https://doi.org/10.1016/j.compag.2009.10.007>
- Pereira, L.S., Perrier, A., Allen, R.G., Alves, I., 1999. Evapotranspiration: Concepts and Future Trends. *J. Irrig. Drain. Eng.* 125, 45–51. [https://doi.org/10.1061/\(ASCE\)0733-9437\(1999\)125:2\(45\)](https://doi.org/10.1061/(ASCE)0733-9437(1999)125:2(45))
- Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environ. Model. Software* 79, 214–232. <https://doi.org/10.1016/j.envsoft.2016.02.008>
- Porwollik, V., Müller, C., Elliott, J., Chryssanthacopoulos, J., Iizumi, T., Ray, D.K., Ruane, A.C., Arnoeth, A., Balkovič, J., Ciais, P., Deryng, D., Folberth, C., Izaurralde, R.C., Jones, C.D., Khabarov, N., Lawrence, P.J., Liu, W., Pugh, T.A.M., Reddy, A., Sakurai, G., Schmid, E., Wang, X., de Wit, A., Wu, X., 2017. Spatial and temporal uncertainty of crop yield aggregations. *Eur. J. Agron.* 88, 10–21. <https://doi.org/10.1016/j.eja.2016.08.006>
- R Core Team, 2022. R: A Language and Environment for Statistical Computing.
- Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009. AquaCrop - The FAO Crop Model to Simulate Yield Response to Water: II. Main Algorithms and Software Description. *Agron. J.* 101, 438–447. <https://doi.org/10.2134/agronj2008.0140s>
- Ramirez-Villegas, J., Koehler, A.-K., Challinor, A.J., 2017. Assessing uncertainty and complexity in regional-scale crop model simulations. *Eur. J. Agron.* 88, 84–95. <https://doi.org/10.1016/j.eja.2015.11.021>

- Rasche, L., Taylor, R.A.J., 2019. EPIC-GILSYM: Modelling crop-pest insect interactions and management with a novel coupled crop-insect model. *J. Appl. Ecol.* 1365-2664.13426. <https://doi.org/10.1111/1365-2664.13426>
- Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. *Nat. Commun.* 6, 5989. <https://doi.org/10.1038/ncomms6989>
- Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agric. For. Meteorol.* 170, 166–182. <https://doi.org/10.1016/j.agrformet.2012.09.011>
- Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J., Trnka, M., 2012. Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. *Field Crop. Res.* 133, 23–36. <https://doi.org/10.1016/j.fcr.2012.03.016>
- Roux, S., Gaudin, R., Tisseyre, B., 2019. Why does spatial extrapolation of the vine water status make sense? Insights from a modelling approach. *Agric. Water Manag.* 217, 255–264. <https://doi.org/10.1016/j.agwat.2019.03.013>
- Saint-Geours, N., Bailly, J.-S., Grelot, F., Lavergne, C., 2014. Multi-scale spatial sensitivity analysis of a model for economic appraisal of flood risk management policies. *Environ. Model. Software* 60, 153–166. <https://doi.org/10.1016/j.envsoft.2014.06.012>
- Saint-Geours, N., Lavergne, C., Bailly, J.-S., Grelot, F., 2012. Change of Support in Spatial Variance-Based Sensitivity Analysis. *Math. Geosci.* 44, 945–958. <https://doi.org/10.1007/s11004-012-9406-5>
- Şalap-Ayça, S., Jankowski, P., 2018. Analysis of the influence of parameter and scale uncertainties on a local multi-criteria land use evaluation model. *Stoch. Environ. Res. Risk Assess.* 32, 2699–2719. <https://doi.org/10.1007/s00477-018-1535-z>
- Scholes, R., Reyers, B., Biggs, R., Spierenburg, M., Duriappah, A., 2013. Multi-scale and cross-scale assessments of social–ecological systems and their ecosystem services. *Curr. Opin. Environ. Sustain.* 5, 16–25. <https://doi.org/10.1016/j.cosust.2013.01.004>
- Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go? *Eur. J. Agron.* 94, 25–35. <https://doi.org/10.1016/j.eja.2018.01.006>
- Sexton, J., Everingham, Y., Inman-Bamber, G., 2016. A theoretical and real world evaluation of two Bayesian techniques for the calibration of variety parameters in a sugarcane crop model. *Environ. Model. Software* 83, 126–142. <https://doi.org/10.1016/j.envsoft.2016.05.014>
- Shelia, V., Šimůnek, J., Boote, K., Hoogenboom, G., 2018. Coupling DSSAT and HYDRUS-1D for simulations of soil water dynamics in the soil-plant-atmosphere system. *J. Hydrol. Hydromech.* 66, 232–245. <https://doi.org/10.1515/johh-2017-0055>
- Shuai, G., Basso, B., 2022. Subfield maize yield prediction improves when in-season crop water deficit is included in remote sensing imagery-based models. *Remote Sens. Environ.* 272, 112938. <https://doi.org/10.1016/j.rse.2022.112938>
- Silva, J.V., Giller, K.E., 2020. Grand challenges for the 21st century: what crop models can and can't (yet) do. *J. Agric. Sci.* 158, 794–805. <https://doi.org/10.1017/S0021859621000150>
- Sinclair, T.R., Seligman, N.G., 1996. Crop Modeling: From Infancy to Maturity. *Agron. J.* 88, 698–704. <https://doi.org/10.2134/agronj1996.00021962008800050004x>
- Soltani, A., Sinclair, T.R., 2015. A comparison of four wheat models with respect to robustness and transparency: Simulation in a temperate, sub-humid environment. *Field Crop. Res.* 175, 37–46. <https://doi.org/10.1016/j.fcr.2014.10.019>
- Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop -The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. *Agron. J.* 101, 426–437. <https://doi.org/10.2134/agronj2008.0139s>
- Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. *European Journal of Agronomy* 18, 289–307. [https://doi.org/10.1016/S1161-0301\(02\)00109-0](https://doi.org/10.1016/S1161-0301(02)00109-0)

- Sudduth, K.A., Drummond, S.T., Kitchen, N.R., 2001. Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. *Comput. Electron. Agric.* 31, 239–264. [https://doi.org/10.1016/S0168-1699\(00\)00185-X](https://doi.org/10.1016/S0168-1699(00)00185-X)
- Tardaguila, J., Stoll, M., Gutiérrez, S., Proffitt, T., Diago, M.P., 2021. Smart applications and digital technologies in viticulture: A review. *Smart Agric. Technol.* 1, 100005. <https://doi.org/10.1016/j.atech.2021.100005>
- Taylor, J.A., Bates, T.R., 2013. A discussion on the significance associated with Pearson’s correlation in precision agriculture studies. *Precis. Agric.* 14, 558–564. <https://doi.org/10.1007/s11119-013-9314-9>
- Taylor, J.A., Tisseyre, B., Leroux, C., 2019. A simple index to determine if within-field spatial production variation exhibits potential management effects: application in vineyards using yield monitor data. *Precis. Agric.* 20, 880–895. <https://doi.org/10.1007/s11119-018-9620-3>
- Tedeschi, L.O., 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89, 225–247. <https://doi.org/10.1016/j.agsy.2005.11.004>
- Teixeira, E.I., Zhao, G., Ruiter, J. de, Brown, H., Ausseil, A.-G., Meenken, E., Ewert, F., 2017. The interactions between genotype, management and environment in regional crop modelling. *Eur. J. Agron.* 88, 106–115. <https://doi.org/10.1016/j.eja.2016.05.005>
- Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., Oomen, R., Russell, G., Ewert, F., Bergez, J.-E., Janssen, S., Wery, J., Van Ittersum, M.K., 2011. Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. *Agric. Ecosyst. Environ.* 142, 85–94. <https://doi.org/10.1016/j.agee.2010.05.007>
- Thorp, K.R., DeJonge, K.C., Kaleita, A.L., Batchelor, W.D., Paz, J.O., 2008. Methodology for the use of DSSAT models for precision agriculture decision support. *Comput. Electron. Agric.* 64, 276–285. <https://doi.org/10.1016/j.compag.2008.05.022>
- Tiefeldorf, M., 2000. Modelling spatial processes - The identification and analysis of spatial relationships in regression residuals by means of Moran’s I, Springer. ed.
- Tisseyre, B., Leroux, C., Pichon, L., Geraudie, V., Sari, T., 2018. How to define the optimal grid size to map high resolution spatial data? *Precis. Agric.* 19, 957–971. <https://doi.org/10.1007/s11119-018-9566-5>
- Tisseyre, B., Mazzoni, C., Fonta, H., 2008. Within-field temporal stability of some parameters in Viticulture: potential toward a site specific management. *OENO One* 42, 27. <https://doi.org/10.20870/oenone.2008.42.1.834>
- Tisseyre, B., McBratney, A.B., 2008. A technical opportunity index based on mathematical morphology for site-specific management: an application to viticulture. *Precis. Agric.* 9, 101–113. <https://doi.org/10.1007/s11119-008-9053-5>
- van Bussel, L.G.J., Ewert, F., Leffelaar, P.A., 2011. Effects of data aggregation on simulations of crop phenology. *Agric. Ecosyst. Environ.* 142, 75–84. <https://doi.org/10.1016/j.agee.2010.03.019>
- van Evert, F.K., Cammarano, D., Kempenaar, C., 2023. Introduction, in: Cammarano, D., van Evert, F.K., Kempenaar, C. (Eds.), *Precision Agriculture: Modelling*. Springer, Cham, pp. 3–24.
- van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—A review. *Field Crop. Res.* 143, 4–17. <https://doi.org/10.1016/j.fcr.2012.09.009>
- van Ittersum, M.K., Donatelli, M., 2003. Modelling cropping systems - Highlights of the symposium and preface to the special issues. *Eur. J. Agron.* 18, 187–197. [https://doi.org/10.1016/S1161-0301\(02\)00095-3](https://doi.org/10.1016/S1161-0301(02)00095-3)
- van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. *Eur. J. Agron.* 18, 201–234. [https://doi.org/10.1016/S1161-0301\(02\)00106-5](https://doi.org/10.1016/S1161-0301(02)00106-5)
- van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* 177, 105709. <https://doi.org/10.1016/j.compag.2020.105709>
- van Oijen, M., Ewert, F., 1999. The effects of climatic variation in Europe on the yield response of spring wheat cv. Minaret to elevated CO₂ and O₃: an analysis of open-top chamber experiments by means of two crop growth simulation models. *Eur. J. Agron.* 10, 249–264.

- Verdugo-Vásquez, N., Acevedo-Opazo, C., Valdés-Gómez, H., Pañitru-De la Fuente, C., Ingram, B., García de Cortázar-Atauri, I., Tisseyre, B., 2022. Identification of main factors affecting the within-field spatial variability of grapevine phenology and total soluble solids accumulation: towards the vineyard zoning using auxiliary information. *Precis. Agric.* 23, 253–277. <https://doi.org/10.1007/s11119-021-09836-5>
- Villa, A., Eckersten, H., Gaiser, T., Ahrends, H.E., Lewan, E., 2022. Aggregation of soil and climate input data can underestimate simulated biomass loss and nitrate leaching under climate change. *Eur. J. Agron.* 141, 126630. <https://doi.org/10.1016/j.eja.2022.126630>
- Wallach, D., 2011. Crop Model Calibration: A Statistical Perspective. *Agron. J.* 103, 1144–1151. <https://doi.org/10.2134/agronj2010.0432>
- Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2019. Working with dynamic crop models - Methods, tools and examples for agriculture and environment, 3rd ed. Elsevier.
- Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P.J., van Ittersum, M., Aggarwal, P.K., Ahmed, M., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., Sanctis, G.D., Dumont, B., Rezaei, E.E., Fereres, E., Fitzgerald, G.J., Gao, Y., Garcia, M., Gayler, S., Girousse, C., Hoogenboom, G., Horan, H., Izaurralde, R.C., Jones, C.D., Kassie, B.T., Kersebaum, C.C., Klein, C., Koehler, K., Maiorano, A., Minoli, S., Müller, C., Kumar, S.N., Nendel, C., Ripoche, D., Rötter, R.P., Semenov, M.A., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Wolf, J., Zhang, Z., 2018. Multimodel ensembles improve predictions of crop–environment–management interactions. *Glob. Chang. Biol.* 00, 1–12. <https://doi.org/10.1111/gcb.14411>
- Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Gourdain, E., Andrianasolo, F., Asseng, S., Basso, B., Buis, S., Crout, N., Dibari, C., Dumont, B., Ferrise, R., Gaiser, T., Garcia, C., Gayler, S., Ghahramani, A., Hiremath, S., Hoek, S., Horan, H., Hoogenboom, G., Huang, M., Jabloun, M., Jansson, P.-E., Jing, Q., Justes, E., Kersebaum, K.C., Klosterhalfen, A., Launay, M., Lewan, E., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Nariman Zadeh, H., Padovan, G., Olesen, J.E., Poyda, A., Priesack, E., Pullens, J.W.M., Qian, B., Schütze, N., Shelia, V., Souissi, A., Specka, X., Srivastava, A.K., Stella, T., Streck, T., Trombi, G., Wallor, E., Wang, J., Weber, T.K.D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao, C., Zhu, Y., Seidel, S.J., 2021. The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise. *Environ. Model. Software* 145, 105206. <https://doi.org/10.1016/j.envsoft.2021.105206>
- Wallach, D., Thorburn, P., Asseng, S., Challinor, A.J., Ewert, F., Jones, J.W., Rotter, R., Ruane, A., 2016. Estimating model prediction error: Should you treat predictions as fixed or random? *Environ. Model. Software* 84, 529–539. <https://doi.org/10.1016/j.envsoft.2016.07.010>
- Wallor, E., Kersebaum, K.-C., Ventrella, D., Bindi, M., Cammarano, D., Coucheney, E., Gaiser, T., Garofalo, P., Giglio, L., Giola, P., Hoffmann, M.P., Iocola, I., Lana, M., Lewan, E., Maharjan, G.R., Moriondo, M., Mula, L., Nendel, C., Pohankova, E., Roggero, P.P., Trnka, M., Trombi, G., 2018. The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crop. Res.* 228, 1–19. <https://doi.org/10.1016/j.fcr.2018.08.021>
- Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Liu, L., Müller, C., Naresh Kumar, S., Nendel, C., O’Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ripoche, D., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., Asseng, S., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nat. Plant.* 3, 17102. <https://doi.org/10.1038/nplants.2017.102>
- Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* 236, 111402. <https://doi.org/10.1016/j.rse.2019.111402>
- Xiang, Z., Bailey, R.T., Nozari, S., Husain, Z., Kisekka, I., Sharda, V., Gowda, P., 2020. DSSAT-MODFLOW: A new modeling framework for exploring groundwater conservation strategies in irrigated areas. *Agric. Water Manag.* 232, 106033. <https://doi.org/10.1016/j.agwat.2020.106033>

- Xu, H., Huang, F., Zuo, W., Tian, Y., Zhu, Y., Cao, W., Zhang, X., 2020. Impacts of Spatial Zonation Schemes on Yield Potential Estimates at the Regional Scale. *Agron.* 10, 631. <https://doi.org/10.3390/agronomy10050631>
- Yang, J.M., Yang, J.Y., Liu, S., Hoogenboom, G., 2014. An evaluation of the statistical methods for testing the performance of crop models with observed data. *Agr. Syst.* 127, 81–89. <https://doi.org/10.1016/j.agsy.2014.01.008>
- You, Y., Tian, H., Pan, S., Shi, H., Bian, Z., Gurgel, A., Huang, Y., Kicklighter, D., Liang, X.-Z., Lu, C., Melillo, J., Miao, R., Pan, N., Reilly, J., Ren, W., Xu, R., Yang, J., Yu, Q., Zhang, J., 2022. Incorporating dynamic crop growth processes and management practices into a terrestrial biosphere model for simulating crop production in the United States: Toward a unified modeling framework. *Agric. For. Meteorol.* 325, 109144. <https://doi.org/10.1016/j.agrformet.2022.109144>
- Zeng, W., Wu, J., Hoffmann, M.P., Xu, C., Ma, T., Huang, J., 2016. Testing the APSIM sunflower model on saline soils of Inner Mongolia, China. *Field Crop. Res.* 192, 42–54. <https://doi.org/10.1016/j.fcr.2016.04.013>
- Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. *Precis. Agric.* 13, 693–712. <https://doi.org/10.1007/s11119-012-9274-5>
- Zhang, N., Taylor, R.K., 2000. Applications of a field-level geographic information system (FIS) in precision agriculture. Presented at the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, USA. <https://doi.org/10.13031/2013.6829>
- Zhang, T., Su, J., Liu, C., Chen, W.-H., 2019. Bayesian calibration of AquaCrop model for winter wheat by assimilating UAV multi-spectral images. *Comput. Electron. Agric.* 167, 105052. <https://doi.org/10.1016/j.compag.2019.105052>
- Zhang, Y., 2023. Introducing Bayesian priors to semi-variogram parameter estimation using fewer observations. Presented at the 14th European Conference on Precision Agriculture, Bologna, Italy, pp. 651–658. https://doi.org/10.3920/978-90-8686-947-3_82
- Zhao, G., Hoffmann, H., van Bussel, L., Enders, A., Specka, X., Sosa, C., Yeluripati, J., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Zhao, Z., Nendel, C., Kiese, R., Eckersten, H., Haas, E., Vanuytrecht, E., Wang, E., Kuhnert, M., Trombi, G., Moriondo, M., Bindi, M., Lewan, E., Bach, M., Kersebaum, K., Rötter, R., Roggero, P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015. Effect of weather data aggregation on regional crop simulation for different crops, production conditions, and response variables. *Clim. Res.* 65, 141–157. <https://doi.org/10.3354/cr01301>
- Zhao, G., Hoffmann, H., Yeluripati, J., Xenia, S., Nendel, C., Coucheney, E., Kuhnert, M., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Kiese, R., Eckersten, H., Haas, E., Cammarano, D., Kassie, B., Moriondo, M., Trombi, G., Bindi, M., Biernath, C., Heinlein, F., Klein, C., Priesack, E., Lewan, E., Kersebaum, K.-C., Rötter, R., Paolo Roggero, P., Wallach, D., Asseng, S., Siebert, S., Gaiser, T., Ewert, F., 2016. Evaluating the precision of eight spatial sampling schemes in estimating regional means of simulated yield for two crops. *Environ. Model. Software* 80, 100–112. <https://doi.org/10.1016/j.envsoft.2016.02.022>
- Zhen, X., Huo, W., Tian, D., Zhang, Q., Sanz-Saez, A., Chen, C.Y., Batchelor, W.D., 2023. County level calibration strategy to evaluate peanut irrigation water use under different climate change scenarios. *Eur. J. Agron.* 143, 126693. <https://doi.org/10.1016/j.eja.2022.126693>
- Zhuo, W., Fang, S., Gao, X., Wang, L., Wu, D., Fu, S., Wu, Q., Huang, J., 2022. Crop yield prediction using MODIS LAI, TIGGE weather forecasts and WOFOST model: A case study for winter wheat in Hebei, China during 2009–2013. *Int. J. Appl. Earth Obs. Geoinf.* 106, 102668. <https://doi.org/10.1016/j.jag.2021.102668>

