

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *08/12/2020* par :

IVANA ALEKSOVSKA

**Améliorer les prévisions à court et moyen termes des modèles
agronomiques en prenant mieux en compte l'incertitude des prévisions
météorologiques**

JURY

SYLVAIN COQUILLAT	Professeur, Toulouse 3	Président
LILIANE BEL	Professeur, AgroParisTech	Rapporteur
STÉPHANIE MAHÉVAS	Cadre de Recherche, IFREMER	Rapporteur
PIERRE PINSON	Professeur, Technical University of Denmark	Rapporteur
SAMUEL BUIS	Ingénieur de Recherche, INRAE	Examineur
LAURENT DUBUS	Responsable scientifique, RTE	Examineur

École doctorale et spécialité :

SDU2E : Océan, Atmosphère et Surfaces Continentales

Unité de Recherche :

Nom du laboratoire (CNRM)

Directeur(s) de Thèse :

Laure RAYNAUD (Météo-France), Robert FAIVRE (INRAE) et François BRUN (ACTA)

Rapporteurs :

Liliane BEL, Stéphanie MAHÉVAS et Pierre PINSON

Remerciements

La réalisation de ma thèse a été avant tout le fruit d'un travail d'équipe. Dans ce cadre-là, je souhaite avant tout remercier les trois personnes qui m'ont accompagné tout au long de ce travail de 3 ans. Je voudrais donc dire un grand merci à ma directrice de thèse, Laure RAYNAUD, mais aussi à François BRUN et Robert FAIVRE pour leur encadrement ainsi que leurs précieux conseils que cela soit en présentiel ou à distance et pour leur contribution à mon développement scientifique et personnel.

Avec ma directrice, je retiendrai avant tout nos échanges quotidiens, professionnels ou personnels autour de nos nombreux repas. Elle m'a transmis son penchant pour les prévisions d'ensemble météorologique, une nouvelle monde pour moi si passionnant. La chance d'avoir allée à l'ECMWF pour présenter mes travaux et rencontrer de nombreux chercheurs du monde météorologique représente pour moi une véritable opportunité. Je souhaiterais souligner sa passion pour la photographie que j'ai grandement apprécié et j'ai ainsi bénéficié de ses nombreux conseils.

Jamais avare d'une petite blague pour redonner le sourire, ce fut un plaisir de travailler et collaborer avec Robert. Je le remercie vivement pour nos échanges de qualité sur les nombreux domaines qui ont entouré cette thèse. J'ai particulièrement apprécié nos échanges sur l'analyse de sensibilité, un domaine sur lequel je porte mon intérêt en particulier. Les journées de rencontre Mexico et la conférence mondiale sur ces méthodes SAMO ont été un moyen d'approfondir mes connaissances.

Un dernier merci, pour François. Je retiendrai surtout sa bonne humeur quotidienne et les différents déplacements que nous avons faits pour présenter ce travail notamment à nos partenaires IFV et Arvalis. Les nombreux échanges sur différents sujets ont comblé ma curiosité. Mon esprit de compétition a trouvé son bonheur dans les nombreuses activités diverses auxquelles nous avons participé, que ce soit au Hackathon Api-Agro au salon de l'agriculture ou au HackTaFerme dans le Gers. Programmer pendant 48 heures dans la grange d'une ferme a été une expérience extraordinaire.

Je voulais vous dire que vos encouragements et la poussée dans la dernière ligne droite de la rédaction de la thèse ont été d'un immense soutien scientifique et humain sans lequel ce manuscrit n'aurait pu être d'une aussi grande qualité.

Je souhaite remercier les membres du comité de thèse David Makowski (INRAE), Sophie Ricci (Cerfacs), Rémy Garçon (EDF), Frédéric Chavaux (Météo-France) qui m'ont fait l'honneur d'être les membres du comité de pilotage de ma thèse. Leurs remarques m'ont permis d'envisager mon travail sous un autre angle.

Je remercie ensuite l'ensemble des membres du jury, qui m'ont fait l'honneur de bien vouloir étudier avec attention mon travail.

Mon travail a porté notamment sur deux cas d'études, l'une sur le blé et la seconde sur la vigne. J'aimerais donc porter une affection particulière à Marc Raynal et Christian Debord de l'IFV - Institut Français de la Vigne et du Vin ainsi qu'à Emmanuelle Gourdain et Olivier Deudon d'Arvalis - Institut du Végétal pour leur aide et les discussions sur ces deux sujets particulièrement intéressants.

Cette thèse n'aurait pas non plus été possible sans l'aide de mes collègues de Météo France, merci à Michael Zamo pour la calibration d'EMOS mais aussi à Franck Souverain, Sébastien Prat et Pascal Raynaud pour leur aide concernant l'extraction des données de prévisions météorologiques. J'aurai aussi une pensée pour Grégoire Pigeon qui a été l'instigateur de cette thèse avec François et qui m'a permis de me plonger à la fois dans le domaine de l'agronomie et dans celui de la météorologie.

Et comment ne pas mentionner des gens de mon équipe, remercier mes collègues de l'équipe Recyf à Météo France : La bonne humeur de Carole. La sensation était étrange puis on s'habitue à entendre « dobar den » (bonjour macédonien) tous les jours. Les discussions avec Laurent mon collègue de bureau et nos échanges sur la prévision d'ensemble mais aussi les petits blagues dont il avait le secret. Si on reste dans ce domaine je n'oublierai pas les blagues de Pascal notamment mon "coach sportif" et les nombreuses séances de footing au lac de la Ramé et les séances de natation à la piscine pendant la pause pour décompresser. Le trail du viaduc de Millau et le semi-marathon de Toulouse que nous avons finis sont les preuves de nos talents de sportifs. A notre chef d'équipe Matthieu Plu pour son attention et sa compréhension en particulier qu'il apporte dans notre équipe, de même que notre chef d'unité Alain Joly.

Je ne saurais oublier mes collègues doctorants de Météo France, de INRAE et AGIR et de l'école doctorale notamment ceux de l'association Acadoc dans laquelle je faisais partie. Les moments de partage autour d'un verre ou d'un pique-nique à côté de la Garonne pour profiter du coucher du soleil sur le pont Saint Pierre et l'hôpital de la Grave ont été des temps de décompression et de rire que j'ai savourés dans les périodes de tension.

Mes remerciements iront aussi à ceux qui ont soutenu cette thèse, c'est à dire l'agence nationale de la recherche dans le cadre du programme investissements pour le futur, dénommé anr-16-conv-0004, à DigitAg. En premier lieu Véronique Bellon-Maurel et toute l'équipe de Digitag pour les nombreux échanges constructifs sur le monde du numérique et de l'agronomie. Les rencontres annuels "digital agriculture convergence lab" (www.hdigitag.fr/en) étaient très intéressantes et permettaient une rencontre entre doctorants. Cette thèse a également été soutenue par le projet METEOPREC, financé par CASDAR, "l'apport d'une prévision précise aux agriculteurs" (les CASDAR sont financés par le ministère français de l'agriculture) et le réseau modélisation et analyse des données pour l'agriculture (www.modelia.org).

Je tiens également à remercier l'Université Paul Sabatier et l'école doctorale SDU2E en premier lieu avec Geneviève Soucail.

Je remercie également mes amis qui ont pu venir et m'ont apporté leurs soutiens mais aussi pouvoir, ensemble, profiter de la ville rose. J'ajouterai quelques mots tout particulièrement pour ma famille, qui m'a encouragé dans tous les moments et situations.

Table des matières

Introduction	1
I Données, outils et objectif	3
1 La Prévision Numérique du Temps (PNT)	7
1.1 Principe de la PNT	7
1.2 Vers une prévision des incertitudes de prévision : la prévision d'ensemble . . .	8
1.3 Calibration des prévisions d'ensemble	9
1.4 Les modèles de PNT mobilisés	10
2 Outils d'aide à la décision (OAD)	13
2.1 Principe des OAD	13
2.2 Les cas d'études des OAD utilisés	14
2.3 Simulations agronomiques d'ensemble	17
3 Objectifs et plan de la thèse	21
II Méthodes	25
4 Calibration des prévisions météorologiques d'ensemble	29
4.1 Méthodes paramétriques	29
4.2 Méthodes non-paramétriques	31
4.3 Prévision d'ensemble calibrée	31
5 Construction d'une prévision d'ensemble cohérente	33
5.1 Principe de la construction d'un ensemble cohérent	33

5.2	Méthodes de construction d'un ensemble cohérent	34
5.3	La distance "Déformation Temporelle Dynamique"	35
6	Vérification de la prévision probabiliste - application aux prévisions d'ensemble météorologique et agronomique	39
6.1	Qu'est-ce qu'une bonne prévision ?	40
6.2	Vérification des prévisions probabilistes : approche orientée distribution . . .	40
6.3	Vérification de prévision probabiliste : approche orientée mesures	42
6.4	Test des rangs signés de Wilcoxon	46
III	Résultats	47
7	Plan d'expérience numérique	51
7.1	Mise en œuvre des simulations agronomiques	51
7.2	Vérification probabiliste des prévisions météorologiques et agronomiques . . .	55
7.3	Synthèse	56
8	Apport de la prévision d'ensemble pour les simulations agronomiques	59
8.1	Analyse graphique des données de prévision d'ensemble et des données fréquentielles	60
8.2	Propagation des incertitudes météorologiques dans les OAD	63
8.3	Évaluation agronomique de la prévision d'ensemble par rapport aux données fréquentielles	69
8.4	Propagation des incertitudes des données météorologiques de PNT à haute résolution pour des simulations agronomiques	76
8.5	Conclusion	77
9	Construction des prévisions d'ensemble cohérentes	79
9.1	Vérification des prévisions calibrées	80
9.2	Construction et caractéristiques des prévisions d'ensemble cohérentes	82

9.3	La calibration : étape incontournable dans la construction des scénarios cohérents	85
9.4	Conclusion	87
10	Évaluation météorologique et agronomique des prévisions d'ensemble cohérentes	91
10.1	Vérification météorologique des prévisions d'ensemble cohérentes	91
10.2	Vérification agronomique des prévisions d'ensemble cohérentes	95
10.3	Conclusion	102
	Conclusion, discussion et perspectives	105
A	Plateformes disponibles actuellement utilisées incluant les deux cas d'étude de la thèse	109
A.1	EPICURE	109
A.2	Taméo	109
B	Article vulgarisé - Végéphy1 - 12e Conférence Internationale sur les Maladies des Plantes (apparu en décembre 2018)	113
C	Article crop protection - version en cours de révision après la soumission (soumis en décembre 2019)	127
D	Article Weather and forecasting - version soumise (soumis en février 2021)	143
E	Travaux complémentaires	191
E.1	Analyse de sensibilité du modèle SeptoLIS aux variables d'entrée agronomiques et météorologiques	191
E.2	Prise en compte des incertitudes des prévisions météorologiques à court terme pour les outils d'aide à la décision pour la vigne et le maïs	193
	Bibliographie	198

Table des figures

1.1	Illustration du principe de la prévision d'ensemble (EPS-Ensemble Prediction System)). Une prévision différente (représentée ici sous la forme d'une série temporelle de température en un point donné) est générée à partir de chaque état initial perturbé (en pratique, d'autres sources d'incertitude, comme celle liée au modèle de prévision, peuvent être prises en compte mais ne sont pas représentées sur ce graphique).	9
2.1	Cycle de développement avec les différents stades du ver de la grappe, ravageur de la vigne. Source : (Maher <i>et al.</i> , 2006)	15
2.2	Présentation du modèle SeptoLIS, schéma fonctionnel adapté Source : (Hervé, 2011)	17
2.3	Représentation schématique des deux étapes : analyse et prévision	18
2.4	Propagation des incertitudes des prévisions météorologiques. Un ensemble de simulations agronomiques est généré par propagation des prévisions météorologiques d'ensemble dans un modèle agronomique : en sortie, chaque membre agronomique utilise une prévision météorologique différente.	19
5.1	Représentation schématique du raccordement selon l'approche "NN". Les trajectoires de la température AROME-EPS (les courbes violettes) sont couplées à J+2 avec les prévisions les plus proches d'ARPEGE-EPS (courbes toutes pleines oranges) ou d'IFS-EPS (courbes toutes pleines vertes). A J+4, seules les trajectoires ARPEGE-EPS sont couplées avec les trajectoires IFS-EPS disponibles. La flèche rouge indique les courbes qui sont choisies pour la jointure. Les courbes en pointillé représentent des membres non sélectionnés pendant le processus.	35
5.2	a) Deux séries temporelles S_1 et S_2 , b) alignement des séries S_1 et S_2 , le point i de la série S_1 est associé au point j de la série S_2 , c) matrice de distance cumulée dans un voisinage (fenêtre de Sakoe-Chiba en vert) et chemin optimal en noir. Source : (Najmeddine <i>et al.</i> , 2012).	36
6.1	A gauche la distribution de probabilité de la prévision (courbe rouge) et la référence (trait vertical noir), à droite les fonctions de répartition associées. Le CRPS correspond à l'aire entre les deux fonctions de répartition (surface grisée). L'axe horizontal représente la variable à prévoir. Source : Hersbach, 2000	44

6.2	Représentation schématique de la courbe ROC(en rouge) et le seuil s (point noir), par rapport à la bissectrice (en noir). En abscisse est représenté le taux de fausse alarme et en ordonnée le taux de détection.	45
7.1	Sites utilisés pour l'évaluation des prévisions agronomiques. Gauche : vignobles dans des zones présentant des problématiques de ver de la grappe. Droite : parcelles de blé dans des zones avec des problématiques de septoriose.	52
7.2	Simulation avec EVA pour le site de Châteauneuf-du-Pape. La partie analyse (en noir) utilise les observations horaires de la température à 2 mètres entre la date d'initialisation (20/09/2017) et la veille de la date de prévision (24/07/2018). La partie prévision (en bleu) utilise les données de prévision de température à 2 mètres d'un membre de la prévision d'ensemble IFS-EPS du 24/07/2018 jusqu'au 07/08/2018 (échéance maximale des prévisions IFS-EPS). En ordonnée : le cumul du stade œuf. La date de traitement est obtenue en fonction du franchissement de seuil choisi, les seuils 2% et 15% sont représentés ici.	54
8.1	Données de température à 2 mètres pour le site de Bordeaux. Les données observées sont en gris. (a) Prévision d'ensemble du modèle IFS-EPS (en turquoise) et prévision déterministe du modèle IFS (en noir), initialisées le 15/06/2018, sur leurs échéances respectives de 15 et 10 jours. (b) Données fréquentielles (observations des années 1998 à 2017) du 15/06 au 29/06 (en bleu).	61
8.2	Données de précipitations quotidiennes sur le site de Poitiers Biard (86). Les données observées sont en gris. (a) Prévision d'ensemble du modèle IFS-EPS (sous la forme de boîtes à moustache, en turquoise) et prévision déterministe du modèle IFS (en noir), initialisées le 10/04/2019, sur leurs échéances respectives de 15 et 10 jours. (b) Données fréquentielles (observations des années 1998 à 2017) du 10/04 au 24/04 (en bleu).	62
8.3	(a) Résultats des simulations d'ensemble EVA pour l'évolution du stade œuf à Bordeaux, obtenus en utilisant la prévision déterministe IFS (en noir) et les prévisions d'ensemble IFS-EPS (en turquoise). La simulation de référence, couplée uniquement aux observations, est superposée en gris. Les simulations débutent le 15 juin 2018 et couvrent des échéances de 10 jours (avec IFS) ou 15 jours (avec IFS-EPS). Histogrammes des dates de déclenchement du traitement (b) pour le seuil 2 % et (c) pour le seuil 15 %. La catégorie "beyond_det" indique que le seuil n'a pas été franchi à la fin de la dernière date de validité de la prévision déterministe.	65

8.4	(a) Résultats des simulations d'ensemble EVA pour l'évolution du stade œuf à Bordeaux, obtenus en utilisant les données fréquentielles. La simulation de référence, couplée uniquement aux observations, est superposée en gris. Les simulations débutent le 15 juin 2018 et couvrent des échéances de 15 jours. Histogrammes des dates de déclenchement du traitement (b) pour le seuil 2% et (c) pour le seuil 15%. La catégorie "beyond_eps" indique que le seuil n'a pas été franchi à la fin de la dernière date de validité de la prévision probabiliste.	67
8.5	Histogramme des dates du traitement T1 prévues par Septo-LIS [®] à Poitiers Biard, obtenu en utilisant les prévisions IFS-EPS à 15 jours d'échéance. La date obtenue en utilisant la prévision déterministe IFS est en noir et la date de référence en gris. Les simulations débutent le 10 avril 2019.	68
8.6	Histogramme des dates du traitement T1 prévues par Septo-LIS [®] à Poitiers Biard, obtenu en utilisant les données fréquentielles (en bleu) et la date de référence (en gris). Les simulations débutent le 10 avril 2019.	68
8.7	Résultats des simulations d'ensemble Septo-LIS [®] à Poitiers Biard, obtenus en utilisant les données de prévisions d'ensemble (colonne gauche) et les données fréquentielles (colonne droite). La date de début des données probabiliste est : (a) et (b) 2019/04/12, (c) et (d) 2019/04/15, (e) et (f) 2019/04/16, (g) et (h) 2019/04/17, (i) et (j) 2019/04/18	70
8.8	Biais de la moyenne des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations EVA d'ensemble utilisant les prévisions IFS-EPS (0.BRUT, en turquoise) et les données fréquentielles (1.FREQ, en bleu), pour la deuxième génération et en utilisant le seuil de décision de 15%. Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	71
8.9	Couverture d'observation des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Elle est calculée à partir des simulations EVA d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération pour un seuil de décision 15%. Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. . .	72

8.10	Médiane des IPR90 des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations EVA d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération et en utilisant le seuil de décision 15%. Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	73
8.11	CRPS des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les CRPS sont calculés à partir des simulations EVA d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération pour un seuil de décision de 15%. Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	73
8.12	Biais de la moyenne des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les biais sont calculés à partir des simulations Septo-LIS [®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	74
8.13	Médiane des IPR90 des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations Septo-LIS [®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	75
8.14	Couverture d'observation des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Elle est calculée à partir des simulations Septo-LIS [®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019.	75

8.15	CRPS des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les CRPS sont calculés à partir des simulations Septo-LIS [®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.	76
8.16	(a) Prévisions des températures horaires fournies par les systèmes AROME, ARPEGE, IFS et leurs équivalents ensemblistes AROME-EPS, ARPEGE-EPS et IFS-EPS (boîtes à moustaches) pour les 2 et 3 juillet 2018 à Bordeaux. (b) Prévisions des températures quotidiennes moyennes fournies par les systèmes AROME, ARPEGE, IFS et leurs équivalents ensemblistes AROME-EPS, ARPEGE-EPS et IFS-EPS (boîtes à moustaches) pour les 2 et 3 juillet 2018 à Bordeaux. Les prévisions sont initialisées le 2 juillet 2018. (c) Proportions d'œufs associées, simulées par EVA.	78
9.1	Biais (à gauche) et RMSE (à droite) de la moyenne d'ensemble des prévisions de température à 2 mètres, en fonction de l'échéance de prévision. Les scores sont calculés pour les 11 sites de tableau 7.1 sur la période 01/06/2018-31/05/2019, avec les ensembles bruts (ligne continue) et calibrés (ligne pointillée). En haut : Arome-EPS, au milieu : ARPEGE-EPS et en bas : IFS-EPS. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble post-traité est statistiquement meilleure (resp. moins bonne) que la performance de l'ensemble brut, selon le test de Wilcoxon.	80
9.2	CRPS des prévisions de température de 2 mètres en fonction de l'échéance de prévision. Les scores sont calculés pour les 11 sites décrits dans le tableau 7.1 sur la période 01/06/2018-31/05/2019 pour les trois ensembles bruts (haut) et calibrés (bas). Les signes + (resp. -) indiquent que la performance d'Arome-EPS et d'ARPEGE-EPS est statistiquement meilleure (resp. moins bonne) que la performance de d'IFS-EPS, selon le test de Wilcoxon.	81
9.3	Les distances entre membres AROME et ARPEGE calculées sur la période commune (de 0h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1). (a) Distributions des distances entre les membres AROME-EPS et leurs couples (bleu), entre les membres AROME-EPS et le membre ARPEGE-EPS le plus proche (turquoise) et entre les membres AROME-EPS et tous les membres ARPEGE-EPS (rouge), calculées sur la période 01/02/2018-28/02/2019 en moyenne pour tous les sites du tableau 7.1.	83
9.4	Distribution de la taille effective des ensembles NN, calculée pour 11 sites du tableau 7.1 sur la période 06/01/2018-31/05/2019, gauche : après le raccordement à 48h et droite : après le raccordement à 96h.	84

9.5	Distributions (sous forme de boîtes à moustaches) des différences absolues de température prévue (moyennées sur tous les membres de l'ensemble) entre les échéances 48h et 49h (en noir) et les échéances 96h et 97h (en gris), calculées pour 11 sites sur la période 06/01/2018-31/05/2019. Ces distributions sont présentées, de gauche à droite, pour les membres IFS-EPS et pour les membres des ensembles cohérents RN, NN et HU.	85
9.6	Ensemble de scénarios cohérents construit avec (a) la méthode HU, (b) la méthode NN. Date de début de la prévision 25/05/2018, à Turckheim. Chaque membre cohérent est initialisé par une prévision AROME-EPS (en rose) puis raccordé à un membre ARPEGE-EPS (en orange) ou IFS-EPS (en bleu) à 48h. Un deuxième raccordement à un membre IFS-EPS est effectué à 96h pour certains membres. Les distances entre membres sont calculées sur les 7 dernières heures de la période commune (de 42h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1)	86
9.7	Date de début de la prévision 25/05/2018 à Turckheim (a) ACP-Analyse en composante principale pour les 3 modèles AROME-EPS (ARO, en rose), ARPEGE-EPS (ARP, en orange) et IFS-EPS (IFS, en turquoise), (b) PVE - Pourcentage de la variance expliquée.	88
9.8	(a) Ensemble de scénarios cohérents construit avec la méthode NN. Date de début de la prévision 25/05/2018, à Turckheim. Chaque membre cohérent est initialisé par une prévision AROME-EPS (en rose) puis raccordé à un membre ARPEGE-EPS (en orange) ou IFS-EPS (en bleu) à 48h. Un deuxième raccordement à un membre IFS-EPS est effectué à 96h pour certains membres. Les distances entre membres sont calculées sur les 7 dernières heures commune (de 42h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1). (b) Histogrammes des membres choisis à J+2. (c) Histogrammes des membres choisis à J+4.	89
10.1	Évolution du score de Brier des prévisions d'ensemble de température à 2 mètres en fonction des échéances, calculé pour 11 sites du tableau 7.1 sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision pour l'évènement $\epsilon := \mathbb{T} \geq 15^\circ\text{C}$. Les résultats d'IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) Une valeur par jour (valide à 24h, 48h, etc.), (b) Zoom sur les premières 48h d'échéance. Les signes + (resp. -) indiquent que la performance d'un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d'IFS-EPS, selon le test de Wilcoxon.	93

10.2	Évolution du CRPS des prévisions d'ensemble de température à 2 mètres en fonction des échéances, calculé pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision. Les résultats d'IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) Une valeur par jour (valide à 24h, 48h, etc.), (b) Zoom sur les premières 48h d'échéance. Les signes + (resp. -) indiquent que la performance d'un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d'IFS-EPS, selon le test de Wilcoxon.	94
10.3	Courbes ROC des prévisions d'ensemble de température à 2 mètres calculées pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision pour l'évènement $\epsilon := \mathbb{T} \geq 15^\circ\text{C}$. Les résultats d'IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) ROC calculée sur les échéances 0 à 48h; (b) ROC calculée sur les échéances 0 à 360h.	96
10.4	Évolution de l'aire sous la courbe ROC (AUC) pour les prévisions de température à 2 mètres supérieures 15°C , calculée pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision (une valeur par jour). Les résultats d'IFS-EPS et des trois ensembles de scénarios cohérents sont superposés. Les signes + (resp. -) indiquent que la performance d'un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d'IFS-EPS, selon le test de Wilcoxon.	97
10.5	Diagramme de fiabilité pour les prévisions de température à 2 mètres supérieures à 15°C , calculé pour 11 sites sur la période 06/01/2018-31/05/2019. (a) calculé sur les échéances 0 à 48h, (b) calculé sur les échéances 0 à 360h. Les résultats d'IFS-EPS et des trois ensembles scénarios cohérents sont superposés.	98
10.6	Évolution du CRPS des prévisions de température à 2 mètres, calculé pour 11 sites sur la période 01/06/2018-31/05/2019, en fonction de l'échéance de prévision. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue). Les signes + (resp. -) indiquent que la performance du raccordement à 24h est statistiquement meilleure (resp. moins bonne) que celle du raccordement à 48h, selon le test de Wilcoxon.	99
10.7	Évolution de l'aire sous la courbe ROC (AUC) pour les prévisions de température de 2 mètres supérieures à 15°C , calculée pour les 11 sites sur la période 06/01/2018-31/05/2019, en fonction du jour de début de la prévision. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue). Les signes + (resp. -) indiquent que la performance de la fusion sur 24h est statistiquement meilleure (resp. pire) que la performance du IFS-EPS, selon le test de Wilcoxon.	99

10.8	Diagramme de fiabilité pour les prévisions de température à 2 mètres au-dessus de 15°C, calculé pour les 11 sites sur la période 06/01/2018-31/05/2019 pour les jointures à 24h et 48h. (a) échéances 0-48h, (b) échéances 0-360h. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue).	100
10.9	Biais de la date moyenne de traitement prévue par l'ensemble de simulations EVA, en fonction de l'intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les scores sont moyennés sur les 8 sites et les années 2018 et 2019. Les biais moyens sont calculés à partir de simulations d'EVA d'ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d'ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).	101
10.10	CRPS des dates de traitement prévues par l'ensemble de simulations EVA, en fonction de l'intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les scores sont calculés en moyenne sur les 8 sites et les années 2018 et 2019. Les CRPS moyens sont calculés à partir de simulations d'EVA d'ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d'ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).	102
10.11	Évolution de la couverture des observations, en fonction de l'intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les pourcentages de couverture d'observation sont calculés en moyenne sur les 8 sites et les années 2018 et 2019. Les couvertures des observations sont calculés à partir de simulations d'EVA d'ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d'ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).	103
A.1	Exemple des sorties de EPICURE	110
A.2	Exemple des sorties de Taméo	111

Liste des tableaux

1.1	Modèles de prévision déterministe utilisés de façon opérationnelle à Météo-France sur la période mars 2018 - juin 2019 considérée dans la suite du travail.	11
1.2	Systèmes de prévision d'ensemble utilisés de façon opérationnelle à Météo-France sur la période mars 2018 - juin 2019 considérée dans la suite du travail.	12
6.1	Relations entre les critères de qualité, la distribution et les scores	42
7.1	Les points spatiaux utilisés pour des simulations agronomiques sur les deux cultures de vigne et de blé. Les lignes en gris correspondent aux parcelles qui ont été choisies pour des évaluations météorologiques.	53
7.2	Récapitulatif des données météorologiques et des sorties agronomiques associées, utilisées dans la suite du travail. Les données météorologiques sont : la température horaire de l'air à 2 mètres (T), la température minimale (T_{\min}), maximale (T_{\max}), et le cumul des précipitations ($\mathbb{R}\mathbb{R}$) au pas de temps journalier. Les prévisions déterministes proviennent des modèles AROME, ARPEGE et IFS, et les prévisions d'ensemble des modèles AROME-EPS, ARPEGE-EPS et IFS-EPS. Les observations sont extraites des archives de Météo-France (T) et d'Arvalis (T_{\min} , T_{\max} , $\mathbb{R}\mathbb{R}$). Pour les données de prévision, elles proviennent des bases de données de Météo-France (AROME, AROME-EPS, ARPEGE, ARPEGE-EPS) et du CEPMMT (IFS et IFS-EPS).	57
8.1	Moyennes et écarts types des IPR90 sur les données de température (en degré) calculés pour la période d'intérêt d'EVA (de 1 mai-19 septembre) pour les années 2018 et 2019 sur l'ensemble des sites présentés dans le tableau 7.1. Les valeurs sont présentées en fonction de l'échéance de prévision, pour la prévision d'ensemble et les données fréquentielles.	63
8.2	Moyennes et écarts types des IPR90 sur les données de précipitations quotidiennes (en mm) calculés pour le période d'intérêt de Septo-LIS [®] (de 10 avril-30 avril) pour les années 2018 et 2019 sur l'ensemble des sites présentés dans le tableau 7.1. Les valeurs sont présentées en fonction de l'échéance de prévision, pour la prévision d'ensemble et les données fréquentielles.	64
9.1	Pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le raccordement à 48h avec les différentes stratégies d'affectation, pour quatre saisons. Les valeurs théoriques pour la sélection aléatoire sont de 40,7% pour ARPEGE-EPS et 59,3% pour IFS-EPS, en raison de leur taille respective. . .	84

10.1	CRPS moyen des différentes prévisions d'ensemble cohérentes et d'IFS-EPS pour différentes saisons et échéances (D-D+1 représente les échéances 0-24h, D+1-D+2 les échéances 24-48h etc.).	92
10.2	Pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le raccordement à 24h avec les différentes stratégies d'affectation, pour quatre saisons. Les valeurs théoriques pour la sélection aléatoire sont de 40,7% pour ARPEGE-EPS et 59,3% pour IFS-EPS, en raison de leur taille respective. . .	97

Introduction

L'agriculture est l'une des activités économiques les plus dépendantes des conditions météorologiques : la mise en place des cultures, le développement et la croissance des plantes, la gestion de l'irrigation ainsi que la protection des cultures sont des exemples parmi d'autres. Ainsi, la consultation des prévisions météorologiques par les agriculteurs est une préoccupation majeure pouvant aller à plusieurs fois par semaine, d'après l'enquête "Agrinautes Agrisurfeurs 2014" BVA Ticagri (Isaac et Pouyat, 2015).

En ce qui concerne la protection des cultures, les maladies et les ravageurs des cultures évoluent en fonction des conditions météorologiques qui conditionnent leurs dynamiques et l'importance des épidémies (Chaulagain *et al.*, 2019; Moyer *et al.*, 2016; Pietravalle *et al.*, 2003; Te Beest *et al.*, 2008). Des traitements phytosanitaires permettent de maîtriser une partie des épidémies, mais ils présentent des impacts importants sur l'environnement et la santé humaine. Afin de limiter drastiquement l'usage de ces moyens de lutte, leur application doit se faire selon des spécifications réglementaires et pratiques, en raisonnant leur recours uniquement aux situations l'exigeant. En France, la mise en œuvre du plan « Ecophyto » depuis 2008 permet de répondre à ces demandes sociétales et comporte des engagements sur la réduction des pesticides et les outils pour y arriver. Le plan « Ecophyto II » a revu les objectifs afin d'atteindre une réduction de l'usage des produits phytopharmaceutiques de 50 % d'ici 2025. Sur les moyens, le plan renforcé « Ecophyto II+ » encourage le développement et l'utilisation des outils numériques dont les outils de modélisation prédictive, en s'appuyant sur la recherche et l'innovation. Ces outils doivent permettre d'améliorer l'analyse du risque et aider ainsi les agriculteurs à réduire l'application des produits phytosanitaires (Écophyto, 2020).

Afin de réduire l'usage de produits phytosanitaires, différents outils peuvent aider les agriculteurs à appuyer leur raisonnement à la fois au niveau stratégique (choix d'un traitement chimique ou de biocontrôle), tactique et opérationnel (Rossi *et al.*, 2012). Au niveau tactique et opérationnel, en cours de campagne, il est important de prendre en compte la situation sanitaire courante et son évolution à venir afin de prendre une décision d'intervention ou non. Un enjeu important concerne le positionnement, au bon moment, d'un éventuel traitement en fonction de la dynamique épidémique et du mode d'action des produits. Pour cela, différents outils d'aide à la décision (OAD) sont proposés pour accompagner les agriculteurs dans ce processus de décision (Pertot *et al.*, 2017). Ces outils, aux formalismes variés, s'appuient sur la modélisation des dynamiques épidémiques et prennent comme entrées les conditions météorologiques, notamment la température, la pluie ou encore l'humidité relative qui pilotent en grande partie les dynamiques des ravageurs et des maladies. S'agissant de phénomènes cumulatifs qui dépendent des conditions météorologiques passées mais aussi à venir, il est nécessaire de disposer des séries temporelles multivariées combinant à la fois des observations météorologiques et des prévisions. Afin d'affiner le raisonnement à la parcelle, les données météorologiques à haute résolution spatiale sont déterminantes pour prédire l'évolution sanitaire et faciliter la gestion en prenant en compte les différentes contraintes, des prévisions à court

et moyen terme sont également nécessaires.

La fiabilité de ces OAD reste largement imparfaite, notamment en lien avec la complexité des systèmes biologiques qu'ils représentent et des incertitudes sur les entrées considérées. Concernant l'incertitude liée aux conditions météorologiques, celle-ci semble particulièrement importante par rapport aux autres composantes de l'incertitude (Rossing *et al.*, 1994). Plusieurs stratégies ont été proposées pour prendre en compte cette incertitude dans les OAD, incluant notamment l'approche dite "fréquentielle" (Linville, Drye *et al.*, 1995), qui utilise plusieurs scénarios météorologiques extraits de l'archive d'observations, ou l'utilisation de générateurs stochastiques de données météorologiques (Lawless et Semenov, 2005 ; Semenov, 2004) afin d'avoir un aperçu de l'ensemble des possibles météorologiques. La principale limitation de ces approches est qu'elles échantillonnent une variabilité inter-annuelle plutôt que l'incertitude associée aux conditions météorologiques actuelles.

Afin d'estimer l'incertitude "du jour", plusieurs centres météorologiques dans le monde, dont Météo-France, ont développé des systèmes de prévision probabiliste, dont l'objectif est de prévoir la distribution de probabilité de l'état de l'atmosphère, contrairement aux prévisions déterministes qui ne fournissent qu'une seule valeur de prévision. Ces prévisions probabilistes fournissent un complément d'information utile pour aider à la prise de décision et sont devenues aujourd'hui un outil majeur pour les prévisionnistes dans des domaines aval (Fundel *et al.*, 2019), en particulier l'hydrologie (Bellier *et al.*, 2017, 2018 ; Roulin et Vannitsem, 2005 ; Roulin, 2006), la gestion du réseau électrique (Pinson *et al.*, 2009), le routage des avions (Cheung *et al.*, 2015), et les feux de forêt (Worsnop *et al.*, 2019). Les applications en agriculture sont en revanche beaucoup plus réduites. A l'échelle de la prévision saisonnière on peut citer les travaux de Zhao *et al.* (2019) et Ogutu *et al.* (2018) et le potentiel des prévisions probabilistes a été examiné par Christ *et al.* (2015) dans le cas de la conduite de parcelles de coton.

Ce travail de thèse vise à évaluer l'intérêt des prévisions d'ensemble pour le monde agricole. En particulier, il s'agira de rechercher par quel moyen les OAD utilisés en protection des cultures peuvent bénéficier des informations issues des prévisions d'ensemble pour mieux prendre en compte les incertitudes associées aux entrées météorologiques. Ce travail se situe ainsi à l'interface entre les prévisionnistes météorologistes et tout particulièrement ceux impliqués dans les prévisions probabilistes et les ingénieurs et conseillers agricoles proposant des outils de gestion comme les OAD basés sur des modèles pour accompagner les agriculteurs vers une gestion raisonnée de leurs traitements phytosanitaires.

Première partie

Données, outils et objectif

Plan de la partie I

Dans l'introduction nous avons présenté le contexte de ce travail de thèse au travers des deux enjeux principaux, socio-environnementaux et météorologiques, ainsi que les problématiques qui ont motivé ce travail de recherche. Cette partie est consacrée à la présentation des données météorologiques (chapitre 1) et des outils agronomiques (chapitre 2) utilisés pour ce travail, ainsi qu'à la présentation détaillée des objectifs de la thèse (chapitre 3).

La Prévision Numérique du Temps (PNT)

Sommaire

1.1	Principe de la PNT	7
1.2	Vers une prévision des incertitudes de prévision : la prévision d'ensemble	8
1.3	Calibration des prévisions d'ensemble	9
1.4	Les modèles de PNT mobilisés	10
1.4.1	Modèles de prévision déterministes	10
1.4.2	Systèmes de prévision d'ensemble	10

Ce chapitre présente les fondamentaux de la prévision numérique du temps, ainsi que les concepts de prévision déterministe et probabiliste. Le principe de la prévision d'ensemble, qui est au cœur de ce travail de thèse, est également détaillé. Enfin, les caractéristiques des différents systèmes de prévision utilisés dans la suite sont présentées.

1.1 Principe de la PNT

La prévision du temps opérationnelle s'appuie sur la modélisation mathématique : il s'agit de résoudre numériquement un système d'équations complexes (Navier-Stokes), qui gouverne l'évolution de l'atmosphère au cours du temps. Ces équations constituent le cœur des modèles de prévision, et permettent de simuler l'évolution de variables telles que le vent, la pression, la température et l'humidité. Il est également nécessaire d'estimer l'état initial de la prévision, en utilisant notamment les observations *in situ* (e.g., stations de mesure au sol et en mer, radiosondes) et satellitaires disponibles sur le globe, et des algorithmes dits d'assimilation de données en temps réel (Courtier *et al.*, 1994; Thépaut et Courtier, 1991).

La résolution de ces équations fait appel à des méthodes numériques qui consistent à discrétiser l'atmosphère sur un maillage tridimensionnel et temporel. La finesse de cette discrétisation détermine l'échelle spatiale des phénomènes représentés explicitement par le modèle : plus la taille de la maille (également appelée résolution spatiale) est petite, meilleure est la représentation des phénomènes de petite échelle, mais plus le calcul est coûteux. L'échéance

d'une prévision, définie par l'intervalle de temps entre le début et la fin de validité de la prévision, est un autre élément de configuration important. Les caractéristiques d'un modèle de prévision et de sa résolution numérique répondent ainsi à un compromis, permettant d'atteindre une performance satisfaisante en des temps de calcul raisonnables.

Des prévisions météorologiques sont exécutées quotidiennement par plusieurs centres de prévision dans le monde, chacun utilisant un modèle particulier, configuré sur des domaines et des grilles variés et une procédure d'assimilation de données propre à chacun.

1.2 Vers une prévision des incertitudes de prévision : la prévision d'ensemble

Depuis la mise en place des premiers modèles, les prévisions ont connu des améliorations régulières (Bauer *et al.*, 2015), notamment grâce au développement des sources d'observations (en particulier les satellites) qui permettent de mieux contraindre le système, à une meilleure compréhension et représentation des processus dynamiques et physiques dans les modèles, et à l'augmentation des ressources informatiques qui permet d'améliorer la finesse des grilles de calcul. Néanmoins, ces prévisions restent largement imparfaites pour deux raisons principales :

- il est impossible de connaître parfaitement l'état initial en chaque point de l'atmosphère ;
- les modèles ne sont qu'une représentation approchée du comportement réel de l'atmosphère.

Les travaux de Lorenz (1963) montrent par ailleurs que l'atmosphère est un système chaotique, sensible à des perturbations très faibles de l'état initial : deux états initiaux très proches peuvent, au bout d'un certain temps (appelé horizon de prévisibilité), conduire à des solutions très différentes. Une prévision unique, dite déterministe, n'a donc de sens que si l'on considère l'incertitude qui lui est associée. Estimer ces incertitudes nécessite de recourir à une approche probabiliste de la prévision du temps (Leutbecher et Palmer, 2008).

Contrairement à la prévision déterministe qui fournit une estimation de l'état le plus probable de l'atmosphère, la prévision probabiliste fournit une estimation de la distribution de probabilité de l'état de l'atmosphère (notée pdf dans la suite). En pratique, cette distribution de probabilité est le plus souvent estimée à partir d'un ensemble de prévisions perturbées. Cette technique, connue sous le nom de *prévision d'ensemble*, est utilisée depuis les années 1990 (Bouttier et Buizza, 2018).

La prévision d'ensemble consiste à exécuter le modèle de prévision plusieurs fois, pour une même situation à prévoir, afin de tenir compte des différentes sources d'incertitude. Chaque prévision utilise notamment des conditions initiales très légèrement perturbées et éventuellement des configurations du modèle légèrement différentes (Figure 1.1). La qualité de la prévision d'ensemble est principalement déterminée par le choix des techniques utilisées pour échantillonner les incertitudes en jeu. Si les perturbations appliquées sont compatibles avec les

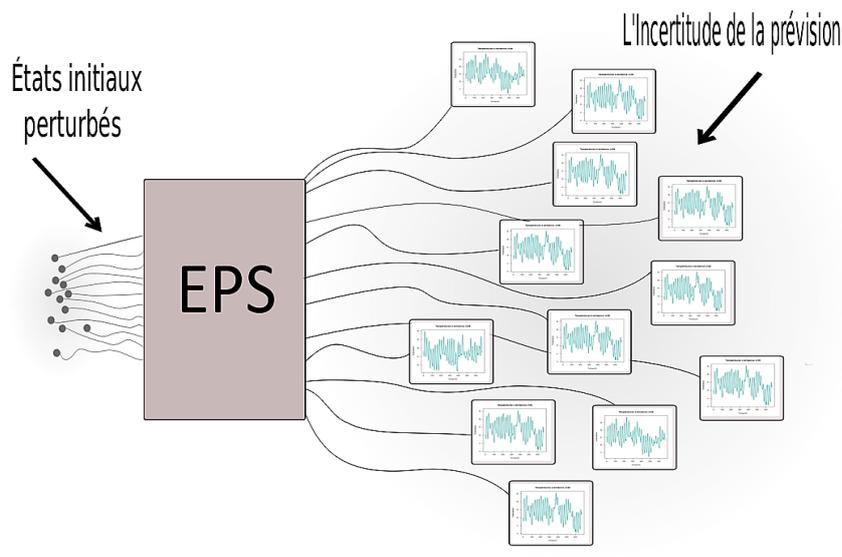


FIGURE 1.1 – Illustration du principe de la prévision d’ensemble (EPS-Ensemble Prediction System)). Une prévision différente (représentée ici sous la forme d’une série temporelle de température en un point donné) est générée à partir de chaque état initial perturbé (en pratique, d’autres sources d’incertitude, comme celle liée au modèle de prévision, peuvent être prises en compte mais ne sont pas représentées sur ce graphique).

erreurs présentes alors l’ensemble de prévisions est représentatif de l’incertitude de prévision. Le lecteur intéressé pourra se référer à (Buizza et Palmer, 1995 ; Raynaud et Bouttier, 2016 ; Leutbecher *et al.*, 2017) pour plus de détails sur les principales techniques de perturbation. Enfin, la taille de l’ensemble est également importante, plus l’ensemble est grand meilleure est l’estimation de la pdf.

Dans la suite du manuscrit chacune des prévisions d’un ensemble sera appelée "membre". Notons que par construction, chaque membre est une réalisation du modèle de prévision et présente donc une cohérence spatio-temporelle.

1.3 Calibration des prévisions d’ensemble

Les prévisions d’ensemble peuvent comporter des erreurs et des biais systématiques, qui peuvent avoir un impact sur la qualité de la prévision. Afin de réduire ces erreurs plusieurs méthodes dites de calibration, visant à recalibrer la distribution prévue sur la distribution observée, ont été proposées dans la littérature (Raftery *et al.*, 2005 ; Gneiting, 2014 ; Gneiting et Raftery, 2005 ; Taillardat *et al.*, 2019). Elles seront présentées plus en détail dans le chapitre 4.

1.4 Les modèles de PNT mobilisés

Deux types de modèle sont classiquement utilisés : les modèles globaux, qui couvrent tout le globe, et les modèles régionaux, qui couvrent un domaine géographique restreint (typiquement un pays). Les modèles régionaux sont également connus sous le nom de modèles à aire limitée (LAM-Limited Area Models). Outre leur couverture géographique ces modèles diffèrent également par leurs résolutions spatiales et leurs échéances. Les modèles régionaux utilisent généralement des maillages plus fins et des échéances plus courtes que les modèles globaux et permettent ainsi de simuler plus précisément les phénomènes de petite échelle. On détaille dans cette partie les modèles utilisés dans la suite du travail.

1.4.1 Modèles de prévision déterministes

Météo-France développe et utilise le modèle global ARPÈGE¹ (Courtier *et al.*, 1991) et le modèle régional AROME² (Seity *et al.*, 2011) pour ses prévisions opérationnelles. Par ailleurs, le modèle global IFS³ du CEPMMT⁴ est également utilisé par les services de production de Météo-France. Les caractéristiques de ces modèles (sur la période d'étude considérée dans la thèse), sont données dans le tableau 1.1.

1.4.2 Systèmes de prévision d'ensemble

En complément de leurs versions déterministes, les modèles ARPEGE, AROME et IFS disposent aussi de leurs versions ensemblistes, suffixées dans la suite EPS (Ensemble Prediction System). Nous utiliserons dans ce travail les systèmes ARPEGE-EPS (Descamps *et al.*, 2015), AROME-EPS (Bouttier *et al.*, 2012, 2016) et IFS-EPS (Molteni *et al.*, 1996 ; Buizza *et al.*, 2000 ; Palmer, 2019), dont les caractéristiques sont données dans le tableau 1.2. Pour des raisons de coût de calcul ces systèmes de prévision d'ensemble utilisent des résolutions horizontales moins fines que les modèles déterministes.

1. Action de Recherche Petite Échelle Grande Échelle
2. Application de la Recherche à l'Opérationnel à Méso-Échelle
3. Integrated Forecast System
4. Centre Européen pour les Prévisions Météorologiques à Moyen Terme

Nom	Caractéristiques
IFS	Zone : Globe ; Résolution horizontale : 9 km Résolution verticale : 137 niveaux ; Prévisions journalières : à 00, 12 UTC ; Echéances : 240h (J+10) ; Résolution de la grille des sorties (lat/lon régulier) : $0.5^{\circ} \times 0.5^{\circ}$
ARPèGE	Zone : Globe ; Résolution horizontale variable : 7,5 km sur la France et 37km aux antipodes Résolution verticale : 105 niveaux ; Prévisions journalières : à 00, 06, 12 et 18 UTC ; Echéances : 102h, 72h, 114h et 60h. Résolution de la grille des sorties (lat/lon régulier) : $0.1^{\circ} \times 0.1^{\circ}$
AROME	Zone : Europe de l'Ouest, centrée sur la France (domaine $\sim 1000 \times 1000$ km) ; Résolution horizontale : 1,3 km ; Résolution verticale : 90 niveaux ; Prévisions journalières : à 00, 03, 06, 12 et 18 UTC ; Echéances : 48h, 45h, 42h, 48h, 42h. Conditions aux limites latérales : prévisions ARPèGE. Résolution de la grille des sorties (lat/lon régulier) : $0.025^{\circ} \times 0.025^{\circ}$

TABLE 1.1 – Modèles de prévision déterministe utilisés de façon opérationnelle à Météo-France sur la période mars 2018 - juin 2019 considérée dans la suite du travail.

Nom	Caractéristiques
IFS-EPS	Zone : Globe ; Taille : 51 prévisions (dont une non perturbée) ; Résolution horizontale : 18 km (jusqu'au échéances J+15) et 36 km (au delà) ; Résolution verticale : 91 niveaux ; Prévisions journalières : 00 et 12 UTC ; Échéances : 360h (J+15) ; Perturbations : état initial, modèle (Leutbecher et Palmer, 2008). Résolution de la grille des sorties (lat/lon régulier) : $0.5^{\circ} \times 0.5^{\circ}$
ARPèGE-EPS	Zone : Globe ; Taille : 35 prévisions (dont une non perturbée) ; Résolution horizontale variable : 10km sur la France et 60km aux antipodes ; Résolution verticale : 90 niveaux ; Prévisions journalières : 00, 06, 12 et 18 UTC ; Échéances : 48h, 90h, 48h et 108h ; Perturbations : état initial, modèle (Descamps <i>et al.</i> , 2015) ; Résolution de la grille des sorties (lat/lon régulier) : $0.1^{\circ} \times 0.1^{\circ}$
AROME-EPS	Zone : Europe de l'Ouest, centrée sur la France (domaine $\sim 1000 \times 1000$ km) ; Taille : 12 prévisions ; Résolution horizontale : 2,5 km ; Résolution verticale : 90 niveaux ; Prévisions journalières : 03, 09, 15 et 21 UTC ; Échéances : 45h, 51h, 45h et 51h ; Perturbations : état initial, modèle, conditions de surface (Bouttier <i>et al.</i> , 2016), couplages latéraux (Bouttier et Raynaud, 2018) ; Résolution de la grille des sorties (lat/lon régulier) : $0.025^{\circ} \times 0.025^{\circ}$

TABLE 1.2 – Systèmes de prévision d'ensemble utilisés de façon opérationnelle à Météo-France sur la période mars 2018 - juin 2019 considérée dans la suite du travail.

Outils d'aide à la décision (OAD)

Sommaire

2.1	Principe des OAD	13
2.2	Les cas d'études des OAD utilisés	14
2.2.1	Le ver de la grappe	14
2.2.2	La septoriose du blé	16
2.3	Simulations agronomiques d'ensemble	17

Le but du chapitre précédent était de sensibiliser le lecteur aux concepts de prévision météorologique et à ses incertitudes ainsi qu'au concept de prévision d'ensemble pour mieux représenter cette incertitude. Afin de mieux illustrer le potentiel de la prévision d'ensemble dans le domaine de la protection des cultures, deux cas d'étude agronomiques ont été retenus. Ils seront détaillés dans ce chapitre ainsi que la mise en œuvre des simulations agronomiques d'ensemble.

2.1 Principe des OAD

Les OAD sont des outils apportant une information à l'utilisateur pour répondre à une problématique précise. Dans le domaine agricole et notamment ceux utilisés pour le raisonnement de la protection des cultures, ils sont notamment dépendant des conditions environnementales dont les conditions météorologiques. Ils s'appuient sur la modélisation de l'agro-écosystème afin de simuler l'évolution d'un ravageur ou d'une maladie particulière en s'appuyant sur une modélisation "mécaniste", prenant en compte les processus du système, ou statistique, mettant en relation directement les entrées et les sorties du système (Norton et Mumford, 1993). En s'appuyant sur cette modélisation, ils permettent de proposer des prévisions pour anticiper les situations à risque de quelques jours à une vingtaine de jours à l'avance. Ils s'appuient également sur des règles de décision qui permettent de transformer l'information sur la dynamique épidémique en préconisation de décision, considérée optimale pour la lutte contre le bioagresseur considéré avec des moyens bien définis. Dans le contexte de la protection des cultures, ces décisions sont souvent traduites en dates de traitement proposées pour appliquer des traitements phytosanitaires. Les OAD sont complémentaires des observations de terrain dans la mesure où ils permettent non seulement d'anticiper la situation mais également d'accéder à des variables difficiles à observer sur le terrain. Ainsi, dans le cadre d'une

réduction des usages des produits phytosanitaires, une partie de ces OAD sont utilisés dans le but d'aider les agriculteurs dans leur prise à décision pour positionner au meilleur moment les traitements nécessaires et ainsi permettre d'éviter des traitements superflus (Rossi *et al.*, 2012). Par ailleurs, les nouvelles solutions de lutte de type biocontrôle présentent des niveaux d'efficacité moindres, qui nécessitent en contrepartie plus de technicité et de précision dans leur positionnement (Shipp et Clarke, 1999) en rapport avec la dynamique épidémique.

2.2 Les cas d'études des OAD utilisés

L'Acta, le réseau des instituts techniques agricoles, développe et diffuse des OAD pour accompagner les agriculteurs et leurs conseillers dans leur raisonnement de la protection des cultures. Différents cas d'étude ont été identifiés parmi des OAD utilisés en routine sur le terrain par les agriculteurs et conseillers, en viticulture et en grandes cultures. Ces OAD, qui reposent sur des formalismes de modélisation variés, s'appuient sur différentes entrées météorologiques pour produire une prédiction. Pour l'étude, nous avons finalement retenu deux OAD : EVA et Septolis. **EVA** modélise la dynamique du ver de la grappe en vigne et utilise la **température** horaire comme seule entrée météorologique (Chavent, 1983). Ce premier cas d'usage nous a permis de proposer une démarche complète et des méthodes pour valoriser au mieux les prévisions d'ensemble. **Septolis** modélise la dynamique de la septoriose du blé et nécessite une série temporelle multivariée au pas de temps quotidien avec à la fois la **température** et les **précipitations** (Gouache *et al.*, 2013). La pluie étant une variable majeure pour la plupart des maladies, mais également pour d'autres thématiques agricoles comme la gestion de l'irrigation, il était important d'étudier l'intérêt potentiel d'information sur les incertitudes concernant ce paramètre, qui reste par ailleurs plus difficile à prédire par les PNT.

2.2.1 Le ver de la grappe

Les vers de la grappe sont des insectes ravageurs de la vigne. Deux espèces de lépidoptères co-existent dans les vignobles français : la Cochylys (*Eupoecilia ambiguella*) et l'Eudémis (*Lobesia botrana*). Toutes les deux pondent leurs œufs sur les inflorescences et les baies de raisin, dont les chenilles se nourrissent pour assurer leur développement, générant ainsi des dégâts quantitatifs et qualitatifs (Thiéry, 2008 ; Thiéry *et al.*, 2013).

Ces deux espèces présentent des cycles de développement très similaires, mais nous nous concentrons pour la suite sur la modélisation de Cochylys et de son paramétrage spécifique. Cet insecte se développe en quatre stades successifs au cours de son cycle : imago (l'adulte ailé), œuf, larve (avec 5 stades différents) et chrysalide (nymphe ou nymphe en diapause) (Figure 2.1). Au cours d'une année, il peut avoir plusieurs cycles de reproduction, généralement entre 3 et 5 par an en fonction des conditions météorologiques, mais les 2ème et 3ème cycles sont les plus critiques en terme de dégâts et sont ainsi au cœur de l'attention pour la lutte contre ce ravageur. Le nombre et le moment précis des cycles dépendent beaucoup des

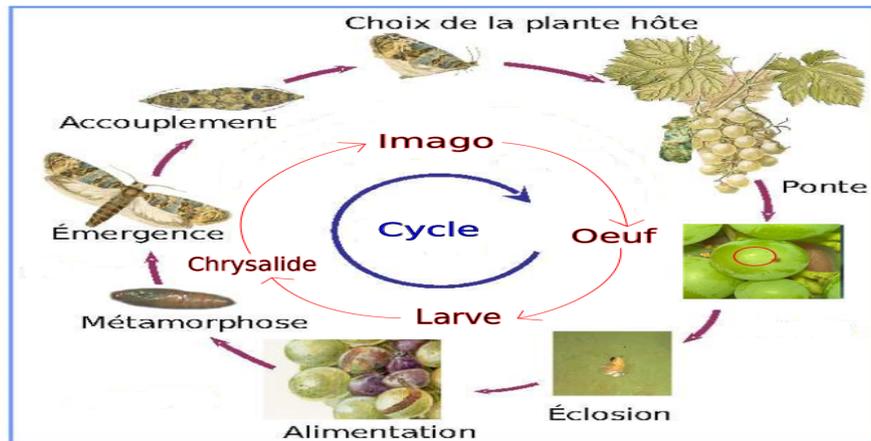


FIGURE 2.1 – Cycle de développement avec les différents stades du ver de la grappe, ravageur de la vigne. Source : (Maher *et al.*, 2006)

conditions météorologiques et demeurent difficiles à prévoir. Pour lutter contre les vers de la vigne, les viticulteurs disposent de trois types de traitement : des produits chimiques, des outils de lutte biologique (trichogrammes et toxines de *Bacillus thuringiensis* (Thiéry *et al.*, 2018)) et la confusion sexuelle. Ces produits s'utilisent à des moments précis, en accord avec le stade optimal de sensibilité du ravageur. Afin d'optimiser la synchronisation des traitements phytosanitaires avec ces stades, des modèles mathématiques sont utilisés pour estimer la dynamique de la population du ver de la grappe.

Il existe différents modèles qui décrivent la simulation du développement des ravageurs. Par exemple, les modèles dits modèles biologiques, qui prennent en compte l'environnement, les ressources alimentaires disponibles ainsi que la synchronisation entre la ponte des femelles et la maturité de la vigne (déterminante pour la survie des chenilles) ont été proposés (Picart, 2009) pour une synthèse). Cependant, pour ce cas d'usage, nous souhaitons nous appuyer sur des modèles prenant en compte les besoins thermiques uniquement pour simuler le cycle de développement du ravageur. Ainsi, nous avons choisi comme cas d'usage le modèle EVA décrit par Chavent (1983) et utilisé de manière opérationnel par l'Institut Français de la Vigne et du Vin (IFV).

EVA est un modèle phénologique qui tente de décrire de façon réaliste le cycle annuel de développement Cochylys/Eudémis. Le principe général du modèle EVA est le suivant. La population totale du ravageur est répartie en autant de compartiments que de stades de développement distincts : œuf, larve (de 1er stade au 5ème stade), chrysalide (nymphe et nymphe diapausante) ou imago. À partir du cumul de la température horaire, le modèle détermine pour chacun des stades une vitesse de développement. Ceci permet de calculer le développement accompli par les individus dans un stade (la proportion et le cumul de chaque stade au cours de temps). Lorsque celui-ci atteint une certaine valeur, une fraction de la population passe au stade suivant. Cette fraction, correspondant à une cohorte d'individus, évolue dans le stade suivant avec la vitesse de développement correspondante. Un facteur biotique, la date de nouaison, intervient dans le développement des larves et permet de moduler la vitesse de

développement. Les cas particuliers sont les stades L5 et imago. Pour le stade L5, la vitesse de développement est modulée par un coefficient de diapause. Selon la valeur de ce coefficient, les larves de 5ème stade évoluent soit vers le stade nymphe, soit vers le stade imago, il ne s’agit pas d’un changement de stade vers le stade œuf mais des pontes sont prises en compte. A partir d’un certain seuil de développement accompli, une décision peut être prise pour lancer un traitement ou non. Sur la base de discussions avec des experts (INRAE et IFV), deux seuils sont pertinents : 2% d’accumulation d’œufs pour un traitement chimique et 15% pour un traitement de biocontrôle (en raison de la faible persistance du produit).

Le modèle EVA développé à l’IFV sur la base du modèle de Chavent (1983) et intégré à la plateforme EPICURE (pour plus de détails, le lecteur est invité à se référer à l’annexe A.1). Le code du modèle EVA a été retranscrit pendant les travaux de la thèse depuis le langage PHP en langage R, en programmation orientée objet, afin de faciliter son utilisation pour de multiples simulations et faciliter le couplage avec les entrées météorologiques et les outils statistiques d’analyse et de combinaison des scénarios. Notez que la température horaire utilisée dans EVA est une sortie directe d’AROME-EPS, tandis que les sorties d’ARPEGE-EPS et d’IFS-EPS sont interpolées linéairement à une fréquence d’une heure.

2.2.2 La septoriose du blé

La septoriose tritiforme (STB) du blé d’hiver (*Triticum aestivum L.*) est une maladie majeure en Europe (Miedaner *et al.*, 2013 ; Kollers *et al.*, 2013 ; Suffert *et al.*, 2011 ; Ghaffary, 2011) et en France (Ghaffary *et al.*, 2011). La maladie a été décrite pour la première fois en France (Sprague, 1938 ; Desmazieres, 1842). Cette maladie foliaire du blé est causée par le champignon *Mycosphaerella graminicola* (Fuckel) Schröter (anamorphe *Septoria tritici*). Cette maladie est l’une des plus importantes au monde (Eyal, 1999), et la plus importante en France avec des pertes de rendement moyennes d’environ 1,5 tonne/ha sans traitement fongicide (Jørgensen *et al.*, 2008).

ARVALIS - Institut du végétal, a développé un modèle épidémiologique (Gouache *et al.*, 2013) pour le développement de la septoriose sur blé intégré dans un outil, Septo-LIS[®], pour optimiser le positionnement du premier traitement contre cette maladie (Gouache et Couleaud, 2009 ; Gouache, Couleaud *et al.*, 2009). Ce modèle combine un module qui prédit le stade d’apparition des feuilles et leur croissance et d’un module qui simule l’évolution de la maladie (Figure A.2), notamment les processus de contamination des feuilles basses vers les feuilles au dessus, la latence et l’extension des lésions. L’objectif de contrôle est de protéger la surface verte des deux dernières feuilles qui contribuent pour une grande part au rendement final. Ce premier traitement constitue donc un compromis entre l’attente de l’émergence de ces deux feuilles et un contrôle suffisamment précoce de la maladie. Ainsi, Septo-LIS[®] décrit de manière couplée la dynamique de la maladie et celle de la plante. La modélisation des processus de contamination entre les différents stades foliaires prend en compte notamment les précipitations quotidiennes, alors que les processus de latence et d’extension sont plus modélisés en fonction des températures. En conséquence, le modèle simule pour chaque étage foliaire le pourcentage de surface contaminé et le pourcentage de surface exprimant

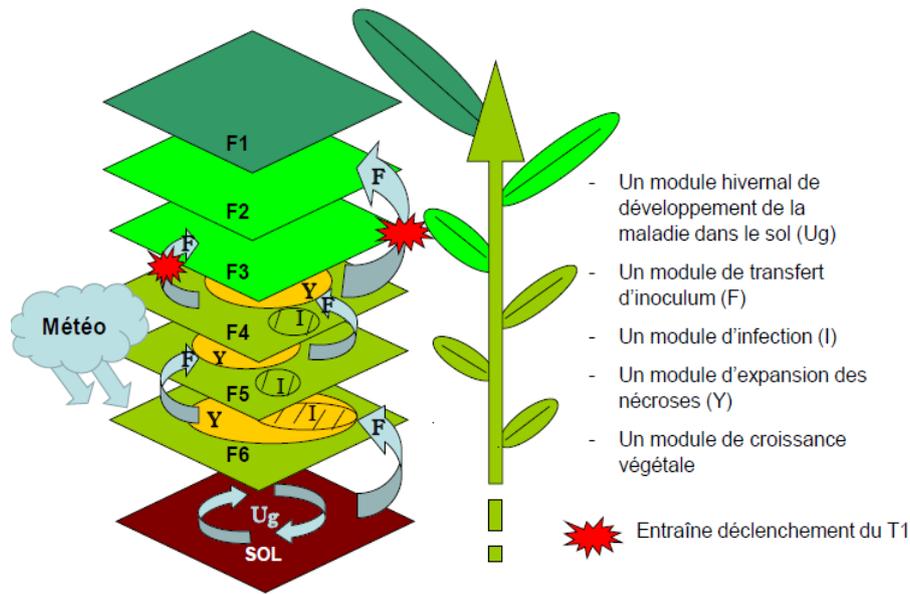


FIGURE 2.2 – Présentation du modèle SeptoLIS, schéma fonctionnel adapté Source : (Hervé, 2011)

les symptômes. Enfin, une règle de décision basée sur les contaminations des trois feuilles supérieures permet alors de calculer la date pour laquelle le premier traitement est considéré comme optimal. Septo-LIS[®] nécessite plusieurs entrées météorologiques au pas quotidien : les précipitations cumulées en 24 heures, les températures à 2 mètres minimale et maximale. Le modèle Septo-LIS[®] est également en opérationnel disponible pour les agriculteurs dans l'application Taméo. Pour plus de détails, le lecteur peut se référer dans Annexe A.2.

2.3 Simulations agronomiques d'ensemble

Les modèles agronomiques, comme les modèles météorologiques, sont des modèles dynamiques qui doivent être alimentés avec une condition initiale le jour de la prévision (le niveau de développement du ravageur ou de la maladie ce jour-là). Cependant, ces conditions initiales ne peuvent pas être mesurées et fournies au modèle comme dans le cas de la météorologie, mais sont plutôt calculées. De ce point de vue, le principe de la simulation agronomique est basé sur deux étapes : l'étape d'analyse et l'étape de la prévision.

L'étape d'analyse commence à une date fixe, appelée date d'initialisation du modèle. Au cours de cette étape, les modèles agronomiques sont exécutés à l'aide de séries chronologiques avec des données provenant d'observations jusqu'à la veille de la date de début de la prévision. Cette étape fournit les conditions initiales pour l'étape de prévision. L'étape de prévision prend en compte l'état initial de l'évolution du ravageur ou de la maladie et à partir de la date de début de la prévision elle considère des données de prévisions météorologiques.

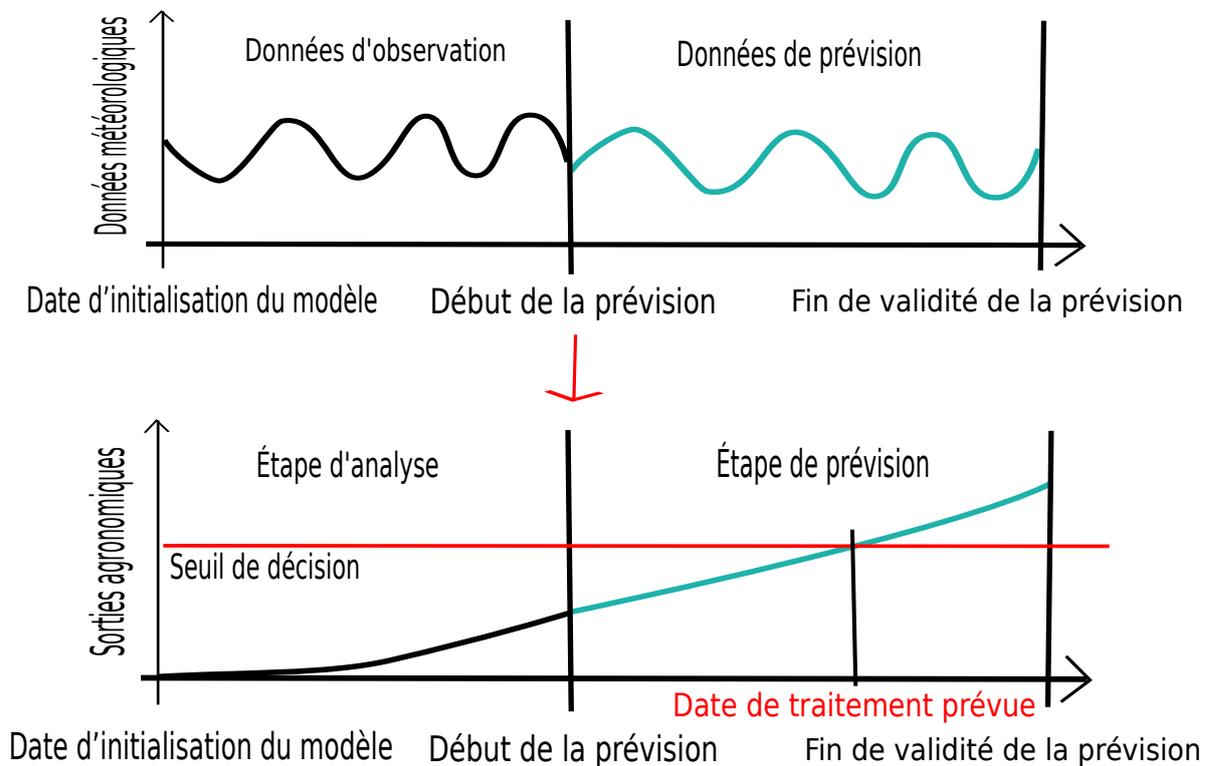


FIGURE 2.3 – Représentation schématique des deux étapes : analyse et prévision

Dans le but de familiariser le lecteur avec ces deux étapes, nous présentons un exemple schématique de simulation agronomique sur la Figure 2.3. Les sorties agronomiques sont obtenues grâce à la série chronologique des conditions météorologiques composée de la partie observation et de la partie prévision. En effet, il sera nécessaire de positionner l'étape de prévision au moment où le développement du ravageur ou de la maladie correspond à la période optimale de traitement. Nous dénommons cette période comme la "période d'intérêt" d'utilisation des prévisions d'ensemble.

La prise en compte de l'incertitude météorologique dans les simulations agronomiques se fait en propageant l'incertitude représentée par la prévision d'ensemble météorologique dans le modèle agronomique. Ainsi, pour la même date, nous obtenons les différentes dates de traitement issus des différents membres de l'ensemble météorologiques : ces différentes dates constituent un ensemble agronomique qui représente l'incertitude des dates prédites (Figure 2.4). Chaque membre de la prévision d'ensemble agronomique fournit alors une estimation de la date de traitement recommandée. Notons que l'ensemble de prévisions obtenu représente uniquement l'incertitude des prévisions météorologiques. D'autres sources d'incertitude, en particulier liées au modèle agronomique (valeur des paramètres, des variables d'entrée,...), pourraient être prises en compte, mais ne seront pas considérées dans la suite de ce travail.

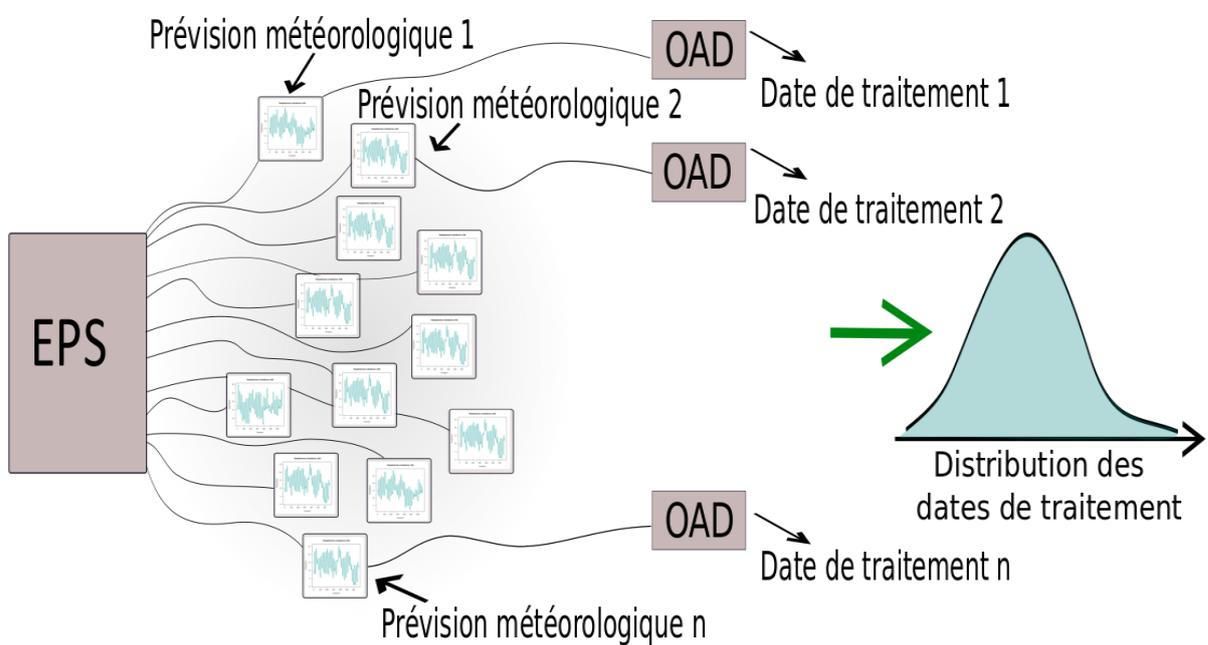


FIGURE 2.4 – Propagation des incertitudes des prévisions météorologiques. Un ensemble de simulations agronomiques est généré par propagation des prévisions météorologiques d'ensemble dans un modèle agronomique : en sortie, chaque membre agronomique utilise une prévision météorologique différente.

Objectifs et plan de la thèse

Objectifs de la thèse

Nous avons vu dans les chapitres précédents que les ingénieurs et conseillers agricoles ont à leur disposition des outils d'aide à la décision pour la gestion de maladies par des produits phytosanitaires basés sur des modèles de dynamique de maladies ayant comme variables d'entrée des entrées météorologiques. Par ailleurs, nous avons vu que les centres de recherche météorologique dont Météo-France développent des modèles de prévision qui fournissent une information riche sur les incertitudes associées aux prévisions météorologiques. L'objectif de la thèse est d'apporter des éléments de réponse à la question générale suivante : " Comment le monde agricole peut-il bénéficier des prévisions d'ensemble ? "

L'utilisation des prévisions d'ensemble dans les modèles de prévision agronomique implique en particulier de porter son attention sur la propagation des incertitudes météorologiques et son impact sur les dates de traitement prévues. L'usage actuel basé sur les données du passé permet de prendre en compte l'incertitude climatique inter-annuelle, mais cela ne reflète pas la situation de l'incertitude "du jour" qui peut être une cause d'erreur sur la précision de l'incertitude associée aux dates de traitement préconisées. Cela implique de se poser la question suivante Q1 : Quels seront l'apport et le potentiel de la prévision d'ensemble sur la quantification de l'incertitude associée à la prévision de la date d'un traitement phytosanitaire par rapport aux approches existantes ?

L'horizon d'une prévision agronomique dépend de l'usage que l'on souhaite en faire. En protection des cultures, la moyenne échéance (jusqu'à environ 2 semaines) est intéressante pour permettre une utilisation pratique dans l'anticipation de l'application des produits phytosanitaires sur les cultures. Cependant, la communauté agronomique a un réel intérêt pour les modèles météorologiques à haute résolution spatiale avec une précision à l'échelle de la parcelle qui pourront répondre au besoin d'ajustement des prévisions des dates de traitement. Sachant que les OAD ont besoin de séries temporelles météorologiques continues, de l'heure qui suit jusqu'à plusieurs jours d'échéance et que ces OAD sont souvent dédiés à la gestion d'une parcelle agricole, la question qui en découle est Q2 : Comment profiter de la richesse des différents modèles de prévision d'ensemble et comment améliorer leur exploitation au sein des OAD ?

Compte tenu du fait que les systèmes de prévision d'ensemble contiennent des erreurs

de biais systématiques notamment en raison d'une prévision des modèles météorologiques à l'échelle d'un pixel (liée à la résolution spatiale des modèles de PNT) alors que le besoin est à l'échelle d'une parcelle agricole, la première façon d'améliorer leur exploitation est d'appliquer des approches de correction statistique aux prévisions fournies par les différents modèles de prévision du temps. La deuxième voie d'amélioration consiste à tirer le meilleur parti de chacun des modèles de PNT et de faire la jointure entre les membres des modèles à haute résolution spatiale disponibles à court terme tout en assurant une continuité temporelle vers des informations à plus faible résolution spatiale à moyen terme en préservant leur cohérence et les caractéristiques météorologiques (le caractère dynamique et l'unicité de chaque membre). Cette jointure sera basée sur un critère de cohérence ; l'ensemble des prévisions obtenu sera nommé dans la suite prévision d'ensemble cohérente (« sans couture » en français, seamless en l'anglais). Cet aspect constitue la question de recherche méthodologique de cette thèse, qui jusque là a été peu abordée dans la littérature : "Comment faire la jointure entre des membres des modèles de PNT de plus haute résolution spatiale à courte échéance et les membres des modèles de PNT de plus faible résolution spatiale mais à plus longue échéance pour générer une prévision d'ensemble cohérente au-delà des échéances du modèle de la plus haute résolution spatiale ?

Enfin, pour évaluer le bénéfice de l'utilisation des données d'ensemble dans les OAD, nous nous posons la question Q3 "Quels seront la qualité et l'impact de notre prévision d'ensemble cohérente d'un point de vue météorologique et de son impact sur les prévisions des OAD ? "

Afin d'atteindre notre objectif général, les réponses aux questions Q1, Q2 et Q3 s'articulent en 3 étapes suivantes :

1. Illustrer et examiner le potentiel de la propagation de prévisions d'ensemble météorologiques dans les modèles agronomiques, dans le cas de la protection des cultures.
2. Appliquer une correction statistique aux prévisions météorologiques d'ensemble et proposer des stratégies de construction d'une prévision d'ensemble cohérente.
3. Évaluer des prévisions d'ensemble cohérentes d'un point de vue météorologique et agronomique.

Plan du manuscrit

Dans la partie II nous présentons les méthodes qui nous ont servi dans nos démarches.

Les méthodes de post-traitement par calibration statistique permettant de corriger les biais et la dispersion des prévisions d'ensemble seront présentées dans chapitre 4.

Le chapitre 5 sera dédié à apporter une réponse méthodologique à la construction des ensembles de membres cohérents. Nous discuterons le principe de la jointure, qui se traduit par un raccordement entre les membres. Nous introduirons également les mesures de distance et les critères de proximité utilisés pour déterminer les meilleures associations entre les membres.

Les prévisions d'ensemble cohérentes, ainsi que les simulations agronomiques associées,

seront soumises à des critères d'évaluation afin d'évaluer leurs performances. Les critères de vérification probabiliste seront présentés dans le chapitre 6.

Dans la partie III nous présentons tout d'abord dans le chapitre 7 le plan de simulation numérique mis en place.

Les résultats numériques associés aux étapes 1) à 3) seront présentés dans les chapitres 8, 9 et 10 respectivement.

Après avoir proposé quelques conclusions à nos travaux, nous les discuterons et apporterons des perspectives dans le chapitre 11.

Les articles soumis seront placés dans les annexes qui compléteront le manuscrit.

Deuxième partie

Méthodes

Plan de la partie II

Dans cette section, nous présentons toutes les méthodes que nous avons utilisées pour atteindre l'objectif de ce travail. Tout d'abord, dans le chapitre 4 nous rappelons les méthodes d'adaptation statistique existantes que nous avons utilisées afin d'améliorer les prévisions d'ensemble brutes. Dans le chapitre 5, la question de recherche principale sera traitée. L'objectif de ce chapitre sera de présenter les méthodes que nous proposons pour la construction d'un ensemble de prévisions "sans couture". Une fois construit, cet ensemble de prévisions sera utilisé pour les évaluations agronomiques. Il sera donc naturel de faire son évaluation, que ce soit d'un point de vue météorologique ou agronomique. Il convient donc d'utiliser des scores de vérification probabilistes. Ce sera l'objet du chapitre 6.

Calibration des prévisions météorologiques d'ensemble

Sommaire

4.1	Méthodes paramétriques	29
4.2	Méthodes non-paramétriques	31
4.3	Prévision d'ensemble calibrée	31

Les modèles de PNT présentent des erreurs et des biais systématiques, qui peuvent être réduits grâce à des méthodes de correction statistique (Hastie *et al.*, 2009; James *et al.*, 2013). L'approche la plus intuitive pour corriger les prévisions météorologiques consiste à établir une relation statistique entre les prévisions et les observations correspondantes. Une fois la relation établie, elle pourra être utilisée pour corriger les prévisions futures (Glahn et Lowry, 1972; Raftery *et al.*, 2005). Ces approches, aussi connues sous le nom d'adaptation statistique (notée AS dans la suite) ou calibration, sont utilisées au quotidien pour améliorer la qualité des prévisions opérationnelles (Ruth *et al.*, 2009). De nombreuses méthodes ont été proposées dans la littérature pour calibrer les prévisions d'ensemble, que l'on peut diviser en deux grands groupes : les méthodes paramétriques et non-paramétriques.

4.1 Méthodes paramétriques

Les méthodes les plus couramment utilisées dans le domaine de la météorologie sont des méthodes paramétriques qui nécessitent une hypothèse sur la distribution de la variable météorologique à calibrer, dont les paramètres dépendent des statistiques de l'ensemble brut. Ce groupe lui-même peut être divisé en deux sous-groupes (Gneiting, 2014) : méthodes dites NR (Non-homogenous Regression)/EMOS (Ensemble Model Output statistiques) (Gneiting et Raftery, 2005) et BMA (Bayesian Model Averaging) (Raftery *et al.*, 2005).

Notons y la variable météorologique d'intérêt et x_1, \dots, x_N les N prévisions correspondant aux membres de l'ensemble. La méthode EMOS s'intéresse à l'estimation de la distribution conditionnelle de y étant données les prévisions des différents membres :

$$Y \mid x_1, \dots, x_N \sim p(Y \mid x_1, \dots, x_N).$$

Par exemple, pour la température et la pression, il est classique d'utiliser une densité gaussienne, dont la moyenne et la variance sont des fonctions supposées affines de la moyenne et de la variance de l'ensemble brut, notées respectivement μ_e et σ_e^2 :

$$Y \mid x_1, \dots, x_N \sim \mathcal{N}(\alpha_0 + \alpha_1 \mu_e, \beta_0 + \beta_1 \sigma_e^2),$$

$$\mu_e = \frac{1}{N} \sum_{n=1}^N x_n,$$

$$\sigma_e^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - \mu_e)^2.$$

Les coefficients α_0 et α_1 permettent de corriger les biais systématiques de la moyenne d'ensemble, tandis que les coefficients β_0 et β_1 permettent d'adapter la dispersion de l'ensemble de manière à représenter correctement l'incertitude des prévisions. Ces coefficients sont estimés par régression à partir d'un échantillon historique de prévisions et des observations correspondantes sur une période d'apprentissage, en minimisant un critère de performance. Dans plusieurs études récentes, la maximisation de la fonction de vraisemblance (Aldrich *et al.*, 1997) tend à être remplacée par la minimisation du Continuous Ranked Probability Score (CRPS, défini dans la section 6.3.3, (Gneiting *et al.*, 2005 ; Gebetsberger *et al.*, 2018)), qui conduit généralement à de meilleurs résultats.

La BMA quant à elle suppose que la distribution conditionnelle du paramètre météorologique Y sachant les prévisions d'ensemble x_1, \dots, x_N suit une loi mixte de type combinaison linéaire de plusieurs densités paramétriques p_n dépendant de la prévision spécifique des membres d'ensemble x_n .

$$Y \mid x_1, \dots, x_N \sim \sum_{n=1}^N \omega_n p_n(Y \mid x_n) \quad (4.1)$$

où $\omega_n \geq 0$, $n = 1, \dots, N$ sont les poids associés aux membres n tels que $\sum_{n=1}^N \omega_n = 1$. Ces poids sont estimés par maximum de vraisemblance sur un échantillon d'apprentissage.

Les deux approches, NR/EMOS et BMA, sont les techniques les plus répandues pour corriger les erreurs des prévisions d'ensemble. Les coefficients inconnus dans les densités paramétriques sont à estimer à partir des prévisions passées et des observations correspondantes sur une période d'apprentissage. Dans le domaine des statistiques, les paramètres de distribution ont traditionnellement été estimés pour la plupart à l'aide de maximum de vraisemblance (Aldrich *et al.*, 1997). Bien que l'estimateur du maximum de vraisemblance ait certaines propriétés optimales, une estimation des paramètres NR en minimisant le score CRPS a été proposé par Gneiting *et al.* (2005). Depuis, cette approche d'estimation a été utilisée fréquemment dans les études de post-traitement pour les applications météorologiques avec des

performances généralement meilleures qu’avec l’estimation du maximum de vraisemblance (Gebetsberger *et al.*, 2018). Dans la suite de ce travail, la méthode EMOS sera privilégiée.

4.2 Méthodes non-paramétriques

Contrairement aux méthodes paramétriques, les méthodes non-paramétriques ne font pas d’hypothèses sur les distributions des variables météorologiques et offrent ainsi un cadre moins contraint pour la calibration. Meinshausen (2006) propose par exemple une méthode de post-traitement utilisant la technique des forêts aléatoires (Breiman, 2001). Taillardat (2017) montre que cette méthode améliore les prévisions de l’EMOS notamment pour la température à 2 mètres et le vent à 10 mètres. En revanche elle nécessite des périodes d’apprentissage plus longues. D’autres méthodes existent, comme l’ajustement des membres individuels de l’ensemble ou des quantiles de la distribution. Le lecteur intéressé pourra se référer à Wilks (2018).

4.3 Prévision d’ensemble calibrée

Le post-traitement de type EMOS fournit une distribution de probabilité calibrée continue, tandis que la prévision d’ensemble fournit une distribution de probabilité échantillonnée par les membres de l’ensemble. Disposer de membres est important pour la plupart des applications, qui utilisent en entrée les données de ces prévisions au lieu d’une distribution de probabilité. Il est donc courant d’ajouter en aval du post-traitement une étape d’échantillonnage de la distribution de probabilité calibrée, permettant de passer du continu au discret sous la forme de membres calibrés. Les deux méthodes les plus couramment utilisées effectuent un tirage aléatoire ou une sélection de quantiles équidistants (Wilks, 2015).

Par ailleurs, le post-traitement des prévisions d’ensemble est le plus souvent effectué de façon univariée, c’est-à-dire, indépendamment pour chaque variable, chaque lieu et chaque échéance. Au contraire, les prévisions d’ensemble issues d’un modèle de PNT sont des scénarios présentant une cohérence temporelle (entre échéances successives), spatiale (entre localisations voisines) et inter-variables. Ces cohérences étant importantes en prévision météorologique mais également pour la plupart des applications aval, il est possible de les reconstruire avec des méthodes de réarrangement, dont l’objectif est de reproduire une structure de dépendance donnée. On se limite dans la suite à la reconstruction de la cohérence temporelle, qui sera la seule considérée dans ce travail (les sites utilisés pour les simulations agronomiques sont considérés indépendants, et le cas d’étude utilisé pour l’évaluation de la calibration est celui du ver de la grappe, qui ne dépend que de la température). Il s’agit alors, pour chaque lieu, de réarranger les membres calibrés de manière à préserver la structure temporelle de référence.

Notons \tilde{x}_i^t , $i = 1, \dots, N$, $t = 1, \dots, T$ les N membres calibrés, en un lieu donné, pour les échéances 1 à T , et z_i^t , $i = 1, \dots, N$, $t = 1, \dots, T$ les données de référence correspondantes, présentant la structure temporelle à reproduire. Pour chaque échéance t :

- 1. les composants z_1^t, \dots, z_N^t sont ordonnés tels que $z_{(1)}^t \leq \dots \leq z_{(N)}^t$, où (i) indique le $i^{\text{ème}}$ membre tiré, et on en déduit les permutations $\pi_t(m) = \text{rank}(z_m^t)$
- 2. les membres $\tilde{x}_1^t, \dots, \tilde{x}_N^t$ sont ordonnés tels que $\tilde{x}_{(1)}^t \leq \dots \leq \tilde{x}_{(N)}^t$ et les membres sont réordonnés d'après les permutations de l'étape 1 : $\tilde{x}_{(\pi_t(1))}^t, \dots, \tilde{x}_{(\pi_t(N))}^t$.

Les deux méthodes de réarrangement les plus courantes sont l'ensemble copula coupling (ECC) (Scheffzik *et al.*, 2013) et le Schaake shuffle (Clark *et al.*, 2004). L'ECC reproduit la structure de dépendance de l'ensemble brut, tandis que le Schaake Shuffle reproduit la structure de dépendance d'un ensemble d'observations passées tirées aléatoirement. Dans le cas de l'ECC la taille de l'ensemble calibré est donc la même que celle de l'ensemble brut, alors que le Schakee Shuffle permet de générer des ensembles de n'importe quelle taille.

Pour illustrer cette étape de réarrangement prenons un exemple simple. Pour une échéance t fixée, considérons un ensemble de 3 membres dont les températures avant calibration sont $z_1 = 22^\circ C$, $z_2 = 21^\circ C$ et $z_3 = 20^\circ C$ et les températures après calibration sont $\tilde{x}_1 = 21.3^\circ C$, $\tilde{x}_2 = 22.1^\circ C$ et $\tilde{x}_3 = 21.7^\circ C$. Les valeurs brutes sont ordonnées de la façon suivante $z_3 \leq z_2 \leq z_1$, on en déduit les permutations : $\pi(1) = 3$, $\pi(2) = 2$ et $\pi(3) = 1$. Les valeurs calibrées sont ordonnées selon $\tilde{x}_1 \leq \tilde{x}_3 \leq \tilde{x}_2$, soit $\tilde{x}_{(1)} = \tilde{x}_1$, $\tilde{x}_{(2)} = \tilde{x}_3$, $\tilde{x}_{(3)} = \tilde{x}_2$. En appliquant les permutations, les membres sont ré-ordonnés tels que $\tilde{x}_1 = \tilde{x}_{(3)} = 22.1^\circ C$, $\tilde{x}_2 = \tilde{x}_{(2)} = 21.7^\circ C$, $\tilde{x}_3 = \tilde{x}_{(1)} = 21.3^\circ C$. Le réarrangement conduit donc bien à retrouver l'ordre décroissant des membres de l'ensemble brut.

On résume ainsi le processus de calibration des prévisions d'ensemble en trois étapes principales :

- application de la méthode EMOS aux prévisions d'ensemble brutes, indépendamment pour chaque lieu, échéance et variable ;
- échantillonnage de la distribution de probabilité calibrée par tirage aléatoire ou sélection de quantiles équidistants ;
- reconstruction des cohérences temporelles (éventuellement spatiales et multivariées) des membres calibrés par des méthodes de réarrangement.

Construction d'une prévision d'ensemble cohérente

Sommaire

5.1	Principe de la construction d'un ensemble cohérent	33
5.2	Méthodes de construction d'un ensemble cohérent	34
5.2.1	La méthode du plus proche voisin - NN ("Nearest Neighbor")	34
5.2.2	La méthode hongroise - HU ("HUngharian method")	35
5.3	La distance "Déformation Temporelle Dynamique"	35

Les simulations agronomiques à moyen terme sont particulièrement intéressantes pour les agriculteurs afin de prendre des décisions suffisamment à l'avance pour la protection des cultures (Calanca *et al.*, 2011). L'utilisation des systèmes AROME-EPS et ARPEGE-EPS, avec des échéances respectives de 2 et 4 jours, est bien adaptée pour préciser la prévision à courte échéance, mais pas pour une planification anticipée des traitements, pour laquelle le système IFS-EPS sera préféré. Dans la perspective de tirer le meilleur parti des différentes prévisions météorologiques disponibles, on propose dans cette partie une méthode simple pour combiner les prévisions de ces trois systèmes sous la forme de séries temporelles "sans couture".

5.1 Principe de la construction d'un ensemble cohérent

L'objectif de l'approche présentée dans cette section est de construire un ensemble de prévisions utilisant l'information haute-résolution des systèmes AROME-EPS et ARPEGE-EPS à courte échéance, puis l'information plus basse résolution des membres IFS-EPS à plus longue échéance. Cette problématique étant peu discutée dans la littérature, on choisit de la traiter dans la suite comme un simple problème de raccordement entre prévisions, en utilisant l'approche suivante :

- l'ensemble est composé des 12 membres AROME-EPS uniquement sur les échéances 0-48h ;
- au-delà, chaque membre AROME-EPS est raccordé à un membre ARPEGE-EPS ou IFS-EPS selon certains critères.

Cette configuration est motivée par les meilleures performances d'AROME-EPS sur ses échéances que nous documenterons dans le chapitre 9. Notons que la taille de l'ensemble ainsi

construit est limitée par la taille d'AROME-EPS (12 membres). Par ailleurs, cette procédure d'association est appliquée une ou deux fois, en fonction des membres sélectionnés à 48h (la fin de validité des prévisions d'AROME-EPS). Si un membre AROME-EPS est connecté à un membre IFS-EPS à l'échéance $t_{\text{merge}} = 48\text{h}$, le raccordement n'est appliqué qu'une seule fois pour ce membre. S'il est connecté à un membre ARPEGE-EPS, la procédure est répétée à l'échéance $t_{\text{merge}} = 96\text{h}$ afin de connecter ce membre ARPEGE-EPS à un membre IFS-EPS. On note dans la suite "ensemble HR" (Haute Résolution) l'ensemble des membres à raccorder et "ensemble BR" (Basse Résolution) l'ensemble des membres candidats au raccordement.

Seule la contrainte de cohérence temporelle des prévisions "sans couture" est considérée dans ce travail. Dans ce cadre simplifié, la procédure de raccordement décrite précédemment se résume donc à une comparaison de séries temporelles. Deux règles de raccordement de membres basées sur des mesures de distances entre des séries temporelles sont proposées et évaluées dans ce travail, que l'on détaille ci-après.

5.2 Méthodes de construction d'un ensemble cohérent

Tout d'abord nous considérons l'approche tirage aléatoire (notée RN - "Random Neighbor"), qui consiste simplement à sélectionner aléatoirement et sans répétition des membres de l'échantillon cible. Elle est utilisée notamment par Wetterhall et Di Giuseppe, 2018 pour une application hydrologique, et servira ici de référence pour l'évaluation de stratégies plus avancées.

Les autres méthodes proposées sont basées sur des mesures de distance, et consistent à sélectionner les membres cible de manière à minimiser une distance entre les membres haute et basse résolution. C'est une approche intuitive, qui permet d'assurer une transition douce entre deux prévisions.

Etant donnée une mesure de distance d , on note d_{ij} la distance entre le membre i de l'ensemble HR (de taille N) et le membre j de l'ensemble BR (de taille M). Deux règles de décision sont considérées, que nous détaillons ci-dessous.

5.2.1 La méthode du plus proche voisin - NN ("Nearest Neighbor")

Cette méthode attribue à chaque membre i de l'ensemble HR le membre j_i^* le plus proche de l'ensemble BR (Figure 5.1) :

$$j_i^* = \arg \min_j \{d_{ij}\}, i = 1, \dots, N, j = 1, \dots, M.$$

Cette méthode effectue donc une optimisation individuelle pour chaque membre de l'ensemble HR. Un inconvénient est que plusieurs membres de l'ensemble HR peuvent être reliés au même membre de l'ensemble BR, ce qui conduit à un ensemble de prévisions dont la taille effective

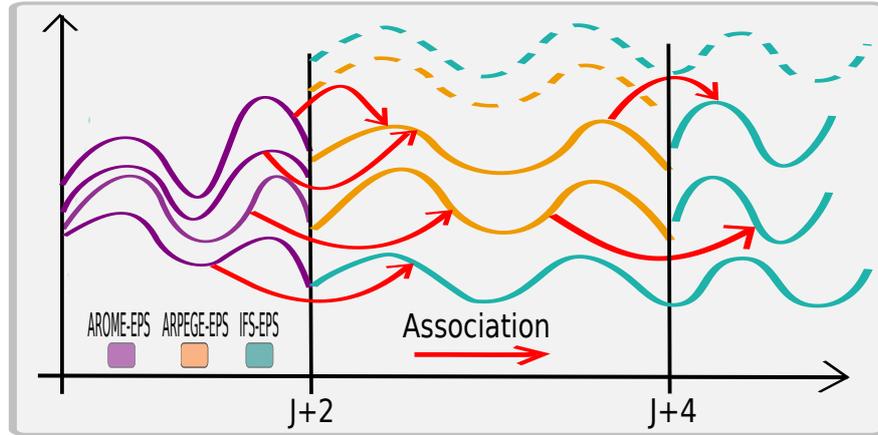


FIGURE 5.1 – Représentation schématique du raccordement selon l'approche "NN". Les trajectoires de la température AROME-EPS (les courbes violettes) sont couplées à J+2 avec les prévisions les plus proches d'ARPEGE-EPS (courbes toutes pleines oranges) ou d'IFS-EPS (courbes toutes pleines vertes). A J+4, seules les trajectoires ARPEGE-EPS sont couplées avec les trajectoires IFS-EPS disponibles. La flèche rouge indique les courbes qui sont choisies pour la jointure. Les courbes en pointillé représentent des membres non sélectionnés pendant le processus.

est inférieure à N .

5.2.2 La méthode hongroise - HU ("HUNGarian method")

La méthode hongroise utilise l'algorithme de Kuhn-Munkres (Kuhn, 1955) pour trouver la correspondance bijective optimale entre deux échantillons, en minimisant la distance totale de l'affectation :

$$j_i^* = \arg \min_j \left\{ \sum_{i=1}^N d_{ij} \right\}, j = 1, \dots, M.$$

Contrairement à l'approche du plus proche voisin, cette méthode sélectionne N membres BR différents (il n'y a donc pas de répétition de membres).

5.3 La distance "Déformation Temporelle Dynamique"

Les méthodes proposées sont génériques et peuvent s'appliquer à différentes variables météorologiques. Cependant, il est nécessaire de définir une mesure de distance appropriée, qui dépend notamment des caractéristiques de ces variables et de leur impact sur la simulation agronomique. Par exemple, dans le cas des modèles de dynamique épidémique qui nous intéressent dans ce travail, deux prévisions de température très proches mais présentant un léger décalage temporel auront le même effet sur le développement du ravageur. Il est donc naturel de choisir une distance qui tolère une certaine déformation temporelle.

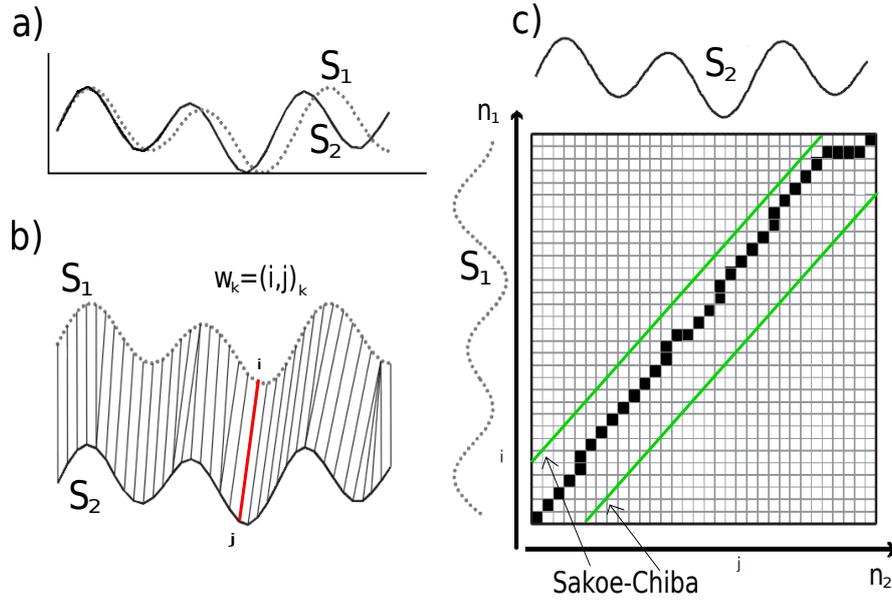


FIGURE 5.2 – a) Deux séries temporelles S_1 et S_2 , b) alignement des séries S_1 et S_2 , le point i de la série S_1 est associé au point j de la série S_2 , c) matrice de distance cumulée dans un voisinage (fenêtre de Sakoe-Chiba en vert) et chemin optimal en noir. Source : (Najmeddine *et al.*, 2012).

La distance Déformation Temporelle Dynamique (DTW - Dynamic Time Warping) initialement proposée par Berndt et Clifford (1994) répond au critère cité précédemment. La DTW mesure la similarité entre deux séries temporelles en déterminant, pour chaque élément d'une série, le meilleur élément correspondant dans l'autre série. Elle réalise ainsi un alignement global optimal en recherchant, parmi tous les alignements possibles, celui qui minimise les coûts d'association. La Figure 5.2 présente un exemple d'alignement effectué par la DTW dans le cas de deux sinusoides légèrement déphasées.

Considérons deux séries temporelles S_1 de taille n_1 et S_2 de taille n_2 :

$$S_1 = s_{1,1}, s_{1,2}, \dots, s_{1,i}, \dots, s_{1,n_1}, S_2 = s_{2,1}, s_{2,2}, \dots, s_{2,j}, \dots, s_{2,n_2} \quad (5.1)$$

et Ω un chemin de déformation qui aligne les éléments de S_1 et S_2 dans une séquence définie comme ceci :

$$\Omega = \omega_1, \omega_2, \dots, \omega_k, \dots, \omega_K \quad (5.2)$$

où $\omega_k = (i, j)_k$. La distance DTW entre deux points $s_{1,i}, s_{2,j}$ est définie par :

$$d_{\text{DTW}}(s_{1,i}, s_{2,j}) = \begin{cases} 0 & \text{si } i = j = 0 \\ \infty & \text{si } i = 0 \mid j = 0 \\ d_{\delta}(s_{1,i}, s_{2,j}) + \min\{d_{\text{DTW}}(s_{1,i-1}, s_{2,j-1}), \\ d_{\text{DTW}}(s_{1,i-1}, s_{2,j}), d_{\text{DTW}}(s_{1,i}, s_{2,j-1})\} & \text{sinon} \end{cases} \quad (5.3)$$

où $d_{\delta}(s_{1,i}, s_{2,j}) = s_{1,i} - s_{2,j}$ ou $d_{\delta}(s_{1,i}, s_{2,j}) = (s_{1,i} - s_{2,j})^2$. L'algorithme de la distance DTW détermine le chemin de déformation Ω entre les séries temporelles S_1 et S_2 tel que la distance cumulée soit minimale :

$$d_{\text{DTW}}(S_1, S_2) = \min_k \sum_{k=1}^K d_{\text{DTW}}(\omega_k) \quad (5.4)$$

où $d_{\text{DTW}}(\omega_k) = d_{\text{DTW}}(s_{1,i}, s_{2,j})$.

Par ailleurs, il est courant d'ajouter des contraintes pour le calcul du chemin de déformation. En particulier, la contrainte de fenêtre illustrée sur la Figure 5.2 (c) permet d'éviter les alignements pathologiques en limitant la recherche des points à un voisinage de forme et de taille données. Le voisinage proposé par Sakoe *et al.* (1990) est un des plus utilisés. Il est défini par la taille de fenêtre qui détermine le plus grand décalage temporel autorisé autour de la diagonale dans la direction de la plus longue série temporelle.

Pour ce travail, les prévisions d'ensemble cohérentes sont réalisées dans le cas de la température à 2 mètres en utilisant la distance DTW. Ainsi que le choix de distance, le type de la fenêtre de déformation et sa largeur, qui contrôle la limite du décalage temporel autorisé, sont des paramètres de réglage. Dans ce qui suit, la fenêtre Sakoe-Chiba est utilisée avec une fenêtre d'une largeur de décalage temporel d'une heure. L'algorithme de DTW est appliqué sur une période W avant le moment de la jointure t_{merge} . Ce critère de jointure est donc uniquement basé sur les informations antérieures au moment de l'affectation des membres. Les paramètres W et t_{merge} peuvent également être considérés comme des paramètres à régler. Dans notre travail, le choix retenu est $W = 7$ heures, suivant les performances obtenues avec différentes valeurs de W entre 1 et 48 heures. Enfin, le DTW est implémenté en utilisant l'algorithme R proposé par Giorgino *et al.* (2009).

Vérification de la prévision probabiliste - application aux prévisions d'ensemble météorologique et agronomique

Sommaire

6.1	Qu'est-ce qu'une bonne prévision ?	40
6.2	Vérification des prévisions probabilistes : approche orientée distribution	40
6.3	Vérification de prévision probabiliste : approche orientée mesures	42
6.3.1	Biais et erreur quadratique moyenne	42
6.3.2	Score de Brier (BS)	43
6.3.3	Continuous Ranked Probability Score (CRPS)	43
6.3.4	La courbe ROC (Receiver Operating Characteristic Curve)	43
6.3.5	Diagramme de fiabilité	44
6.3.6	La couverture des observations	45
6.3.7	IPR	45
6.4	Test des rangs signés de Wilcoxon	46

La prévision d'ensemble fournit une estimation de la distribution de probabilité, qu'il s'agisse, dans ce travail, de la distribution des conditions météorologiques ou de la distribution des dates prévues d'application des traitements phytosanitaires. Ce chapitre présente les outils de vérification de ces prévisions d'ensemble, classiquement utilisés pour en évaluer la performance. Contrairement à la vérification des prévisions déterministes qui s'appuie généralement sur des mesures simples telles que le biais et l'erreur quadratique moyenne, la vérification des prévisions d'ensemble requiert des mesures de performance spécifiques pour mesurer certaines propriétés des distributions de probabilité vis-à-vis d'une référence.

Par la suite, les notations suivantes seront employées :

- X la variable aléatoire qui désigne la cible recherchée (ex. la prévision de la température, ou la date de traitement préconisées par un OAD) et x une réalisation de la variable aléatoire qui désigne une prévision (prévision de la température le lendemain ou date préconisée de traitement) ;

— \mathbb{X}_o la référence vue comme une variable aléatoire. La réalisation de cette variable aléatoire, x_o représentera la référence observée

Dans la suite, nous noterons p les distribution marginales et conditionnelles, μ la moyenne, σ l'écart type de la variable aléatoire et $\text{supp}(\mathbb{X}) = \{x \in \mathbb{X} \mid p(x) \neq 0\}$ le support de la distribution.

6.1 Qu'est-ce qu'une bonne prévision ?

D'un point de vue très général, trois critères d'évaluation d'une prévision sont mentionnés dans les travaux de Murphy (1993). Il s'agit de la consistance (cohérence entre la prévision jugée la meilleure par le prévisionniste et la prévision diffusée), la valeur (le profit que l'utilisateur peut tirer de l'utilisation de la prévision) et la qualité (l'accord entre la prévision et la référence correspondante). Nous nous intéressons dans la suite à l'évaluation de la qualité des prévisions. Les principaux critères permettant d'évaluer la qualité d'un système de prévision probabiliste sont : précision, fiabilité, résolution, discrimination et acuité. Ils seront détaillés dans la section suivante.

6.2 Vérification des prévisions probabilistes : approche orientée distribution

La vérification d'une prévision probabiliste par une approche orientée distribution part du principe que toute l'information disponible se trouve dans la distribution conjointe $p(\mathbb{X}, \mathbb{X}_o)$ (Murphy et Winkler, 1987). La distribution conjointe du système de prévision et de référence, en utilisant la définition des probabilités conditionnelles, permet les factorisations suivantes :

$$p(\mathbb{X}, \mathbb{X}_o) = p(\mathbb{X} \mid \mathbb{X}_o) p(\mathbb{X}_o) \quad (6.1)$$

$$p(\mathbb{X}, \mathbb{X}_o) = p(\mathbb{X}_o \mid \mathbb{X}) p(\mathbb{X}) \quad (6.2)$$

Les deux factorisations mettent en évidence cinq distributions disponibles pour analyser la qualité d'un système de prévision : la distribution conjointe du système de prévision et de référence $p(\mathbb{X}, \mathbb{X}_o)$, la distribution conditionnelle du système de prévision sachant la référence (la fréquence à laquelle les différentes prévisions sont données avant qu'une valeur particulière soit observée) $p(\mathbb{X} \mid \mathbb{X}_o)$, la distribution conditionnelle de la référence sachant le système de prévision (le nombre de fois où une référence différente a été enregistrée lors d'une prévision spécifique) $p(\mathbb{X}_o \mid \mathbb{X})$, la distribution marginale du système de prévision (la probabilité des valeurs de prévision différentes) $p(\mathbb{X})$ et de la référence (la fréquence des différentes valeurs observées) $p(\mathbb{X}_o)$.

Précision Le critère de précision donne de l'information sur la correspondance entre la moyenne prévue par la prévision (ou le support de la prévision) et la référence ; Cela répond à la question : Est-ce que la référence est proche de la moyenne de la prévision d'ensemble ? Est-ce que la référence appartient à l'une des valeurs prévues par l'ensemble ? La réponse positive correspond à une prévision précise. Ce critère visera à déterminer : $\mu(\mathbb{X}) - x_o$, $\mu[(\mathbb{X} - x_o)^2]$ et $x_o \in \text{supp}(\mathbb{X})$.

Fiabilité La fiabilité indique l'adéquation entre les probabilités prévues et les fréquences observées de l'évènement, sachant la prévision. Cela représente le degré de correspondance entre la probabilité conditionnelle $p(\mathbb{X}_o | \mathbb{X})$ et \mathbb{X} pour toutes les valeurs de \mathbb{X} , c.a.d., lorsqu'une certaine prévision est émise avec 30% de confiance, il est préférable que l'évènement se produise 30 % du temps. La fiabilité est une mesure de l'efficacité de cette prévision dans la réalité. Cela répond à la question : Est-ce que la moyenne des températures observées sachant que nous avons prévu 20°C est égale à 20°C ? Plus elle est proche ou égale, plus le système est fiable. La fiabilité, pour une prévision probabiliste, est exprimée sous la forme suivante : $\mu(\mathbb{X}_o | \mathbb{X}) = \mathbb{X}$.

Résolution La résolution mesure la capacité d'un système de prévision à classer différentes références en fonction de sa prévision. Il s'agit donc de déterminer la différence entre $\mu(\mathbb{X}_o | \mathbb{X})$ et $\mu(\mathbb{X}_o)$ pour toutes les valeurs de \mathbb{X} . Cela répond à la question : Dans quelle mesure les moyennes conditionnelles de la référence correspondant à une prévision de température de 20°C et 30°C diffèrent l'une de l'autre et de la moyenne globale de la référence ? Plus la différence entre elles est grande, plus la résolution est bonne.

Discrimination Désigne la capacité d'un système de prévision à varier sa prévision selon la référence. Ce critère cherche à trouver le degré de correspondance des prévisions entre les évènements observés $\mu(\mathbb{X} | \mathbb{X}_o = 1)$ et non observés $\mu(\mathbb{X} | \mathbb{X}_o = 0)$. Cela répond à la question : Est-ce que le système de prévision fait la distinction entre les évènements et les non-évènements ? Idéalement, la distribution des prévisions dans les situations où l'évènement prévu se produit devrait être différente de la distribution correspondante dans les situations où il n'a pas lieu. Moins ces deux distributions coïncident, plus le système de prévision est discriminant.

Acuité Informe sur la variabilité des prévisions décrites par leur distribution. Il sera donc important de déterminer l'indicateur de dispersion de \mathbb{X} , $\sigma(\mathbb{X})$. Cela répond à la question : Quelle est la plage de valeurs prévue par l'ensemble ? Ce critère illustre la capacité du système à prévoir des probabilités extrêmes. Un système de prévision a une bonne acuité si les probabilités fournies par le système ont des valeurs extrêmes, c'est-à-dire proche de 0 ou de 1. Au contraire, un système produisant un grand nombre de probabilités proches de 0.5 a une mauvaise acuité.

Dans le tableau 6.1 nous présentons les particularités de ces distributions et les liens entre

Critère de qualité	Distributions utilisées	Nom du score
Précision	$p(\mathbb{X}), p(\mathbb{X}_o)$	Biais/MSE/ La couverture des observations
Fiabilité	$p(\mathbb{X}_o \mathbb{X}), p(\mathbb{X})$	BS, CRPS, Diagramme de fiabilité
Résolution	$p(\mathbb{X}_o \mathbb{X}), p(\mathbb{X}_o)$	BS, CRPS
Discrimination	$p(\mathbb{X} \mathbb{X}_o), p(\mathbb{X}_o)$	ROC/AUC
Acuité	$p(\mathbb{X})$	IPR

TABLE 6.1 – Relations entre les critères de qualité, la distribution et les scores

elles qui permettent d'étudier la qualité d'un système de prévision sous différents aspects en indiquant la ou les distributions qui contiennent l'information pertinente (Murphy, 1993).

6.3 Vérification de prévision probabiliste : approche orientée mesures

L'approche traditionnelle de la vérification s'appuie des mesures également appelées scores. Les règles de score sont des fonctions qui associent un scalaire à un couple prévision/référence (Stanski *et al.*, 1989). Ce scalaire est l'estimateur quantitatif de la qualité de la prévision. Dans la suite nous présentons les scores utilisés pour mesurer la qualité de la prévision probabilité.

6.3.1 Biais et erreur quadratique moyenne

Les scores déterministes biais et erreur quadratique moyenne (Mean Square Error - MSE) peuvent s'appliquer à la moyenne d'ensemble. Notons $\mu[\mathbb{X}_l(t)]$ la moyenne de l'ensemble pour l'échéance t et le lieu l :

$$\mu[\mathbb{X}_l(t)] = \frac{1}{N} \sum_{n=1}^N x_{n,l}(t), \quad (6.3)$$

où N est la taille de l'ensemble et $x_{n,l}(t)$ le membre n à l'échéance t et au lieu l . Le biais et la MSE sont définis par :

$$BIAIS(t) = \frac{1}{L} \sum_{l=1}^L (\mu[\mathbb{X}_l(t)] - x_{o,l}(t)) \quad (6.4)$$

$$MSE(t) = \frac{1}{L} \sum_{l=1}^L (\mu[\mathbb{X}_l(t)] - x_{o,l}(t))^2, \quad (6.5)$$

où $x_{o,l}(t)$ désigne la référence pour échéance t et lieu l et L le nombre de lieux considérés. La racine carrée de l'erreur quadratique moyenne est notée RMSE (Root Mean Square Error).

6.3.2 Score de Brier (BS)

Ce score, initialement proposé par Brier (1950), est utilisé pour la vérification d'une prévision probabiliste d'un évènement donné ϵ . Par exemple, on s'intéresse à l'occurrence d'un évènement tel que "la température est supérieure à un certain seuil s_ϵ ", noté comme $\epsilon := \mathbb{X} \geq s_\epsilon$. Notons $p_{s_\epsilon} = \mathbb{P}[\mathbb{X} \geq s_\epsilon]$ la probabilité prévue (c'est-à-dire le rapport entre le nombre de membres prévoyant l'évènement et le nombre total de membres, $p_{s_\epsilon} \in [0, 1]$) et $p_{o,s_\epsilon} = \mathbf{1}_{[\mathbb{X}_o \geq s_\epsilon]}$ la probabilité observée de l'évènement (p_{o,s_ϵ} vaut 1 si l'évènement est observé, 0 sinon). Le score de Brier mesure l'erreur quadratique moyenne entre la probabilité prévue et la probabilité observée :

$$BS_{s_\epsilon}(t) = \frac{1}{L} \sum_{l=1}^L (p_{s_\epsilon,l}(t) - p_{o,s_\epsilon,l}(t))^2. \quad (6.6)$$

6.3.3 Continuous Ranked Probability Score (CRPS)

Un score fréquemment employé dans la vérification des prévisions probabilistes d'une variable continue est le CRPS (Candille et Talagrand, 2005 ; Matheson et Winkler, 1976). Soit $\mathbb{F} = \mathbb{P}[\mathbb{X} \leq x]$ la fonction de répartition associée à une prévision d'ensemble et $\mathbb{F}_o = \mathbf{1}_{\{\mathbb{X}_o \leq x\}}$ la fonction de répartition associée à la référence x_o (qui est alors une fonction en escalier), le CRPS est défini par :

$$CRPS(\mathbb{F}, \mathbb{F}_o) = \int_{\mathbb{R}} (\mathbb{F}(x) - \mathbb{F}_o)^2 dx. \quad (6.7)$$

Une représentation visuelle de CRPS est donnée sur la Figure 6.1. Le score CRPS permet de mesurer la performance globale d'une prévision d'ensemble et peut être considéré comme l'intégrale du score de Brier sur tous les seuils possibles. Plus le score est petit, meilleure est la prévision.

6.3.4 La courbe ROC (Receiver Operating Characteristic Curve)

La courbe ROC (Hanley, 1982 ; Hanley et McNeil, 1982) évalue la capacité d'une prévision d'ensemble à discriminer un évènement donné (de type dépassement de seuil par exemple) pour différents seuils de décision s : si la probabilité prévue p_{s_ϵ} de l'occurrence de l'évènement est supérieure à s alors l'évènement est prévu, si le $p_{o,s_\epsilon} = 1$ alors l'évènement était observé.

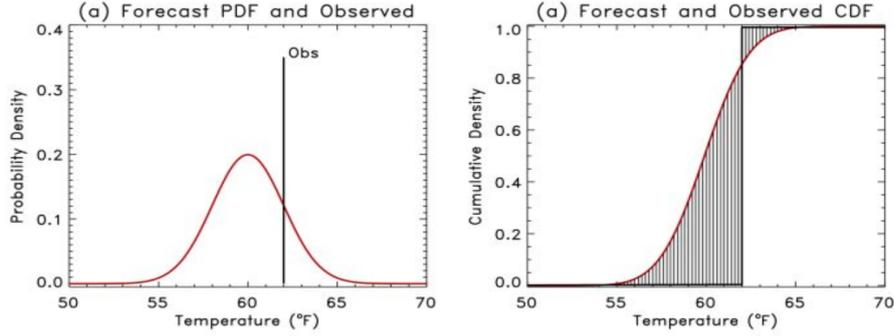


FIGURE 6.1 – À gauche la distribution de probabilité de la prévision (courbe rouge) et la référence (trait vertical noir), à droite les fonctions de répartition associées. Le CRPS correspond à l’aire entre les deux fonctions de répartition (surface grisée). L’axe horizontal représente la variable à prévoir. Source : Hersbach, 2000

Pour construire la courbe ROC, il faudrait donc définir deux statistiques qui résument la probabilité conditionnelle de la prévision sachant la référence pour le seuil s : le taux de détection $H(s)$ (la probabilité d’avoir bien prévu l’occurrence lorsque l’évènement a eu lieu) et le taux de fausses alarmes $F(s)$ (la probabilité d’avoir prévu l’occurrence alors que l’évènement n’a pas eu lieu).

$$H(s) = p(p_{s_\epsilon} \geq s \mid p_{o,s_\epsilon} = 1) , \quad (6.8)$$

$$F(s) = p(p_{s_\epsilon} \geq s \mid p_{o,s_\epsilon} = 0) , \quad (6.9)$$

La courbe ROC consiste à tracer le taux de détection (en ordonnée) et le taux de fausses alarmes (en abscisse) pour différents seuils de décision. Si la courbe ROC est alignée sur la première bissectrice, cela signifie que le système de prévision n’est pas discriminant (entre différentes situations de l’occurrence ou non d’un évènement). Une courbe proche du coin supérieur gauche (taux de détection = 1 et taux de fausses alarmes = 0) correspond à une prévision discriminante. Il est également possible de résumer cette information par l’aire sous la courbe ROC (notée AUC, Area under curve), comprise entre 0 et 1. Une AUC proche de 1 témoigne d’une prévision bien discriminante (Figure 6.2).

6.3.5 Diagramme de fiabilité

Soit p_{s_ϵ} la probabilité prévue et p_{o,s_ϵ} la probabilité observée de l’évènement. Les valeurs prévues sont réparties dans différents intervalles B_k , $k = 1, \dots, K$. Soit I_k un ensemble d’indices tel que $I_k = \{i : p_{s_\epsilon} \in B_k\}$. La fréquence relative observée pour l’intervalle k est définie par $\bar{o}_{k,s_\epsilon} = \frac{\sum_{i \in I_k} \mathbf{1}_{[p_{o,s_\epsilon, i} = 1]}}{|I_k|}$ et la moyenne des probabilités prévues par intervalle est définie par

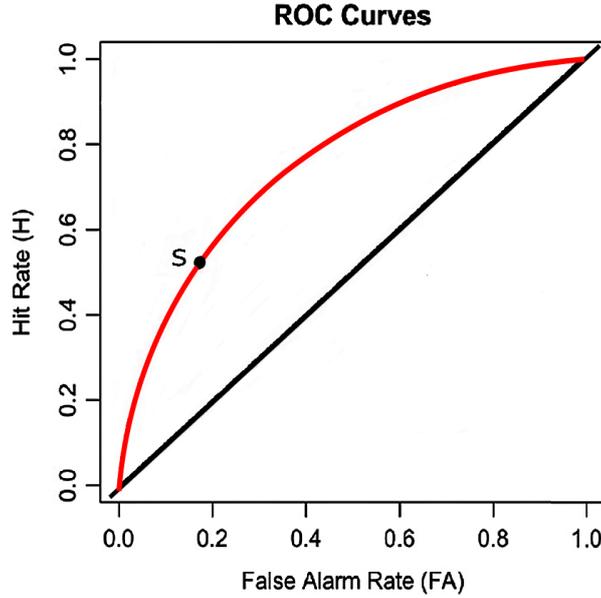


FIGURE 6.2 – Représentation schématique de la courbe ROC(en rouge) et le seuil s (point noir), par rapport à la bissectrice (en noir). En abscisse est représenté le taux de fausse alarme et en ordonnée le taux de détection.

$$\hat{p}_{k,s_\epsilon} = \frac{\sum_{i \in I_k} p_{s_\epsilon, i}}{|I_k|}, \text{ où } |I_k| \text{ est le nombre cardinal de l'ensemble } I_k.$$

Le diagramme de fiabilité représente la fréquence relative observée \bar{o}_{s_ϵ} (en ordonnée) et la moyenne de la probabilité prévue \hat{p}_{s_ϵ} (en abscisse) (Bröcker et Smith, 2007). Le système de prévision est fiable si l'évènement se produit avec une fréquence relative observée compatible avec la moyenne des valeurs prévues. Une courbe proche de la diagonale est le signe d'un système fiable. En revanche, un écart à la diagonale est le signe d'un biais.

6.3.6 La couverture des observations

La couverture des observations mesure le pourcentage d'observations qui se situent dans la fourchette des valeurs prévues par l'ensemble. Pour un ensemble de taille N , sous l'hypothèse de modèle parfait, la couverture des observations est égale à $\frac{N-1}{N+1} \times 100\%$.

6.3.7 IPR

Une façon de mesurer la dispersion de l'ensemble est de calculer les écarts entre les quantiles de la distribution prévue (notés IPR-Interpercentile Ranges). On utilisera principalement l'IPR90, qui représente la différence entre les 5ème et 95ème quantiles. Plus l'IPR90 est élevé, plus l'incertitude de la prévision est grande.

6.4 Test des rangs signés de Wilcoxon

Lorsque l'on compare deux prévisions il est important de déterminer si la différence de score observée est statistiquement significative. Pour cela nous utiliserons le test des rangs signés de Wilcoxon (Wilcoxon *et al.*, 1970). Il s'agit d'un test statistique non paramétrique qui compare deux échantillons. Il a comme objectif de déterminer si ces deux échantillons proviennent de la même distribution. Supposons que N est la taille des échantillons dans chaque groupe et que $\nu_{1,i}$, $\nu_{2,i}$ sont les échantillons. On fait les deux hypothèses suivantes :

- H_0 (nulle) : la différence entre les paires suit une distribution symétrique autour de zéro (c.a.d. les échantillons sont issus de la même distribution)
- H_1 (alternative) : la différence entre les paires ne suit pas une distribution symétrique autour de zéro

La statistique W sera ensuite considérée.

$$W = \sum_{i=1}^{N_r} [\text{sgn}(\nu_{2,i} - \nu_{1,i}) \cdot R_i]$$

où R_i est le rang des paires dans la série $|\nu_{2,i} - \nu_{1,i}|$, N_r est la taille de la séries des paires après avoir exclu les paires telles que $|\nu_{2,i} - \nu_{1,i}| = 0$. Sous l'hypothèse H_0 , et pour $N_r \geq 20$ la statistique W suit une distribution normale de moyenne 0 et de variance $\frac{N_r(N_r+1)(2N_r+1)}{6}$. On calcule ensuite la valeur p qui est la preuve contre une hypothèse nulle. Plus la valeur p est faible, plus la preuve de rejeter l'hypothèse nulle est forte. Par exemple pour un test unilatéral à gauche on calcule la p-value comme suit : $p_{\text{value}} = \mathbb{P}[W \leq w_0]$. Pour α un seuil critique donné, si $p_{\text{value}} < \alpha$ alors l'hypothèse nulle est rejetée ; on dira que la différence entre les deux échantillons est statistiquement significative. Pour ce travail, le test de significativité de Wilcoxon est effectué en utilisant le niveau $\alpha = 0,05$.

Troisième partie

Résultats

Plan de la partie III

Dans la partie II les méthodes existantes et celles que nous proposons ont été expliquées. La partie III sera dédiée aux résultats obtenus grâce à ces méthodes. Dans le chapitre 7 nous définirons les configurations expérimentales mises en œuvre sur les modèles agronomiques. Ensuite, dans le chapitre 8, nous montrons les simulations agronomiques d'ensemble obtenues avec les prévisions d'ensemble brutes. La comparaison avec l'approche fréquentielle permet ainsi de montrer le gain apporté par la prévision d'ensemble qui représente l'incertitude météorologique du jour. Les deux chapitres suivants présentent les résultats liés aux prévisions d'ensemble post-traitées, avec l'étape de calibration et celle d'élaboration de scénarios "sans couture". Dans le cadre de la thèse, ces ensembles post-traités ont été construits pour la température horaire à 2 mètres. Dans le chapitre 9 nous présentons les caractéristiques des scénarios cohérents construits, en termes notamment de continuité temporelle et d'indépendance des membres. Ensuite, dans le chapitre 10, les évaluations météorologiques de ces ensembles "sans couture" sont présentées avant de montrer l'impact de la calibration et des ensembles "sans couture" sur l'application agronomique avec le modèle du ver de la grappe en vigne (EVA).

Plan d'expérience numérique

Sommaire

7.1	Mise en œuvre des simulations agronomiques	51
7.2	Vérification probabiliste des prévisions météorologiques et agronomiques	55
7.3	Synthèse	56

Dans ce chapitre nous présentons tout d'abord la mise en œuvre des simulations agronomiques, et en particulier les données météorologiques (observations et prévisions) mobilisées pour cela. La procédure de vérification des prévisions météorologiques et agronomiques est ensuite présentée. Nous terminons par une synthèse des différentes données et simulations agronomiques disponibles utilisées pour la suite des travaux.

7.1 Mise en œuvre des simulations agronomiques

On rappelle que pour une année donnée, une simulation du modèle EVA (présenté dans le chapitre 2) commence toujours le 20 septembre de l'année précédente. Dans une première étape le modèle est intégré en utilisant les observations horaires de température de l'air à 2 mètres du réseau français d'observation météorologique en temps réel, à partir de cette date d'initialisation jusqu'à la date de début de la prévision. Au-delà, le modèle est intégré en utilisant les prévisions horaires de température à 2 mètres. La période d'intérêt pour les simulations EVA s'étend généralement sur la période d'avril à septembre, lorsque l'impact du ver de la grappe sur les baies de vigne est le plus important. Dans ce travail, les simulations avec EVA ont été effectuées pour chaque jour dans la période d'intérêt pour les années 2018 et 2019 sur 8 sites de vignobles français (Figure 7.1) qui représentent des zones où le ravageur est présent et peut causer des dégâts importants s'il n'est pas contrôlé.

Concernant la simulation avec Septo-LIS[®] (présenté dans le chapitre 2) pour une année donnée, le modèle est initialisé à la date de semis, fixé pour l'étude au 20 octobre de l'année précédant la récolte. De cette date d'initialisation à la date de début de la prévision, le modèle est intégré en utilisant les observations des températures (de l'air à 2 mètres) minimales et maximales, et les précipitations journalières. Ces données proviennent de stations météorologiques standards opérées par Météo-France ou Arvalis. Au-delà, les simulations Septo-LIS[®]

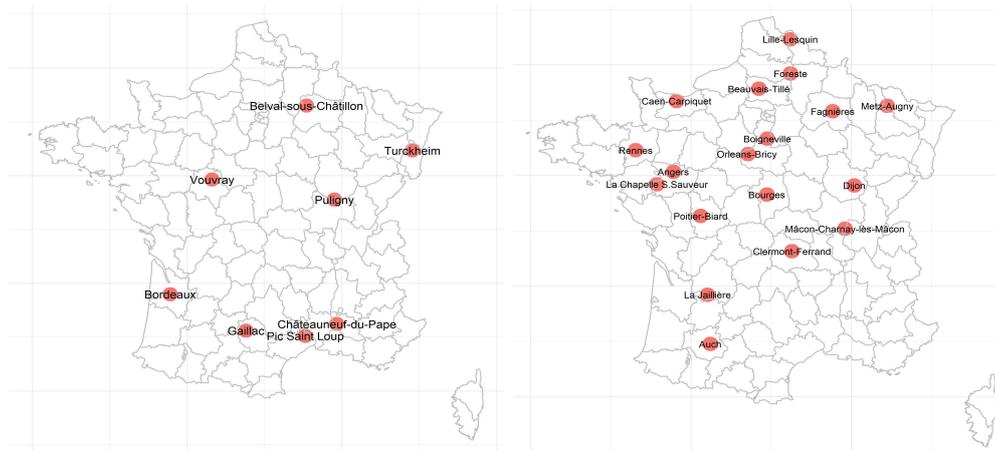


FIGURE 7.1 – Sites utilisés pour l'évaluation des prévisions agronomiques. Gauche : vignobles dans des zones présentant des problématiques de ver de la grappe. Droite : parcelles de blé dans des zones avec des problématiques de septoriose.

sont réalisées avec des prévisions météorologiques des températures minimales et maximales à 2 mètres et des précipitations journalières. Les prévisions Septo-LIS[®] sont effectuées entre le 10 et le 30 avril, période pendant laquelle les premiers traitements (T1) contre la septoriose sont généralement effectués, ce qui constitue la période d'intérêt pour les prévisions issues de Septo-LIS[®]. Les simulations avec ce modèle sont effectuées pour chaque jour dans la période d'intérêt sur les 18 parcelles du blé représentées sur la Figure 7.1, pour les années 2018 et 2019.

Les prévisions des variables météorologiques nécessaires aux simulations EVA et Septo-LIS[®] sont extraites de l'archive opérationnelle des modèles AROME, ARPEGE, IFS et des prévisions d'ensemble AROME-EPS, ARPEGE-EPS et IFS-EPS, sur la période d'intérêt de chacun des deux cas d'étude pour les années 2018 et 2019. Les prévisions au point de grille le plus proche des parcelles considérées sont extraites et utilisées pour l'étude. Lorsque les prévisions au pas de temps horaire ne sont pas directement disponibles dans l'archive, une interpolation temporelle linéaire est effectuée.

Pour les simulations agronomiques réalisées avec l'approche dite "fréquentielle", les prévisions sont données par les séries d'observations des 20 dernières années, de 1998 à 2018, extraites des archives de Météo-France et d'Arvalis selon les points spatiaux. Pour les simulations du modèle EVA, les séries de température (fournies par Météo-France) sont celles de la station la plus proche de la parcelle. En revanche, pour les simulations avec Septo-LIS[®], les données quotidiennes des températures minimales et maximales et des précipitations (fournies par Arvalis) sont obtenues par interpolation spatiale des données des stations les plus proches.

Ces points spatiaux particuliers ont été choisis : cela correspond à adopter le point de vue d'un agriculteur en privilégiant des prévisions ponctuelles associées à des parcelles particulières, sans prendre en compte une dimension spatiale. Par ailleurs, étant donnée la répartition

Nom de la parcelle	Latitude	Longitude	Culture
Belval-sous-Châtillon	49.128	3.856	Vigne
Bordeaux	44.748	-0.552	Vigne
Châteauneuf-du-Pape	44.057	4.831	Vigne
Gaillac	43.902	1.897	Vigne
Pic Saint Loup	43.780	3.813	Vigne
Puligny	46.945	4.753	Vigne
Turckheim	48.087	7.278	Vigne
Vouvray	47.412	0.799	Vigne
Angers	47.483	-0.600	Blé
Auch	43.690	0.600	Blé
Beauvais-Tillé	49.458	2.116	Blé
Boigneville	48.331	2.377	Blé
Bourges	47.066	2.366	Blé
Caen-Carpiquet	49.183	-0.450	Blé
Clermont-Ferrand	45.787	3.148	Blé
Dijon	47.275	5.092	Blé
Fagnières	48.950	4.416	Blé
Foreste	49.802	3.104	Blé
La Chapelle S.Sauveur	47.452	-0.966	Blé
La Jaillière	44.803	0.513	Blé
Lille-Lesquin	50.576	3.099	Blé
Mâcon-Charnay-lès-Mâcon	46.297	4.800	Blé
Metz-Augny	49.072	6.125	Blé
Orleans-Bricy	47.983	1.777	Blé
Poitiers-Biard	46.583	0.300	Blé
Rennes	48.070	-1.727	Blé

TABLE 7.1 – Les points spatiaux utilisés pour des simulations agronomiques sur les deux cultures de vigne et de blé. Les lignes en gris correspondent aux parcelles qui ont été choisies pour des évaluations météorologiques.

sur la France entière, il est également supposé qu’il n’y a pas ou peu de dépendance spatiale entre les points.

Un exemple de simulation agronomique avec le modèle EVA, présentant l’évolution temporelle du cumul du stade œuf du 20/09/2017 au 07/08/2018, est donné sur la Figure 7.2. Les étapes d’analyse et de prévision sont identifiées en noir et bleu respectivement. Le début de la prévision est ici positionné au moment crucial où débute le développement du stade œuf. A partir de cette dynamique temporelle du stade œuf, on en déduit une date de traitement correspondant au franchissement d’un seuil (fixé pour l’étude à 2% et 15% selon le type de traitement). Pour le modèle Septo-LIS[®] la règle de décision est directement intégrée dans le modèle, et nous ne disposons en sortie que de l’information concernant les dates de traitement prévues.

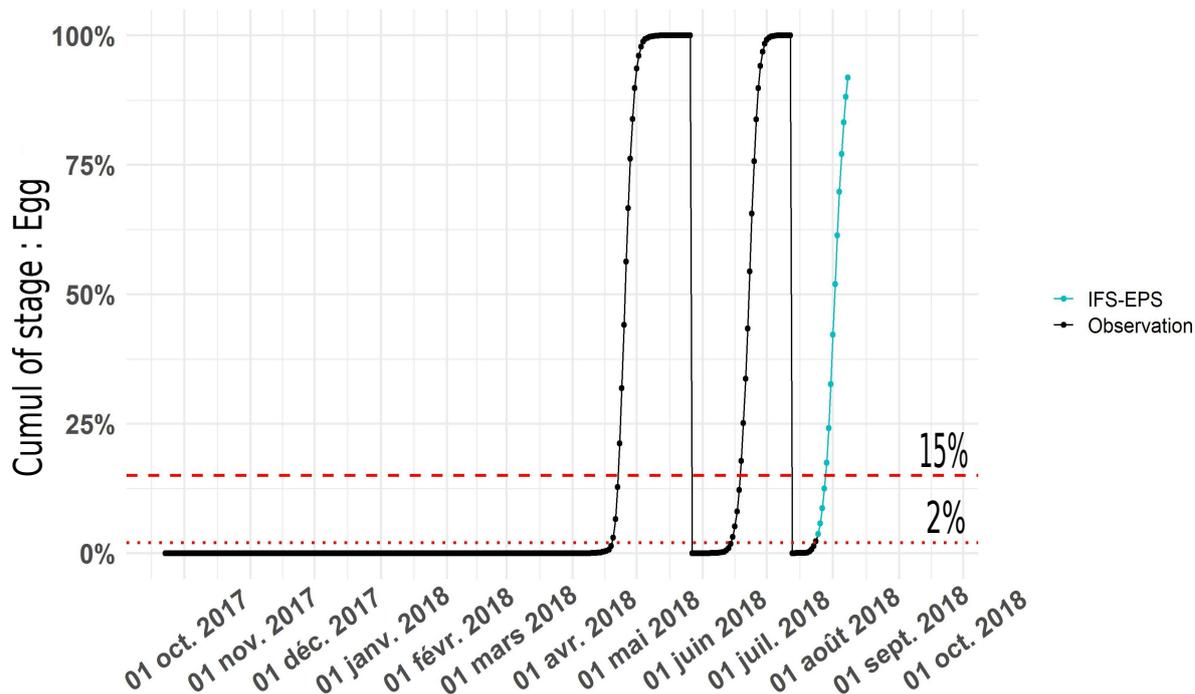


FIGURE 7.2 – Simulation avec EVA pour le site de Châteauneuf-du-Pape. La partie analyse (en noir) utilise les observations horaires de la température à 2 mètres entre la date d’initialisation (20/09/2017) et la veille de la date de prévision (24/07/2018). La partie prévision (en bleu) utilise les données de prévision de température à 2 mètres d’un membre de la prévision d’ensemble IFS-EPS du 24/07/2018 jusqu’au 07/08/2018 (échéance maximale des prévisions IFS-EPS). En ordonnée : le cumul du stade œuf. La date de traitement est obtenue en fonction du franchissement de seuil choisi, les seuils 2% et 15% sont représentés ici.

7.2 Vérification probabiliste des prévisions météorologiques et agronomiques

Les scores de vérification probabilistes présentés dans le chapitre 6 sont utilisés pour les évaluations météorologiques et agronomiques.

L'évaluation météorologique vise à évaluer l'impact de la calibration et des stratégies de construction des prévisions d'ensemble cohérentes. Cette évaluation est effectuée uniquement pour la température à 2 mètres sur la période du 1^{er} juin 2018 au 31 mai 2019, et sur les 11 parcelles présentées en gris dans le tableau 7.1. Ces sites échantillonnent une grande variété de conditions météorologiques dans différentes régions du pays pour la viticulture et la culture du blé tendre affecté par le ravageur du ver de la grappe et la maladie de la septoriose. Les 3 sites associés au modèle du blé représentent des parcelles sur lesquelles les premières études agronomiques et météorologiques ont été réalisées. Les 15 parcelles restantes ont été prises en compte uniquement pour les études agronomiques, pour des raisons d'accès décalé dans le temps et de coût des calculs liés à la calibration et à la construction de prévisions d'ensemble cohérentes. La référence considérée pour cette évaluation est l'observation de terrain fournie par les stations Météo-France les plus proches de ces sites (Jolliffe et Stephenson, 2012). Les observations météorologiques peuvent contenir des erreurs de mesure. Dans le cadre de ce travail, cette erreur n'est pas considérée.

En ce qui concerne la vérification agronomique, les scores sont appliqués à la distribution des dates de traitement prévues comparée à une date de référence. Cette évaluation sera mise en œuvre pour comparer les performances des prévisions d'ensemble et de l'approche fréquentielle, ainsi que pour évaluer l'impact des post-traitements des prévisions d'ensemble (calibration, jointure des ensembles). Les vérifications les plus détaillées sont effectuées dans le cas d'EVA. En agronomie et en particulier en protection des cultures, il est souvent difficile d'obtenir les observations de terrain régulières et sur un vaste territoire. En effet, ces observations nécessitent l'intervention d'experts qui doivent réaliser des observations régulières sur la base de protocoles harmonisés. Si sur nos deux cas d'étude (ver de la grappe et septoriose), il existe bien des données collectés dans le cadre de l'épidémiosurveillance (par exemple, pour le Bulletin de Santé du Végétal), il est compliqué, voire impossible de les faire correspondre aux simulations agronomiques disponibles. Ainsi, quand elles existent, ces données ne correspondent pas forcément aux sorties du modèle. En effet, ici, nous avons le modèle qui prédit une date préconisée de traitement, qui correspond à une date optimale pour assurer l'efficacité d'un traitement, très difficile à estimer à partir d'observation, alors que les observations portent sur des captures de papillon (stade adulte) pour le ver de la grappe ou de l'incidence de la septoriose sur les dernières feuilles. Aussi, nous avons choisi d'utiliser comme référence la date de traitement obtenue en intégrant le modèle agronomique uniquement à partir des observations météorologiques. Cette référence agronomique est possiblement entâchée d'erreur, en raison notamment des erreurs de modélisation agronomique (processus non pris ou mal pris en compte, erreurs sur des paramètres ou d'autres variables d'entrée que la météo), dont il faudrait idéalement tenir compte pour le calcul des scores et prendre l'ensemble des sources d'incertitude en compte. Dans ce travail, cette erreur sur les références est néanmoins

négligée.

Pour les deux cas d'étude agronomiques les prévisions réalisées ont une échéance maximale de 15 jours (correspondant à l'échéance des prévisions IFS-EPS). Si le seuil utilisé pour déterminer la date de traitement n'est pas atteint par certains membres en fin de prévision, le modèle fixe la date de traitement à "au-delà". Afin d'effectuer le calcul des scores, ces dates "au-delà" sont arbitrairement fixées à 17.

7.3 Synthèse

Le tableau 7.2 récapitule les données météorologiques utilisées et les sorties agronomiques associées qui seront utilisées dans les chapitres suivants. La date d'initialisation des simulations agronomiques et les paramètres des modèles EVA et Septo-LIS[®] sont fixés à des valeurs constantes communes à l'ensemble des simulations.

REMARQUE : L'extraction des données météorologiques observées et prédites a nécessité l'interrogation de plusieurs bases de données (internes à Météo-France, hébergée sur le centre européen ECMWF ou encore du système d'information d'Arvalis). La procédure étant spécifique à chaque archive, cela a nécessité le développement de plusieurs requêtes d'extraction. Le processus d'extraction des prévisions est par ailleurs coûteux en temps, quelle que soit la base de données, en particulier les prévisions d'ensemble de plusieurs membres. Les données de la prévision à plusieurs jours d'échéance ont été extraites pour chaque jour de la période d'intérêt et pour 2 ans, ce qui représente un grand volume de données. Cette partie a été une phase importante de ce travail.

Données météorologiques	Sorties agronomiques
Observations des stations météorologiques (\mathbb{T} , \mathbb{T}_{\min} , \mathbb{T}_{\max} , et \mathbb{RR})	Date de traitement de référence (EVA, Septo-LIS [®])
Prévisions déterministes brutes (\mathbb{T} , \mathbb{T}_{\min} , \mathbb{T}_{\max} , et \mathbb{RR})	Date de traitement (EVA, Septo-LIS [®])
Données fréquentielles, i.e. des séries d'observations des stations météorologiques (\mathbb{T} , \mathbb{T}_{\min} , \mathbb{T}_{\max} , et \mathbb{RR})	Distribution des dates de traitement (EVA, Septo-LIS [®])
Prévisions d'ensemble brutes (\mathbb{T} , \mathbb{T}_{\min} , \mathbb{T}_{\max} , et \mathbb{RR})	Distribution des dates de traitement (EVA, Septo-LIS [®])
Prévisions d'ensemble post-traitées (\mathbb{T})	Distribution des dates de traitement (EVA)
Prévisions d'ensemble cohérentes (\mathbb{T})	Distribution des dates de traitement (EVA)

TABLE 7.2 – Récapitulatif des données météorologiques et des sorties agronomiques associées, utilisées dans la suite du travail. Les données météorologiques sont : la température horaire de l'air à 2 mètres (\mathbb{T}), la température minimale (\mathbb{T}_{\min}), maximale (\mathbb{T}_{\max}), et le cumul des précipitations (\mathbb{RR}) au pas de temps journalier. Les prévisions déterministes proviennent des modèles AROME, ARPEGE et IFS, et les prévisions d'ensemble des modèles AROME-EPS, ARPEGE-EPS et IFS-EPS. Les observations sont extraites des archives de Météo-France (\mathbb{T}) et d'Arvalis (\mathbb{T}_{\min} , \mathbb{T}_{\max} , \mathbb{RR}). Pour les données de prévision, elles proviennent des bases de données de Météo-France (AROME, AROME-EPS, ARPEGE, ARPEGE-EPS) et du CEPMMT (IFS et IFS-EPS).

Apport de la prévision d'ensemble pour les simulations agronomiques

Sommaire

8.1	Analyse graphique des données de prévision d'ensemble et des données fréquentielles	60
8.2	Propagation des incertitudes météorologiques dans les OAD	63
8.2.1	Distribution des dates de traitement : Modèle EVA	63
8.2.2	Distribution des dates de traitement : Modèle Septo-LIS [®]	66
8.3	Évaluation agronomique de la prévision d'ensemble par rapport aux données fréquentielles	69
8.3.1	Application au modèle EVA	69
8.3.2	Application au modèle Septo-LIS [®]	72
8.4	Propagation des incertitudes des données météorologiques de PNT à haute résolution pour des simulations agronomiques	76
8.5	Conclusion	77

L'utilisation d'une prévision déterministe pour les simulations agronomiques fournit une date de traitement unique. Cette information peut dans certains cas s'avérer limitante pour la prise de décision dans un contexte incertain. Dans le cas des simulations agronomiques probabilistes, chaque membre de la prévision météorologique d'ensemble produira une date de traitement lors de son utilisation dans un OAD. L'ensemble de ces dates de traitement correspond à une estimation de la distribution des dates de traitement prévisionnelles. L'approche probabiliste apporte donc une information plus riche et potentiellement intéressante pour accompagner la prise de décision.

Comme expliqué dans les chapitres précédents, l'approche probabiliste la plus souvent mise en œuvre en simulation agronomique est l'approche dite "prévision fréquentielle". Cette approche utilise comme entrées météorologiques l'historique des séries observées des années passées (souvent sur 20 ans) sur les points géographiques d'intérêt. Par construction, pour un site, cette distribution reflète donc la variabilité inter-annuelle des paramètres météorologiques plutôt que réellement l'incertitude associée à la prévision liée aux conditions météorologiques d'un moment particulier comme pour les prévisions d'ensemble météorologiques.

Ce chapitre a pour objectif d'examiner et d'évaluer l'impact sur les simulations des modèles EVA et Septo-LIS[®] des prévisions d'ensemble météorologiques par rapport aux deux

pratiques existantes, à savoir la prévision déterministe et la prévision fréquentielle. Pour cela, les sorties des modèles agronomiques seront comparées entre elles ainsi qu'à la date de traitement de référence définie dans la section 7.2. Cette comparaison s'appuie sur l'examen de certaines caractéristiques des distributions et sur les critères de vérification définis dans le chapitre 6.

Ce chapitre s'appuie en grande partie sur l'article (Aleksavska *et al.*, 2018) (Annexe B) et Aleksavska *et al.*, 2020 soumis à *Crop Protection* (Annexe C).

8.1 Analyse graphique des données de prévision d'ensemble et des données fréquentielles

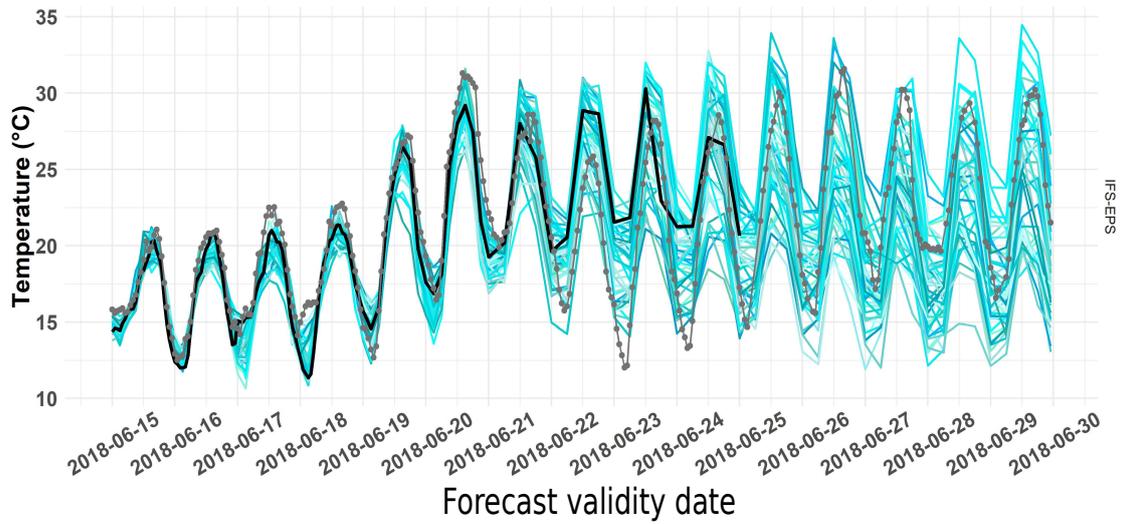
Dans une première étape, nous comparons visuellement les données météorologiques fréquentielles et celles issues de la prévision d'ensemble IFS-EPS afin d'illustrer les conséquences d'utilisation de ces différentes sources sur des cas particuliers. Nous nous intéressons en particulier à deux paramètres météorologiques, la température (horaire à 2 mètres pour le modèle EVA, minimale et maximale pour Septo-LIS[®]) et les précipitations quotidiennes au sol (pour le modèle Septo-LIS[®]).

La Figure 8.1 présente un exemple de séries temporelles de la température de l'air à 2 mètres à Bordeaux. Avec la prévision d'ensemble (Figure 8.1(a)) nous remarquons que la dispersion des membres, et donc l'incertitude des prévisions, augmente avec l'échéance. En revanche l'incertitude échantillonnée par les données fréquentielles (Figure 8.1(b)) est relativement constante dans le temps.

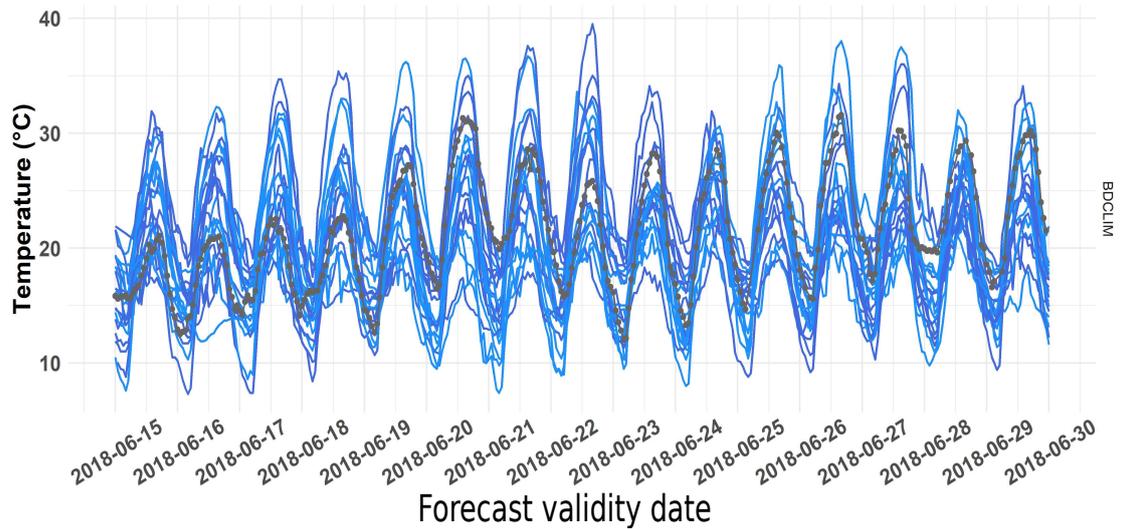
La Figure 8.2 présente un exemple d'une distribution des précipitations quotidiennes à Poitiers Biard, obtenue à partir de la prévision d'ensemble et résumée sous la forme de boîtes à moustaches (Figure 8.2(a)) et des données fréquentielles (Figure 8.2(b)). La distribution échantillonnée par la prévision d'ensemble dépend de la situation météorologique et de l'échéance, tandis que les variations représentées par les données fréquentielles sont constantes dans le temps, avec plus de valeur extrêmes observées.

Localement des différences assez notables entre ces deux sources d'information peuvent être observées. Par exemple, pour les premiers jours du 10 au 13 avril, la prévision d'ensemble prévoit des cumuls nuls et sans aucune variabilité, alors que les données fréquentielles présentent de la variabilité pour cette période. Pour la journée du 18/04/2019, la prévision d'ensemble prévoit des cumuls de pluie supérieurs à 5mm avec une incertitude assez grande, tandis que les données fréquentielles correspondantes prévoient des cumuls majoritairement inférieurs à 2mm.

Afin de généraliser ces conclusions les tableaux 8.1 et 8.2 présentent la moyenne et l'écart type de l'IPR90 calculés sur plusieurs sites pour les années 2018 et 2019, pour la température et les précipitations respectivement. Comme attendu, l'incertitude de la prévision d'ensemble augmente avec l'échéance tandis que celle des données fréquentielles reste relativement



(a)

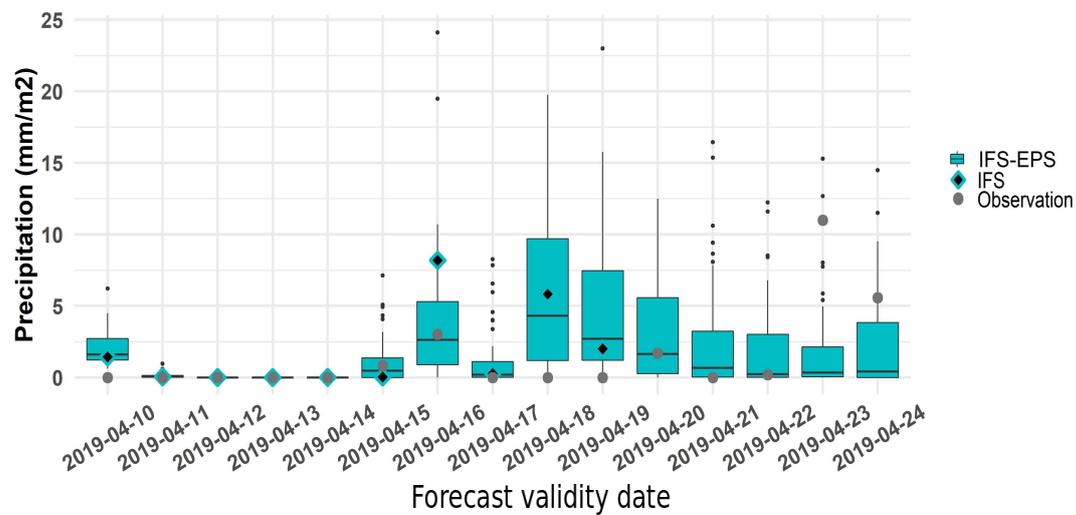


(b)

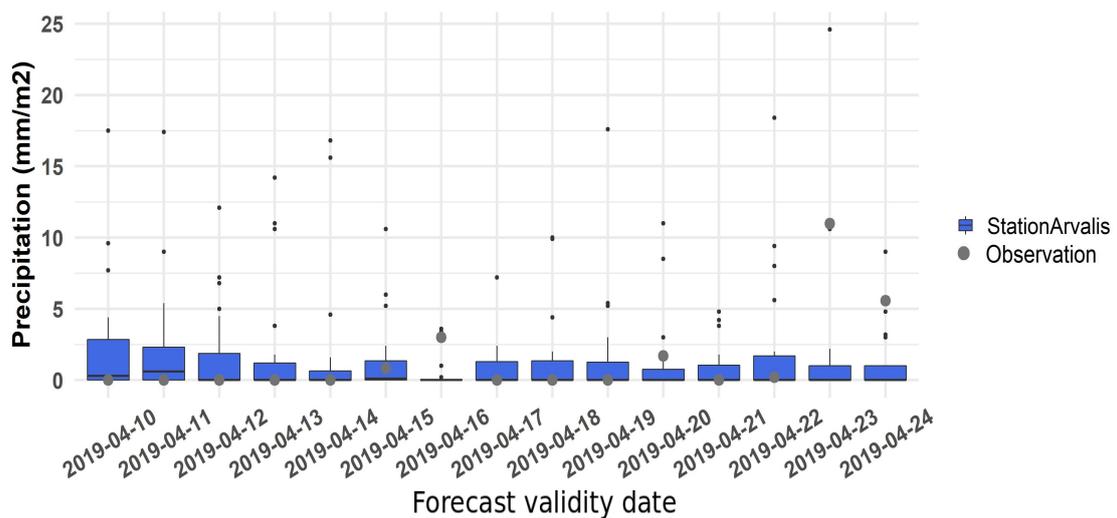
FIGURE 8.1 – Données de température à 2 mètres pour le site de Bordeaux. Les données observées sont en gris. (a) Prévision d'ensemble du modèle IFS-EPs (en turquoise) et prévision déterministe du modèle IFS (en noir), initialisées le 15/06/2018, sur leurs échéances respectives de 15 et 10 jours. (b) Données fréquentielles (observations des années 1998 à 2017) du 15/06 au 29/06 (en bleu).

constante. Pour la température on remarque que l'incertitude de la prévision d'ensemble devient proche de celle des données fréquentielles pour les dernières échéances. Pour les précipitations en revanche l'IPR moyen de la prévision d'ensemble reste bien à inférieur à celui des données fréquentielles pour toutes les échéances.

Ces différences au niveau de la quantification de l'incertitude entre les deux sources d'information météorologique considérées pourraient conduire à des différences significatives sur



(a)



(b)

FIGURE 8.2 – Données de précipitations quotidiennes sur le site de Poitiers Biard (86). Les données observées sont en gris. (a) Prédiction d'ensemble du modèle IFS-EPS (sous la forme de boîtes à moustache, en turquoise) et prédiction déterministe du modèle IFS (en noir), initialisées le 10/04/2019, sur leurs échéances respectives de 15 et 10 jours. (b) Données fréquentielles (observations des années 1998 à 2017) du 10/04 au 24/04 (en bleu).

les sorties agronomiques. Cet impact est exploré dans la section suivante.

Échéance (h)	Données de la prévision d'ensemble		Données fréquentielles	
	μ (IPR90)	σ (IPR90)	μ (IPR90)	σ (IPR90)
24	2.528280	0.8809718	10.27145	2.042471
48	3.079610	1.0102056	10.27244	2.040963
72	3.611566	1.1438567	10.27853	2.039947
96	4.244506	1.3245569	10.27227	2.047976
120	5.009395	1.6007036	10.27544	2.040922
144	5.754874	1.7522555	10.28133	2.045607
168	6.565103	1.8926525	10.27852	2.045506
192	7.363912	2.0064005	10.28415	2.044394
216	8.048783	2.0544344	10.28425	2.042114
240	8.542973	2.0195120	10.29549	2.044212
264	9.033236	1.9598822	10.29320	2.042725
288	9.478821	1.9462306	10.29223	2.048135
312	9.707027	1.9328487	10.29795	2.047751
336	9.926747	1.9102097	10.30596	2.050617
360	10.068888	1.8189365	10.31406	2.054181

TABLE 8.1 – Moyennes et écarts types des IPR90 sur les données de température (en degré) calculés pour la période d'intérêt d'EVA (de 1 mai-19 septembre) pour les années 2018 et 2019 sur l'ensemble des sites présentés dans le tableau 7.1. Les valeurs sont présentées en fonction de l'échéance de prévision, pour la prévision d'ensemble et les données fréquentielles.

8.2 Propagation des incertitudes météorologiques dans les OAD

Nous nous intéressons dans cette section à la propagation au travers des modèles EVA et Septo-LIS[®] des incertitudes météorologiques échantillonnées par la prévision d'ensemble IFS-EPS et par les données fréquentielles. Il s'agit notamment d'examiner si les différences entre ces données mises en évidence dans la section précédente conduisent à des différences notables sur l'estimation des dates de traitement.

8.2.1 Distribution des dates de traitement : Modèle EVA

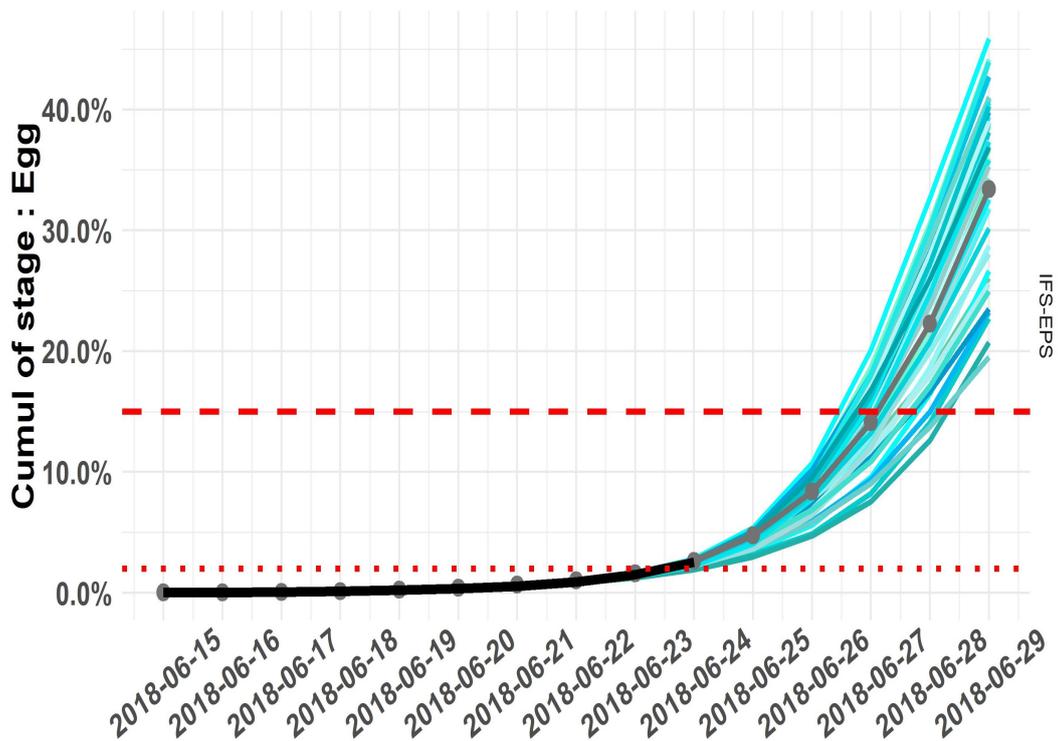
L'utilisation des données météorologiques probabilistes en entrée du modèle EVA permet d'obtenir plusieurs trajectoires d'évolution dynamique du ravageur à partir de la date de la prévision. En appliquant un seuil de décision sur le stade oeuf de 2% (traitement chimique) ou 15% (biocontrôle), nous obtenons la distribution des dates préconisées considérées comme optimale pour appliquer le traitement (Aleksavska *et al.*, 2018). Pour illustrer les résultats probabilistes agronomiques dans le cas d'EVA, nous considérons les prévisions météorologiques d'ensemble d'IFS-EPS et les données fréquentielles à Bordeaux, présentées précédemment (Figure 8.1).

Échéance (jour)	Données de la prévision d'ensemble		Données fréquentielles	
	$\mu(\text{IPR90})$	$\sigma(\text{IPR90})$	$\mu(\text{IPR90})$	$\sigma(\text{IPR90})$
1	1.116820	2.784742	8.130672	4.715871
2	1.416352	3.272913	8.167500	4.780750
3	1.908331	3.860662	8.268266	4.835938
4	2.259061	3.867937	8.352778	4.923318
5	2.749207	4.265386	8.505709	4.981486
6	3.109426	4.471996	8.418934	5.018792
7	3.646486	4.831442	8.459804	5.041437
8	4.193817	5.748336	8.566387	4.978650
9	4.640575	5.810941	8.797964	5.053700
10	4.946049	5.657455	8.818873	5.189894
11	4.815772	5.552894	9.297794	5.091422
12	4.710855	5.357694	9.381618	5.186552
13	4.540505	5.253469	9.591176	5.103852
14	4.510380	5.027863	9.791912	5.143719
15	4.553457	5.055717	10.045588	5.145445

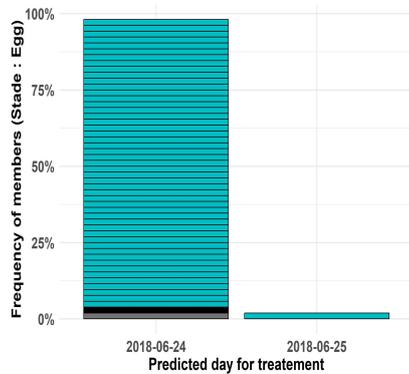
TABLE 8.2 – Moyennes et écarts types des IPR90 sur les données de précipitations quotidiennes (en mm) calculés pour le période d'intérêt de Septo-LIS[®] (de 10 avril-30 avril) pour les années 2018 et 2019 sur l'ensemble des sites présentés dans le tableau 7.1. Les valeurs sont présentées en fonction de l'échéance de prévision, pour la prévision d'ensemble et les données fréquentielles.

Simulation avec la prévision d'ensemble. Les trajectoires probabilistes du pourcentage d'individus au stade œuf, issues des simulations du modèle EVA utilisant la prévision d'ensemble, sont présentées sur la Figure 8.3(a). Nous pouvons observer que la dispersion augmente avec l'échéance. Avec le seuil de 2% la distribution ne comporte que deux dates possibles (Figure 8.3(b)) dont une très clairement majoritaire, tandis qu'avec le seuil de 15% trois dates sont possibles (Figure 8.3(c)). Ceci est lié au fait que le seuil 15 % est atteint avec des prévisions pour des échéances plus lointaines, pour lesquelles l'incertitude est plus grande. Pour les deux seuils, la date de référence fait partie des distributions prévues. Par ailleurs, il peut être remarqué que la date simulée avec la prévision déterministe IFS pour le seuil de 2 % fait partie de la distribution des dates de traitement prévue par la prévision d'ensemble. En revanche, en raison de son échéance maximale plus courte, elle ne permet pas de franchir le seuil de 15%, il est donc positionné dans le cas "beyond".

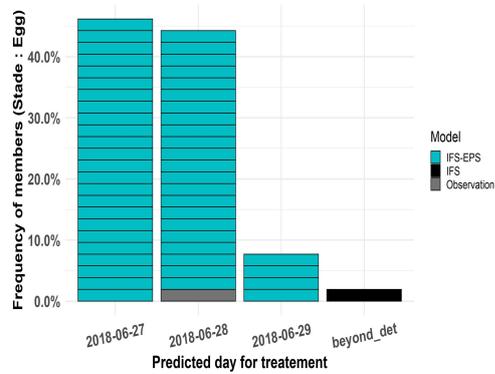
Simulation avec les données fréquentielles. Les sorties probabilistes d'EVA associées aux données fréquentielles et aux deux seuils de décision sont présentées sur la Figure 8.4. L'incertitude de ces prévisions est nettement plus grande que dans le cas précédent, en particulier pour le seuil de 2% pour lequel la distribution des dates de traitement s'étale sur 5 jours (Figure 8.4(b)). Pour le seuil de 15 % (Figure 8.4(c)) il y a 4 dates de traitement prévues et deux trajectoires n'atteignent pas le seuil en fin de prévision (catégorie "beyond"). Cette plus



(a)



(b)



(c)

FIGURE 8.3 – (a) Résultats des simulations d'ensemble EVA pour l'évolution du stade œuf à Bordeaux, obtenus en utilisant la prévision déterministe IFS (en noir) et les prévisions d'ensemble IFS-EPS (en turquoise). La simulation de référence, couplée uniquement aux observations, est superposée en gris. Les simulations débutent le 15 juin 2018 et couvrent des échéances de 10 jours (avec IFS) ou 15 jours (avec IFS-EPS). Histogrammes des dates de déclenchement du traitement (b) pour le seuil 2 % et (c) pour le seuil 15 %. La catégorie "beyond_det" indique que le seuil n'a pas été franchi à la fin de la dernière date de validité de la prévision déterministe.

grande incertitude est directement liée à l'incertitude des températures, qui est plus grande que celle de la prévision d'ensemble sur les courtes échéances. Cet exemple illustre donc bien le potentiel de la prévision d'ensemble par rapport à l'utilisation de données fréquentielles qui fournissent une incertitude certainement surestimée, correspondant à une variabilité climatique inter-annuelle.

8.2.2 Distribution des dates de traitement : Modèle Septo-LIS[®]

La même étude est faite avec le modèle Septo-LIS[®] pour la prévision des dates de traitement T1. Pour cela, nous avons utilisé les prévisions météorologiques d'ensemble d'IFS-EPS et les données fréquentielles à Poitiers Biard, présentées sur les figures 8.2(a) et 8.2(b) respectivement.

Simulation avec la prévision d'ensemble. La Figure 8.5 présente un exemple de dates prévues en utilisant les prévisions IFS et IFS-EPS. Ces simulations débutent le 10 avril 2019, soit 13 jours avant la date de traitement de référence. Nous remarquons que la distribution prévue présente un support étendu sur une semaine, avec la date la plus précoce prévue le 17/04/2019 et la date la plus tardive prévue le 23/04/2019 ; de plus deux trajectoires n'ont pas atteint le seuil nécessaire. Cette incertitude, largement plus grande que celle observée sur les simulations EVA, est potentiellement liée à l'impact des précipitations, dont la prévisibilité est plus faible que celle de la température. Par ailleurs la date de référence et la date prévue par la simulation déterministe se trouvent dans cette distribution. Néanmoins la date déterministe ne coïncide pas avec la date de traitement de référence, ce qui suggère que ne considérer que les données déterministes peut conduire à une mauvaise prise de décision. Dans le cas présent, utiliser la prévision d'ensemble peut a minima alerter sur l'incertitude associée à la simulation déterministe.

Simulation avec les données fréquentielles. La Figure 8.6 présente les dates prévues en utilisant les données fréquentielles. La distribution prévue a un support beaucoup plus large (au minimum 10 jours) que celle de la figure précédente, avec la date la plus précoce prévue le 14/04/2019, ce qui pourrait avoir un impact sur la mise en œuvre d'un traitement phytosanitaire trop précoce par des agriculteurs plus prudents, qui risque de devoir être renouvelé faute d'efficacité suffisante. Par ailleurs, il existe sept trajectoires qui n'ont pas atteint le seuil nécessaire. Cela représente un pourcentage élevé, soit 35 % des membres qui ne fournissent pas d'information. Cette incertitude est beaucoup plus importante que celle observée dans le cas de la prévision d'ensemble, et pose la question de l'utilité d'une information aussi incertaine à cette échelle de temps, pour la prise de décision.

Évolution des incertitudes dans le temps Les incertitudes sur les dates de traitement obtenues par le modèle Septo-LIS[®] sont plus importantes, probablement en raison du caractère plus incertain des précipitations à longue échéance. Des simulations plus proches de la

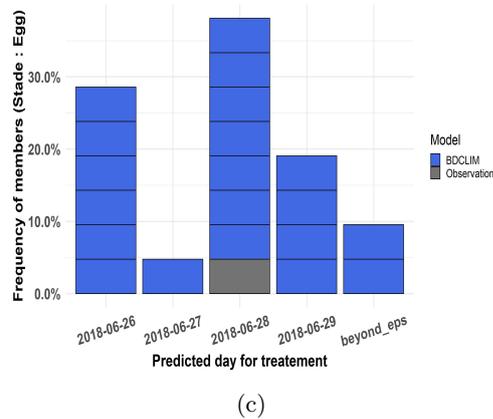
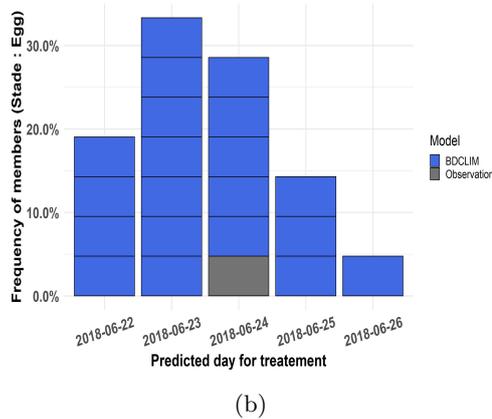
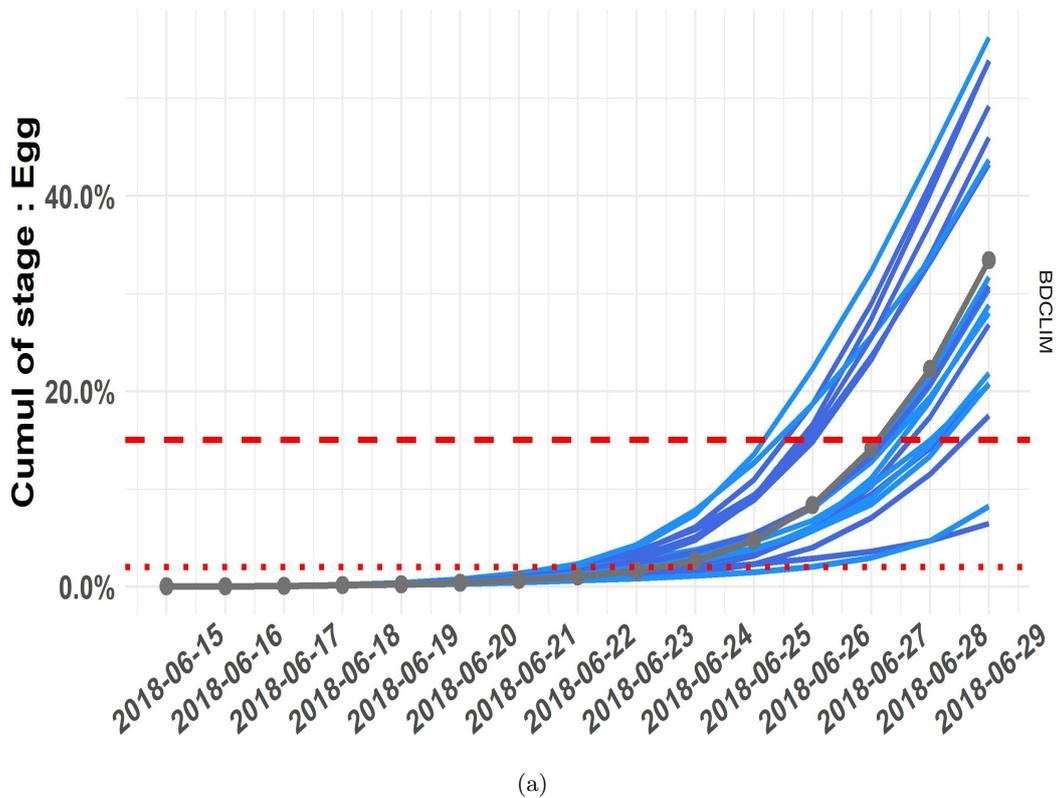


FIGURE 8.4 – (a) Résultats des simulations d’ensemble EVA pour l’évolution du stade œuf à Bordeaux, obtenus en utilisant les données fréquentielles. La simulation de référence, couplée uniquement aux observations, est superposée en gris. Les simulations débutent le 15 juin 2018 et couvrent des échéances de 15 jours. Histogrammes des dates de déclenchement du traitement (b) pour le seuil 2% et (c) pour le seuil 15%. La catégorie "beyond_eps" indique que le seuil n’a pas été franchi à la fin de la dernière date de validité de la prévision probabiliste.

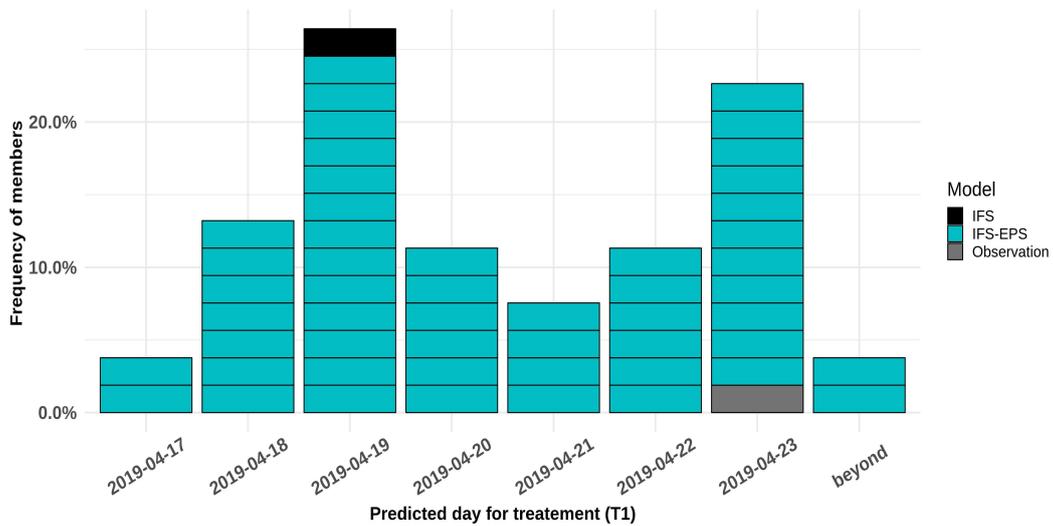


FIGURE 8.5 – Histogramme des dates du traitement T1 prévues par Septo-LIS[®] à Poitiers Biard, obtenu en utilisant les prévisions IFS-EPS à 15 jours d'échéance. La date obtenue en utilisant la prévision déterministe IFS est en noir et la date de référence en gris. Les simulations débutent le 10 avril 2019.

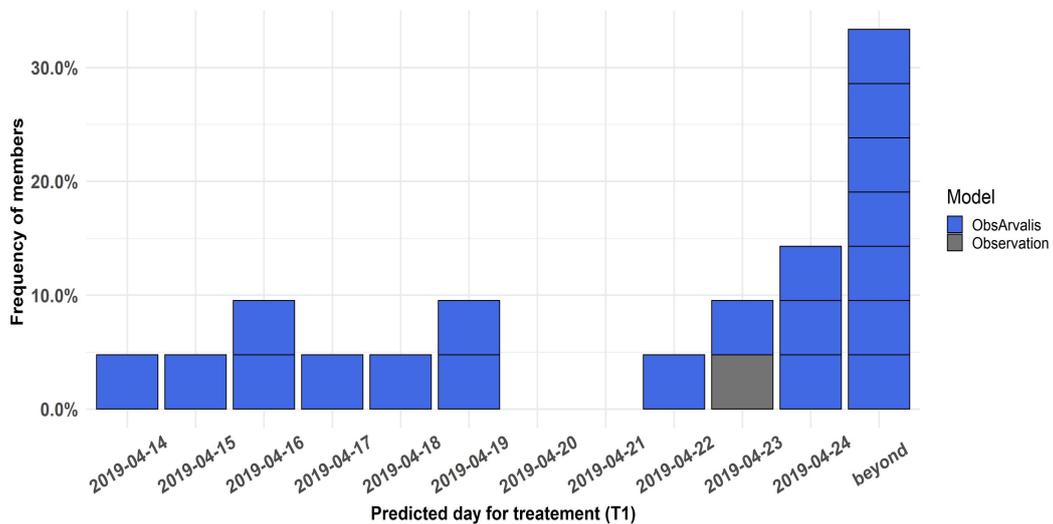


FIGURE 8.6 – Histogramme des dates du traitement T1 prévues par Septo-LIS[®] à Poitiers Biard, obtenu en utilisant les données fréquentielles (en bleu) et la date de référence (en gris). Les simulations débutent le 10 avril 2019.

date de traitement de référence et l'utilisation de la prévision d'ensemble pourraient aider à réduire cette incertitude. En revanche, dans le cas de l'utilisation des données fréquentielles, l'incertitude reste importante même à plus brève échéance (Figure 8.7).

8.3 Évaluation agronomique de la prévision d'ensemble par rapport aux données fréquentielles

La section précédente a permis, au travers de l'analyse de quelques cas d'étude, de mettre en évidence le potentiel apport de la prévision d'ensemble dans les deux OAD étudiés. Il s'agit dans cette partie d'évaluer objectivement les performances de la prévision d'ensemble et des données fréquentielles sur les sorties agronomiques afin de répondre à la question suivante : "la prévision d'ensemble fournit-elle en moyenne une estimation plus précise des dates de traitement que l'approche fréquentielle tout en maintenant une incertitude raisonnable ?"

Cette évaluation utilise les outils de vérification présentés dans le chapitre 6. La précision sera évaluée en considérant le biais de la moyenne des dates prévues et la couverture des observations, tandis que l'incertitude sera évaluée avec l'IPR90. Une bonne prévision probabiliste doit être caractérisée par un faible biais, une couverture des observations élevée et une incertitude réduite.

Nous présentons dans cette section une évaluation des simulations EVA et Septo-LIS[®] pour les années 2018 et 2019. Pour EVA, elle est effectuée sur les 8 vignobles listés dans le tableau 7.1 et elle correspond à la date de traitement pour la 2^{ème} génération de larves de l'année (la plus critique). Pour Septo-LIS[®], elle est réalisée sur les 18 parcelles de blé listées dans le tableau 7.1. Les scores sont moyennés sur tous les sites et leur évolution est représentée en fonction de l'intervalle de temps entre la date de traitement de référence et la date de début de la prévision. Au-delà de l'échéance maximale de la prévision d'IFS-EPS (15 jours), si le seuil de traitement n'est pas atteint par certains membres, les dates prédites de traitement ont été arbitrairement fixées à 17 afin de permettre les calculs des scores sur la totalité de l'ensemble. Mentionnons également que les simulations utilisant la prévision d'ensemble ont 50 membres tandis que celles utilisant les données fréquentielles n'ont que 20 membres. Afin que cette différence de taille n'affecte pas la comparaison, les scores pour la prévision d'ensemble ont été calculés sur un échantillon de 20 membres (tirés aléatoirement dans le cas d'IFS-EPS).

8.3.1 Application au modèle EVA

Biais (précision). Le biais de la date moyenne de traitement est présenté sur la Figure 8.8. Pour les prévisions à court terme correspondant à des intervalles de temps compris entre 1 et 3 jours, le biais est nul quelle que soit la prévision de température utilisée. Pour les intervalles de temps compris entre 4 et 15 jours, le biais augmente vers des valeurs positives. Cela signifie que la date de traitement est légèrement trop tardive par rapport à la date de référence, ce

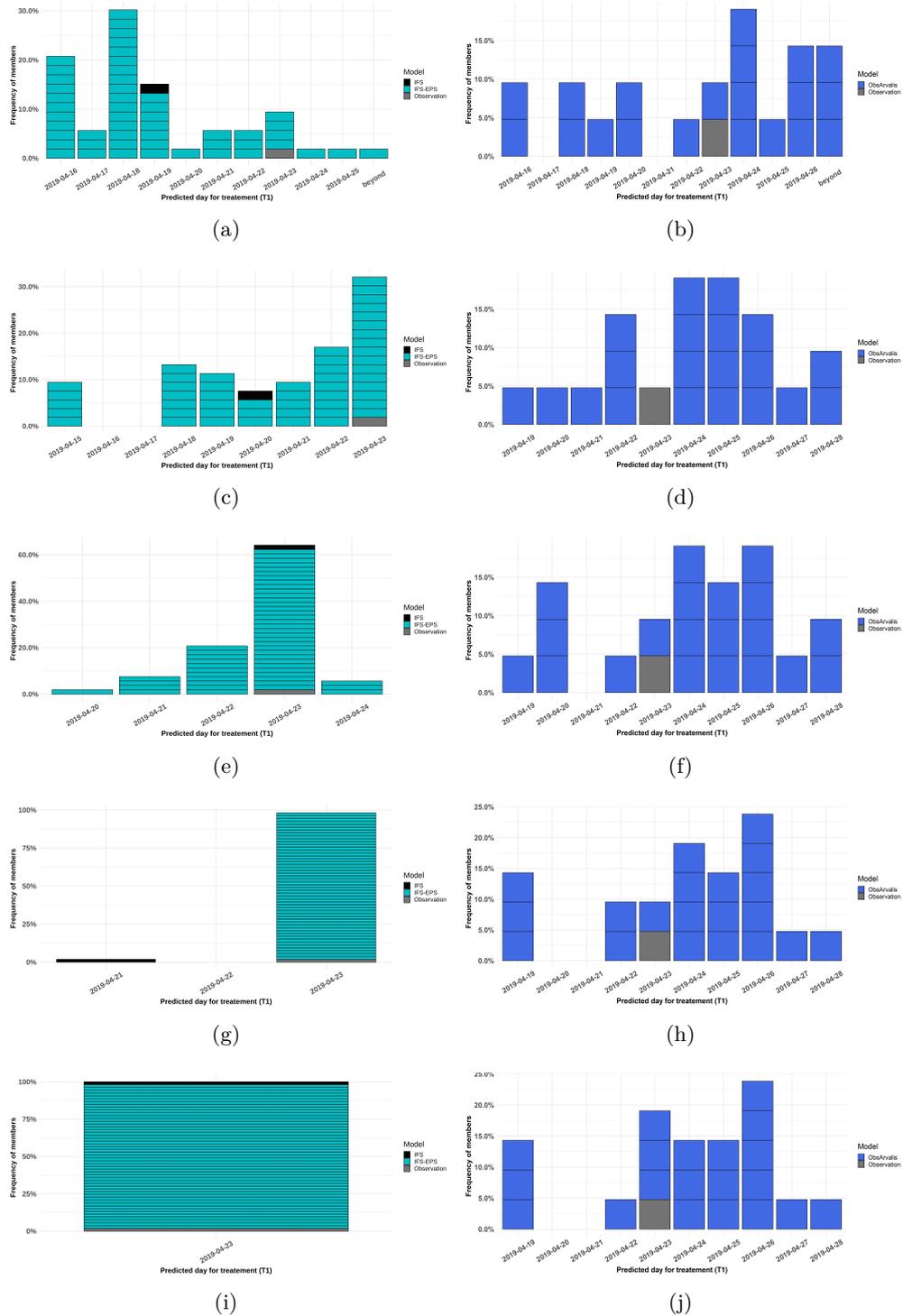


FIGURE 8.7 – Résultats des simulations d'ensemble Septo-LIS® à Poitiers Biard, obtenus en utilisant les données de prévisions d'ensemble (colonne gauche) et les données fréquentielles (colonne droite). La date de début des données probabiliste est : (a) et (b) 2019/04/12, (c) et (d) 2019/04/15, (e) et (f) 2019/04/16, (g) et (h) 2019/04/17, (i) et (j) 2019/04/18

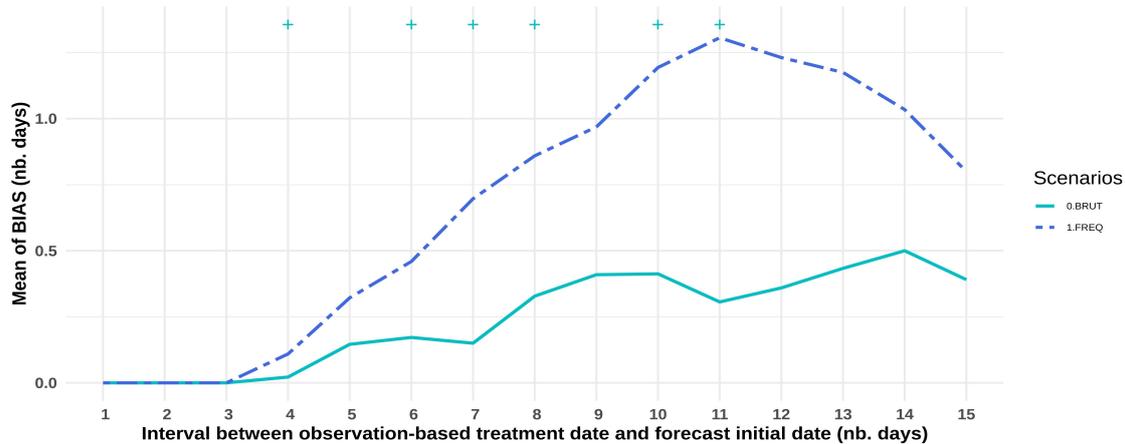


FIGURE 8.8 – Biais de la moyenne des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations EVA d’ensemble utilisant les prévisions IFS-EPS (0.BRUT, en turquoise) et les données fréquentielles (1.FREQ, en bleu), pour la deuxième génération et en utilisant le seuil de décision de 15%. Les résultats sont agrégés sur l’ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l’ensemble de la prévision d’ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

qui pourrait être lié au biais négatif de la température qui retardera le développement du ver de la grappe. Le biais obtenu avec les données fréquentielles est presque deux fois plus élevé que le biais associé aux prévisions IFS-EPS. L’apport de la prévision d’ensemble est statistiquement significatif (au sens du test de Wilcoxon) pour des intervalles de temps entre 4 et 11 jours.

Couverture des observations (précision). Sur la Figure 8.9, on observe le pourcentage des dates de traitement de référence qui appartiennent aux dates prévues. Jusqu’à un intervalle de quatre jours, il n’y a pas de différence entre l’approche fréquentielle et la prévision IFS-EPS. Pour des intervalles plus longs, la couverture d’observation diminue de manière importante, jusqu’à moins de 70% pour les données de prévision. Les données fréquentielles, en raison de leur plus grande variabilité, parviennent à maintenir un niveau élevé de couverture d’observation.

IPR90 (acuité). La Figure 8.10 présente l’IPR90 médian des distributions prévues. Pour des prévisions initialisées jusqu’à 4 jours avant la date de traitement de référence l’IPR90 est égal à 0 avec l’approche fréquentielle. En utilisant la prévision d’ensemble IFS-EPS la valeur de l’IPR90 reste nulle jusqu’à 8 jours avant la date de traitement de référence. Dans les deux cas, plus nous nous éloignons de la date de traitement de référence, plus l’IPR90 augmente. En utilisant IFS-EPS, l’incertitude sur la date de traitement n’est que d’un jour pour les prévisions à une semaine et de deux jours pour les prévisions à deux semaines. Lorsque l’on

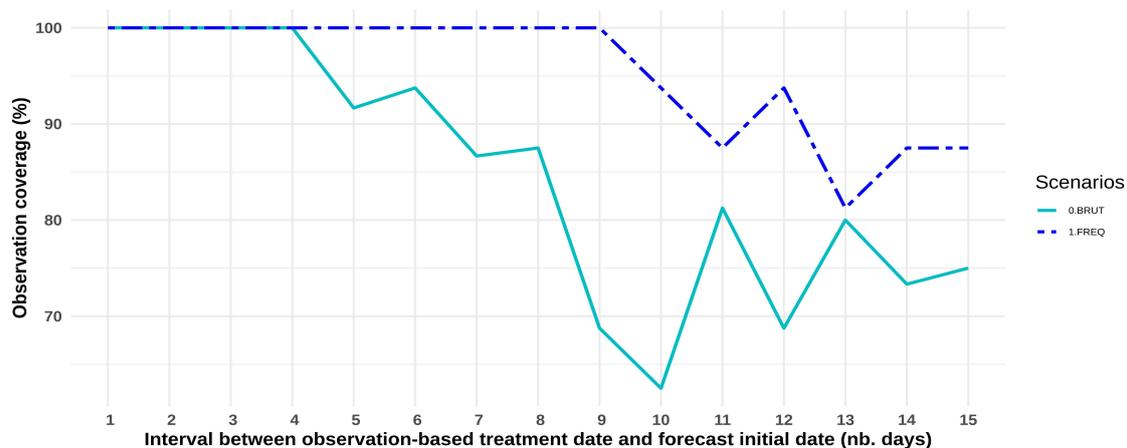


FIGURE 8.9 – Couverture d’observation des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Elle est calculée à partir des simulations EVA d’ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération pour un seuil de décision 15%. Les résultats sont agrégés sur l’ensemble des sites considérés et pour les années 2018 et 2019.

utilise les données fréquentielles, l’incertitude sur les dates de traitement est beaucoup plus élevée, en particulier pour la période comprise entre une et deux semaines avant la date de traitement de référence.

CRPS (performance globale). La Figure 8.11 montre les valeurs de CRPS pour les deux distributions de dates de traitement considérées. Pour les 3 premiers jours, ce score reste nul. Au-delà, il augmente légèrement pour la prévision d’ensemble et beaucoup plus avec les données fréquentielles. L’apport de la prévision d’ensemble est statistiquement significatif pour des intervalles de temps de 4 à 11 jours.

8.3.2 Application au modèle Septo-LIS®

Biais (précision) La Figure 8.12 indique que les deux approches présentent des biais faibles légèrement positifs sur les deux premiers jours. A partir du 3ème jour, le biais devient négatif, ce qui signifie que les dates de traitement prévues sont en moyenne trop précoces. Il est important de noter que le biais moyen de la prévision d’ensemble diminue plus rapidement vers des valeurs négatives et devient significativement moins bon par rapport à l’approche fréquentielle au-delà de 8 jours d’échéance. Ce comportement pourrait être lié à un excès de pluie dans les prévisions d’ensemble, conduisant à une propagation de la maladie plus rapide, mais la complexité du modèle ne nous permet pas de valider cette relation de causalité sur la base de ces seuls résultats.

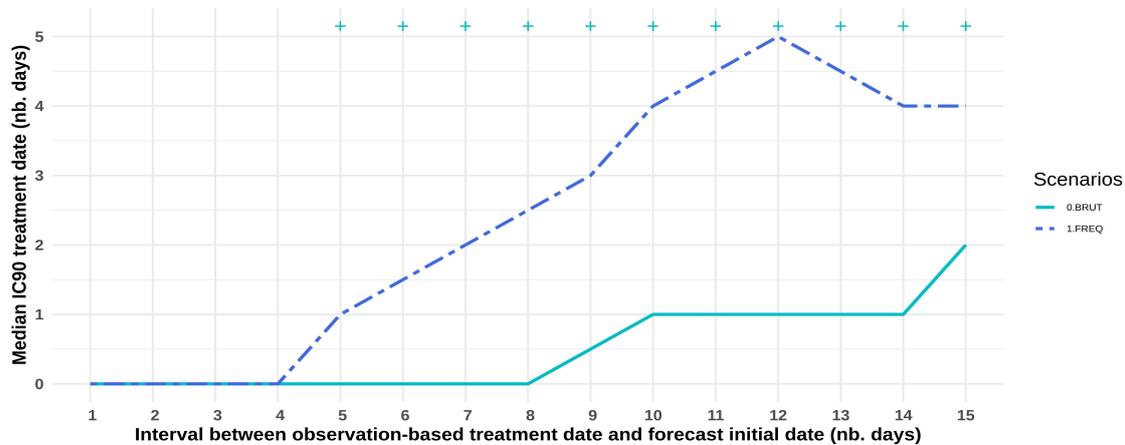


FIGURE 8.10 – Médiane des IPR90 des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations EVA d’ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération et en utilisant le seuil de décision 15%. Les résultats sont agrégés sur l’ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l’ensemble de la prévision d’ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

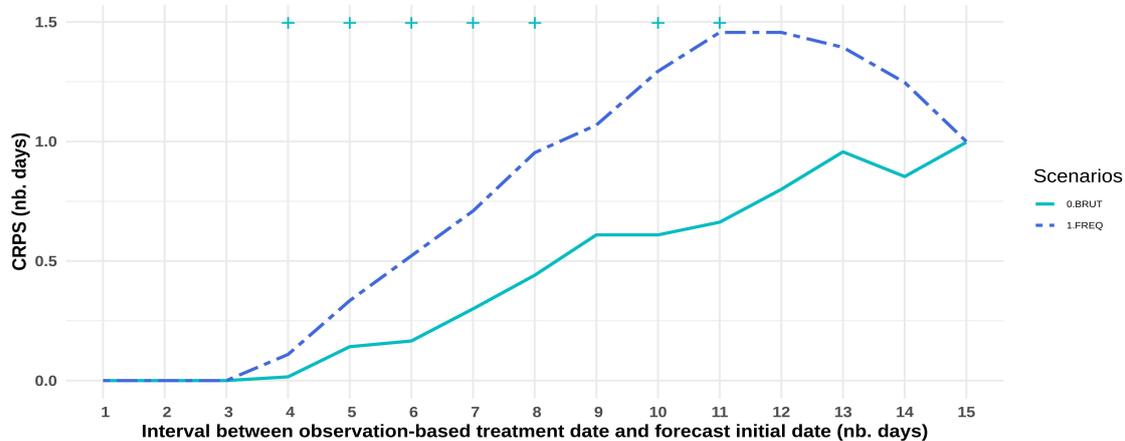


FIGURE 8.11 – CRPS des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les CRPS sont calculés à partir des simulations EVA d’ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu), pour la deuxième génération pour un seuil de décision de 15%. Les résultats sont agrégés sur l’ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l’ensemble de la prévision d’ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

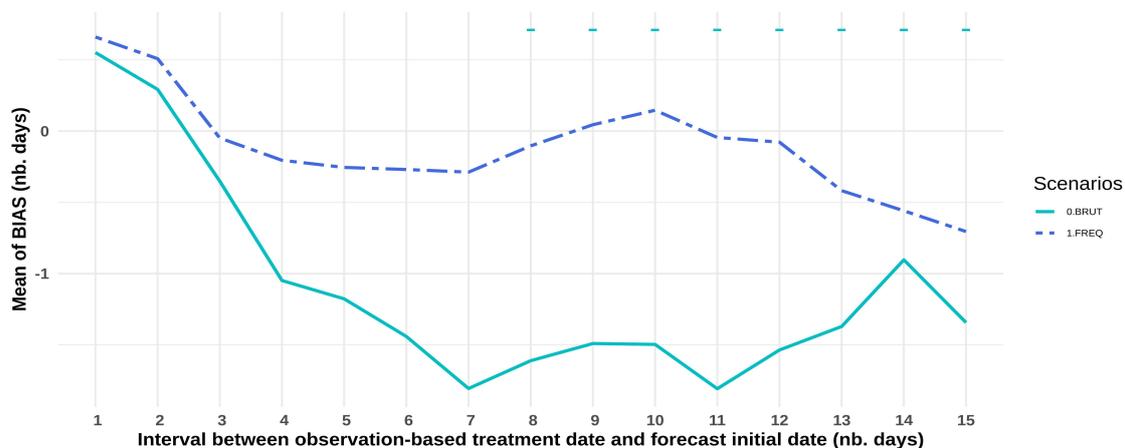


FIGURE 8.12 – Biais de la moyenne des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les biais sont calculés à partir des simulations Septo-LIS[®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

IPR90 (accuité). La Figure 8.13 montre l'évolution de la médiane des IPR90 des dates de traitement prévues. Dans le cas de la prévision d'ensemble, l'IPR90 reste faible au cours de la première semaine et augmente au-delà. Pour l'approche fréquentielle, l'IPR90 est beaucoup plus élevé à toutes les échéances. Il s'agit là d'un résultat attendu des premières analyses visuelles des données d'entrée.

Couverture des observations (précision). La Figure 8.14 indique qu'avec l'approche fréquentielle l'observation fait partie des dates prévues par l'ensemble dans 90% à 100% des cas tout au long des échéances. Ceci est un résultat attendu puisque le support des distributions de dates est grand. En revanche la couverture des observations associée aux prévisions d'ensemble atteint son maximum de 60-70% pour des intervalles de 2 à 6 jours et au delà de 13 jours.

CRPS (performance globale) La Figure 8.15 représente les valeurs moyennes de CRPS pour les deux distributions de dates de traitement considérées. Pour les trois premiers jours, ce score reste proche pour les deux approches probabilistes. Au-delà, jusqu'à environ une semaine il augmente légèrement pour la prévision fréquentielle et il diminue avec la prévision d'ensemble. La prévision d'ensemble est significativement meilleure que l'approche fréquentielle uniquement pour les intervalles de 5 et 6 jours. Au-delà d'une semaine il est plus difficile de tirer des conclusions robustes, mais l'apport de la prévision d'ensemble est moins net.

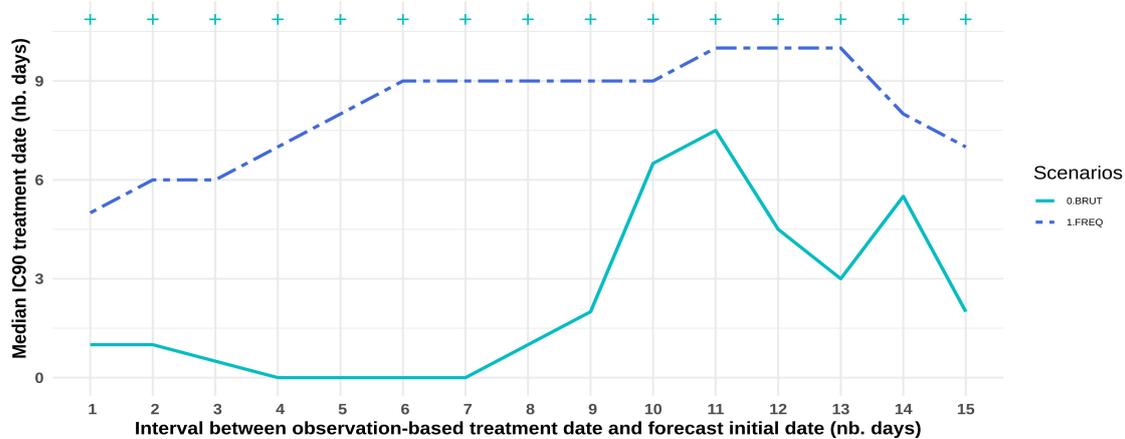


FIGURE 8.13 – Médiane des IPR90 des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les IPR90 sont calculés à partir des simulations Septo-LIS[®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble de la prévision d'ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

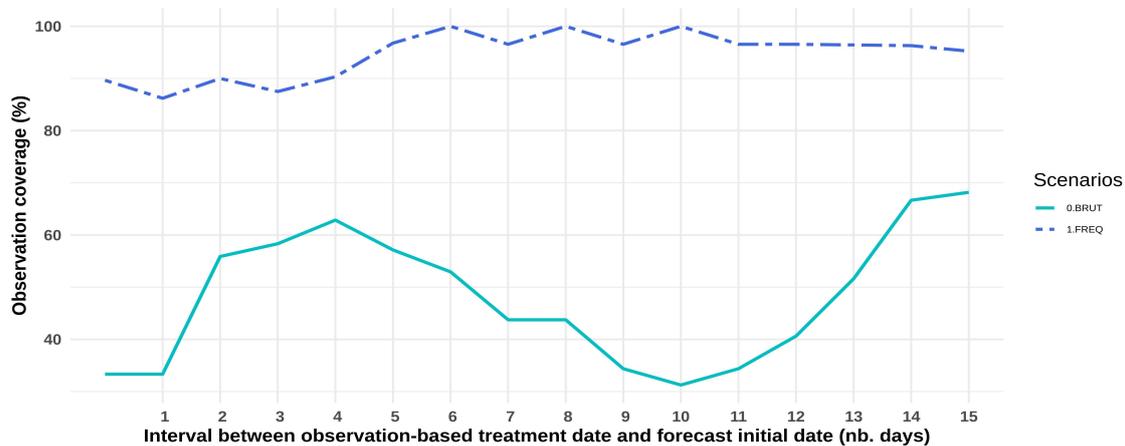


FIGURE 8.14 – Couverture d'observation des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Elle est calculée à partir des simulations Septo-LIS[®] d'ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l'ensemble des sites considérés et pour les années 2018 et 2019.

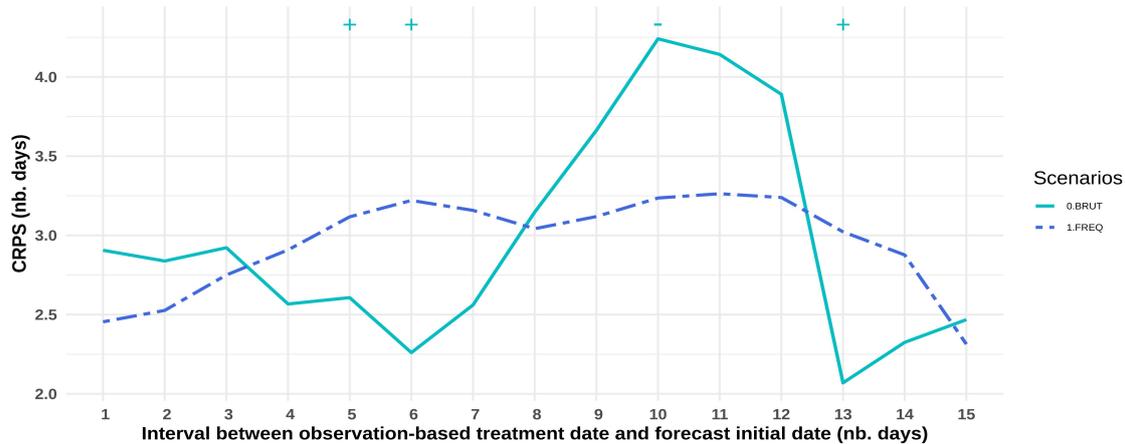


FIGURE 8.15 – CRPS des dates de traitement prévues, en fonction du nombre de jours entre la date de traitement de référence et le début de la prévision. Les CRPS sont calculés à partir des simulations Septo-LIS[®] d’ensemble utilisant les prévisions IFS-EPS (en turquoise) et les données fréquentielles (en bleu). Les résultats sont agrégés sur l’ensemble des sites considérés et pour les années 2018 et 2019. Les signes + (resp. -) en haut indiquent que la performance de l’ensemble de la prévision d’ensemble est statistiquement meilleure (resp. moins bonne) que la performance des données fréquentielles, selon le test de Wilcoxon.

8.4 Propagation des incertitudes des données météorologiques de PNT à haute résolution pour des simulations agronomiques

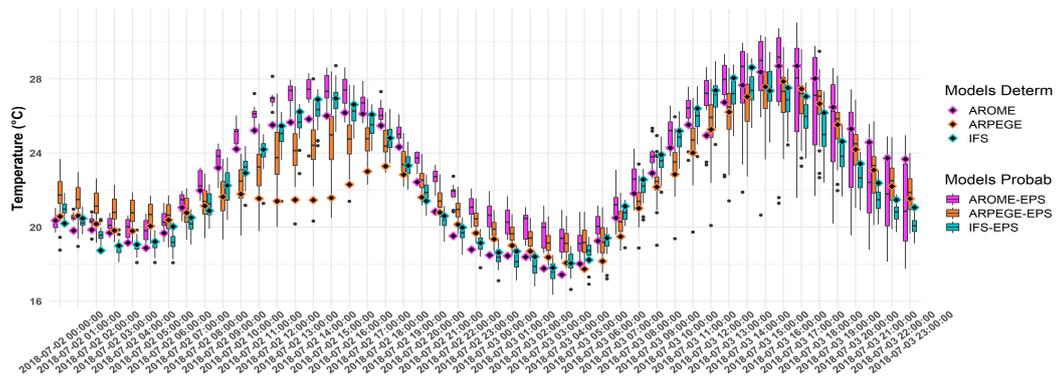
Les prévisions météorologiques issues des modèles de plus haute résolution spatiale avec des échéances plus courtes peuvent également être prises en compte pour les simulations agronomiques. Ces modèles ne permettent pas d’anticiper la situation agronomique à moyenne échéance, mais leur information de fine échelle et leur variabilité plus forte à courte échéance peut permettre de mieux estimer l’incertitude des dates de traitement, notamment dans le cas d’EVA où l’effet cumulatif et la variabilité réduite des prévisions d’ensemble IFS-EPS aux courtes échéances conduisent à une variabilité faible des sorties agronomiques. Un exemple de simulations agronomiques obtenues à partir des prévisions AROME-EPS, ARPEGE-EPS et IFS-EPS est présenté sur la Figure 8.16. Pour cette simulation, nous avons utilisé les observations météorologiques disponibles jusqu’au 1 juillet 2018 compris puis les prévisions ensemblistes au-delà. La Figure 8.16(a) présente les prévisions déterministe et ensembliste (sous forme de boîte à moustaches) de température horaire pour les 2 et 3 juillet 2018 à Bordeaux, pour les trois modèles. La Figure 8.16(b) montre une représentation de température journalière moyenne et la Figure 8.16(c) présente les sorties d’EVA correspondantes, pour la proportion du stade œuf. Un lien entre la distribution des prévisions météorologiques en entrée et la distribution des prédictions agronomiques en sortie peut être remarqué. En effet, différentes distributions de température impliquent différents développements du stade œuf. Il sera donc intéressant de profiter de cette information des modèles à haute résolution pour

affiner l'incertitude des simulations dans les tout premiers jours. Par ailleurs, comme les modèles agronomiques représentent des phénomènes cumulatifs, nous pouvons aussi nous attendre à des conséquences sur des échéances plus longues.

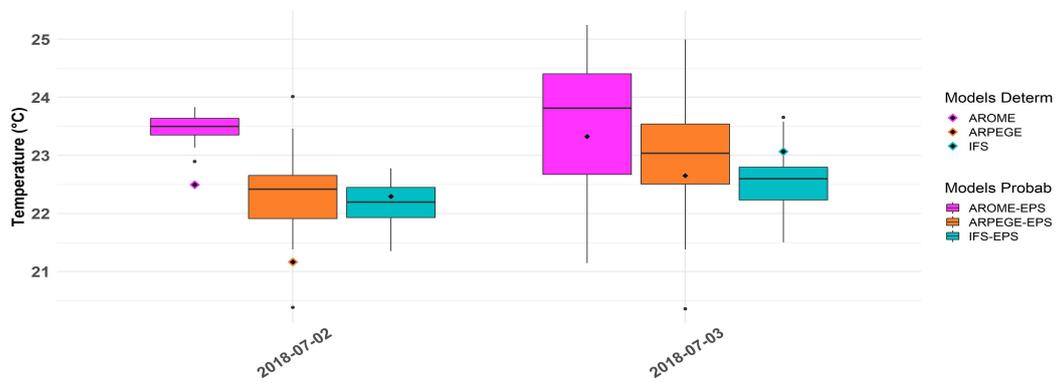
8.5 Conclusion

Ce chapitre a montré le potentiel de la prévision d'ensemble par rapport aux pratiques existantes : utilisation de prévisions déterministes et de données fréquentielles. Nous avons d'abord montré que les informations apportées par les simulations déterministes pouvaient conduire à de mauvaises prises de décision par l'analyse visuelle des résultats. D'autre part, l'approche probabiliste permet de compléter la prévision déterministe par une estimation de la confiance dans les dates de traitement prévues, ce qui pourrait aider à la prise de décision. Nous avons ensuite évalué objectivement les performances de la prévision d'ensemble et des données fréquentielles en utilisant plusieurs scores de vérification. Cette évaluation a montré un apport significatif de la prévision d'ensemble par rapport à l'approche fréquentielle, en particulier pour des échéances de une à deux semaines.

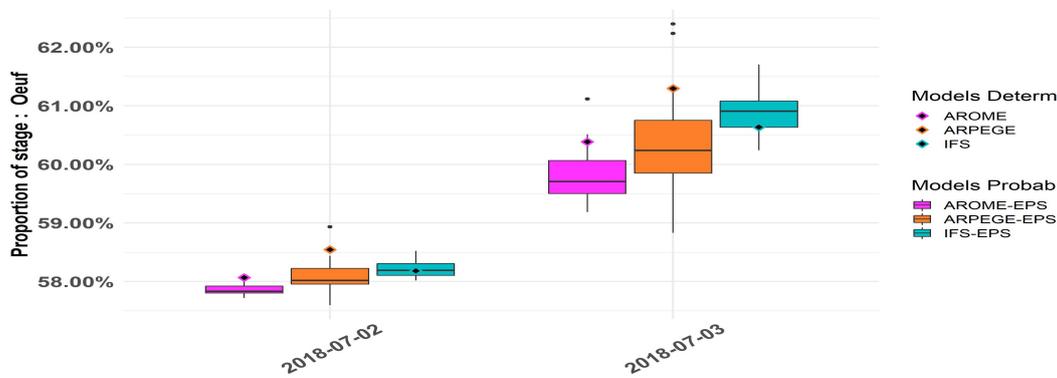
Nous avons enfin illustré le potentiel des prévisions d'ensemble à haute résolution AROME-EPS et ARPEGE-EPS, pour une représentation plus précise de l'incertitude aux courtes échéances. Intégrer ces prévisions dans les prévisions à long terme pourrait être bénéfique pour les simulations agronomiques. Ces aspect est examiné dans le chapitre suivant.



(a)



(b)



(c)

FIGURE 8.16 – (a) Prévisions des températures horaires fournies par les systèmes AROME, ARPEGE, IFS et leurs équivalents ensemblistes AROME-EPS, ARPEGE-EPS et IFS-EPS (boîtes à moustaches) pour les 2 et 3 juillet 2018 à Bordeaux. (b) Prévisions des températures quotidiennes moyennes fournies par les systèmes AROME, ARPEGE, IFS et leurs équivalents ensemblistes AROME-EPS, ARPEGE-EPS et IFS-EPS (boîtes à moustaches) pour les 2 et 3 juillet 2018 à Bordeaux. Les prévisions sont initialisées le 2 juillet 2018. (c) Proportions d’œufs associées, simulées par EVA.

Construction des prévisions d'ensemble cohérentes

Sommaire

9.1	Vérification des prévisions calibrées	80
9.1.1	Biais et RMSE	80
9.1.2	CRPS	80
9.2	Construction et caractéristiques des prévisions d'ensemble cohérentes	82
9.2.1	Première idée : Jointure avec le coupleur	82
9.2.2	Origine des membres sélectionnés à 48h	82
9.2.3	Répétition des membres pour la stratégie NN (critère d'unicité des membres)	83
9.2.4	Analyse des discontinuités de raccordement (critère de continuité des prévisions)	84
9.2.5	Exemple représentatif d'un ensemble de scénarios cohérents	85
9.3	La calibration : étape incontournable dans la construction des scénarios cohérents	85
9.4	Conclusion	87

Le chapitre précédent a mis en évidence l'intérêt potentiel d'utiliser les prévisions d'ensemble à court terme des modèles à haute résolution pour les simulations agronomiques. Leur variabilité à court terme pourrait en effet permettre de mieux représenter la variabilité associée des résultats agronomiques. Cependant, leur utilisation ne permettrait pas d'anticiper la planification de l'application de traitements phytosanitaires à long terme. Dans l'idée de tirer le meilleur parti de chaque modèle de PNT, ce chapitre présente les résultats des deux techniques de post-traitement utilisées dans cette thèse, la calibration et la jointure pour la construction de prévisions d'ensemble cohérentes. Tout d'abord, une évaluation de l'impact de la calibration statistique des prévisions de température de 2 mètres est présentée. Dans un second temps, la construction de la prévision d'ensemble cohérente est détaillée avec une analyse des caractéristiques de chaque méthode de jointure. On vérifiera en particulier que ces scénarios cohérents respectent les critères de continuité temporelle et d'unicité des membres. Les résultats s'appuient en grande partie sur l'article Aleksovska *et al.* 2020 soumis à *Weather and Forecasting* (voir Annexe D).

9.1 Vérification des prévisions calibrées

Dans cette première partie nous présentons l'évaluation des prévisions d'ensemble de température brutes et calibrées au moyen de la technique paramétrique EMOS présentée dans le chapitre 4. Les résultats sont montrés pour les trois systèmes AROME-EPS, ARPEGE-EPS et IFS-EPS.

9.1.1 Biais et RMSE

La Figure 9.1 présente le biais et la RMSE de la moyenne d'ensemble des prévisions de température à 2 mètres d'AROME-EPS, ARPEGE-EPS et IFS-EPS, avant et après la calibration. Comme attendu, la procédure de calibration a un impact positif significatif sur la performance des trois systèmes.

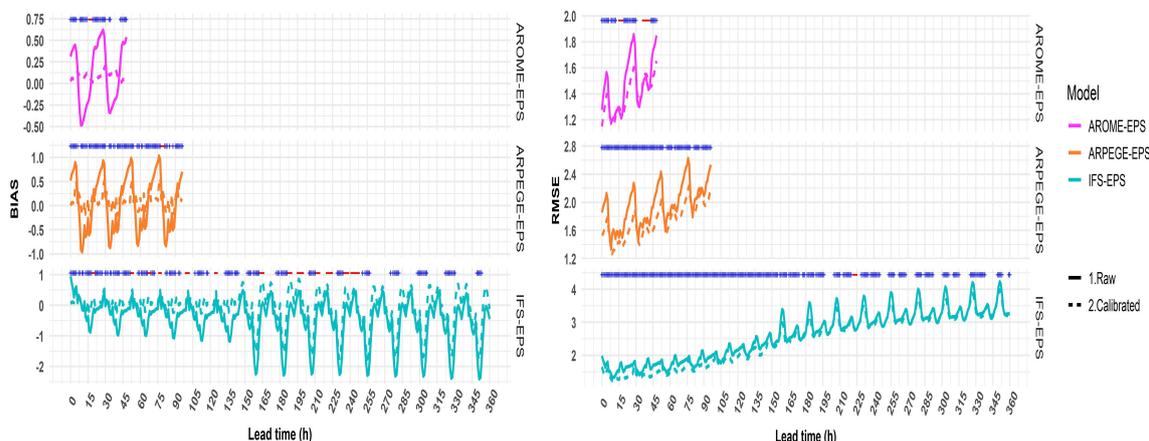


FIGURE 9.1 – Biais (à gauche) et RMSE (à droite) de la moyenne d'ensemble des prévisions de température à 2 mètres, en fonction de l'échéance de prévision. Les scores sont calculés pour les 11 sites de tableau 7.1 sur la période 01/06/2018-31/05/2019, avec les ensembles bruts (ligne continue) et calibrés (ligne pointillée). En haut : Arome-EPS, au milieu : ARPEGE-EPS et en bas : IFS-EPS. Les signes + (resp. -) en haut indiquent que la performance de l'ensemble post-traité est statistiquement meilleure (resp. moins bonne) que la performance de l'ensemble brut, selon le test de Wilcoxon.

9.1.2 CRPS

La Figure 9.2 présente le CRPS pour les trois ensembles bruts et calibrés. Concernant les prévisions brutes, on remarque qu'AROME-EPS a la meilleure performance sur ses échéances, tandis qu'ARPEGE-EPS est le moins bon. Comme attendu la calibration conduit à une diminution du CRPS des trois ensembles ainsi qu'à un rapprochement de leurs performances. AROME-EPS a désormais la meilleure performance jusqu'à environ 20h d'échéance. De 24 à 48 heures, AROME-EPS est meilleur qu'ARPEGE-EPS mais proche ou légèrement moins

bon qu'IFS-EPS. ARPEGE-EPS reste le moins bon sur toutes ses échéances, malgré une résolution horizontale plus fine que celle d'IFS-EPS. Ces résultats supportent donc le choix de construire les ensembles de scénarios cohérents avec les membres AROME-EPS uniquement sur les échéances 0-48h. Pour les échéances 48-96h, ARPEGE-EPS ayant une performance moins bonne qu'IFS-EPS, le raccordement des membres AROME-EPS à 48h est ouvert aux membres ARPEGE-EPS et IFS-EPS (au lieu des membres ARPEGE-EPS uniquement). Par ailleurs, étant donné les meilleures performances des ensembles calibrés, ils sont utilisés dans la suite pour la construction des prévisions d'ensemble cohérentes.

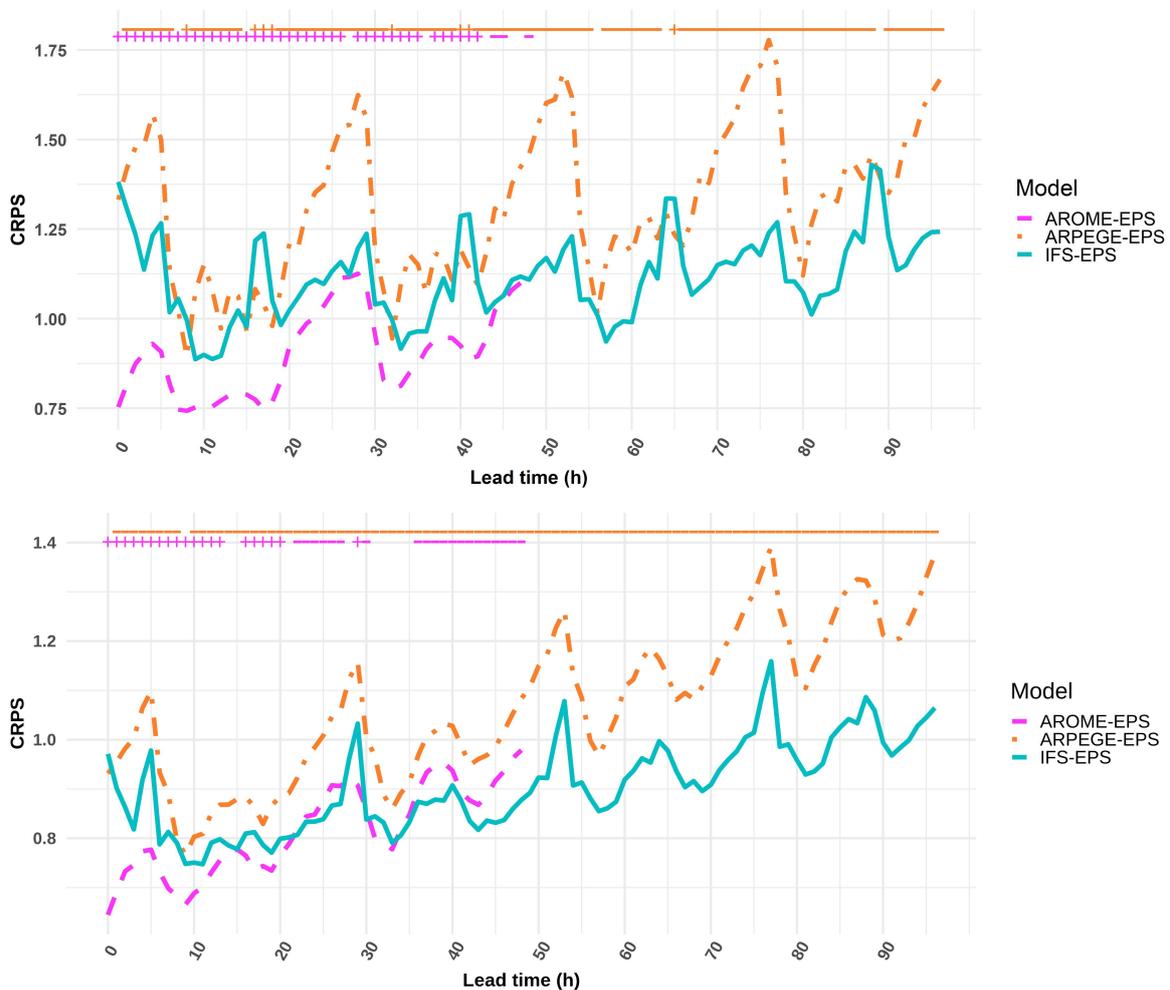


FIGURE 9.2 – CRPS des prévisions de température de 2 mètres en fonction de l'échéance de prévision. Les scores sont calculés pour les 11 sites décrits dans le tableau 7.1 sur la période 01/06/2018-31/05/2019 pour les trois ensembles bruts (haut) et calibrés (bas). Les signes + (resp. -) indiquent que la performance d'AROME-EPS et d'ARPEGE-EPS est statistiquement meilleure (resp. moins bonne) que la performance de d'IFS-EPS, selon le test de Wilcoxon.

9.2 Construction et caractéristiques des prévisions d'ensemble cohérentes

Nous rappelons que les stratégies proposées pour construire les prévisions d'ensemble cohérentes reposent sur deux choix principaux : utiliser les membres AROME-EPS uniquement au cours des 48 premières heures et sélectionner les membres suivants d'ARPEGE-EPS et d'IFS-EPS en fonction de leur proximité avec les membres AROME-EPS.

9.2.1 Première idée : Jointure avec le coupleur

Comme expliqué dans le chapitre 1, chaque membre AROME-EPS utilise comme conditions aux limites une sélection de membres ARPEGE-EPS. Un choix naturel pourrait donc consister à raccorder chaque membre AROME-EPS à son coupleur ARPEGE-EPS entre 48h et 96h d'échéance.

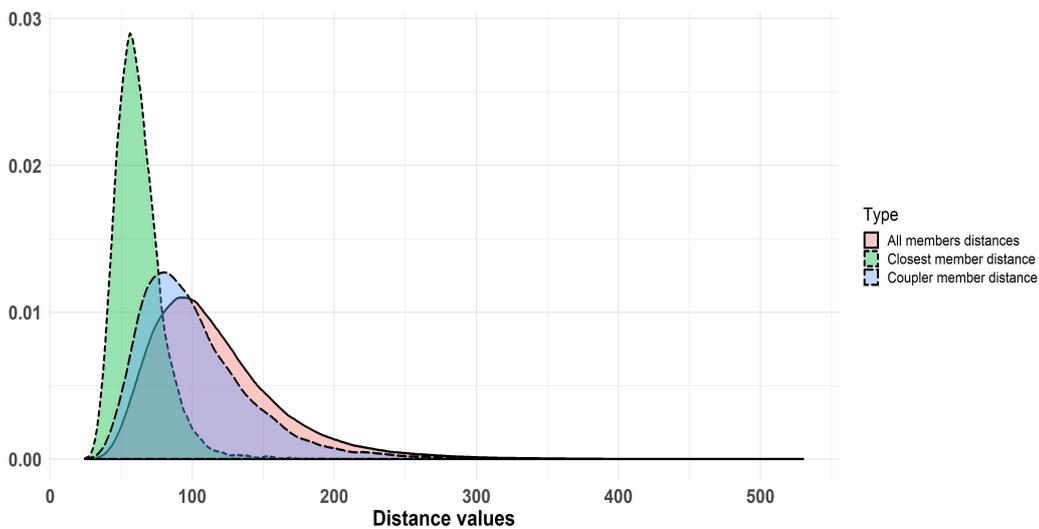
Afin d'évaluer la pertinence de ce choix, la Figure 9.3(a) présente la distribution des distances entre les membres AROME-EPS et leurs coupleurs, la distribution des distances entre les membres AROME-EPS et le membre ARPEGE-EPS le plus proche (au sens de la distance DTW) et la distribution des distances entre tous les membres AROME-EPS et tous les membres de ARPEGE-EPS, pour la période d'hiver 2018/2019. On peut remarquer que les distances avec les membres ARPEGE-EPS les plus proches sont les plus faibles, et sont nettement inférieures aux distances entre les membres AROME-EPS et leurs coupleurs. Ce résultat peut être expliqué par le fait que le couplage est fait pour des paramètres de grande échelle, et les prévisions de température à 2 mètres d'AROME sont plus influencées par des processus de fine échelle que par le couplage de grande échelle.

Sur la base de ces résultats, il a été choisi de ne pas effectuer le raccordement d'AROME-EPS aux membres coupleurs mais de s'orienter sur les membres les plus proches, comme détaillé dans la suite.

9.2.2 Origine des membres sélectionnés à 48h

Le raccordement sur la distance à 48h s'appuie sur le calcul de la matrice des distances dtw entre les membres AROME-EPS et les membres ARPEGE-EPS et IFS-EPS. Un algorithme de type plus proche voisin (NN) ou méthode hongroise (HU) permet ensuite d'affecter à chaque membre AROME-EPS le membre ARPEGE-EPS ou IFS-EPS qui le prolongera. Notons que la distance dtw est calculée sur les 7 dernières heures (c'est-à-dire sur les échéances 42h à 48h), valeur choisie suite à des tests de sensibilité. Pour comparaison, une sélection aléatoire (RN) des membres utilisés pour le raccordement est également effectuée.

Le tableau 9.1 présente le pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le raccordement à 48h, avec les trois stratégies RN, NN et HU, pour différentes saisons. Avec le tirage aléatoire, 60% des membres sélectionnés proviennent d'IFS-EPS en raison de



(a)

FIGURE 9.3 – Les distances entre membres AROME et ARPEGE calculées sur la période commune (de 0h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1). (a) Distributions des distances entre les membres AROME-EPS et leurs coupleurs (bleu), entre les membres AROME-EPS et le membre ARPEGE-EPS le plus proche (turquoise) et entre les membres AROME-EPS et tous les membres ARPEGE-EPS (rouge), calculées sur la période 01/02/2018-28/02/2019 en moyenne pour tous les sites du tableau 7.1.

la plus grande taille de cet ensemble. Les affectations basées sur la distance entre membres permettent une sélection plus équilibrée et certaines variations saisonnières sont observées. Au printemps et en été, les membres d'IFS-EPS sont légèrement plus nombreux que ceux d'ARPEGE-EPS, tandis que le contraire est observé en hiver. En automne, la répartition est proche du 50-50. Ces variations reflètent les distances moyennes entre les membres AROME-EPS et les membres ARPEGE-EPS et IFS-EPS. Il a été vérifié par exemple qu'en hiver les membres ARPEGE-EPS sont en moyenne plus proches des membres AROME-EPS (attention toutefois, plus proches ne signifie pas meilleurs).

On peut enfin remarquer que les pourcentages obtenus avec les méthodes NN et HU sont très similaires. En outre, les différences dans les sélections NN et HU, qui sont uniquement dues à la règle d'affectation (le critère de distance est le même), sont plutôt faibles. L'intersection entre les membres sélectionnés par les approches NN et HU est égale à 87% en moyenne pour ce raccordement à 48h.

9.2.3 Répétition des membres pour la stratégie NN (critère d'unicité des membres)

Comme expliqué dans le chapitre 5, une limitation majeure de l'approche du plus proche voisin est la répétition possible de certains membres ARPEGE-EPS et IFS-EPS. Afin d'exa-

2*Time Period	ARPEGE-EPS			IFS-EPS		
	RN	NN	HU	RN	NN	HU
2018-06-01-2018-08-31	40.7 %	45.65 %	45.41 %	59.3 %	54.35 %	54.59 %
2018-09-01-2018-11-30	40.7 %	49.60 %	49.10 %	59.3 %	50.40 %	50.90 %
2018-12-01-2019-02-28	40.7 %	53.00 %	52.50 %	59.3 %	47.00 %	47.50 %
2019-03-01-2019-05-31	40.7 %	46.27 %	45.87 %	59.3 %	53.73 %	54.13 %

TABLE 9.1 – Pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le raccordement à 48h avec les différentes stratégies d’affectation, pour quatre saisons. Les valeurs théoriques pour la sélection aléatoire sont de 40,7% pour ARPEGE-EPS et 59,3% pour IFS-EPS, en raison de leur taille respective.

miner cet aspect, la Figure 9.4 présente la distribution de la taille effective de l’ensemble (c’est-à-dire, le nombre de membres indépendants de l’ensemble) après les raccordements à 48h et 96h. On observe que la plupart des ensembles comptent entre 9 et 12 membres indépendants, tandis que les ensembles de 12 membres ne représentent que 20% et 10% des cas après les premier et deuxième raccordements respectivement. On remarque également que des ensembles de moins de 6 membres indépendants sont obtenus dans quelques cas extrêmes.

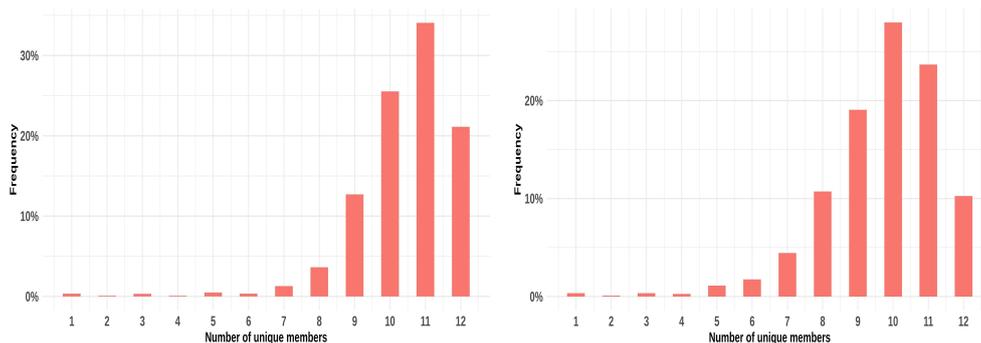


FIGURE 9.4 – Distribution de la taille effective des ensembles NN, calculée pour 11 sites du tableau 7.1 sur la période 06/01/2018-31/05/2019, gauche : après le raccordement à 48h et droite : après le raccordement à 96h.

9.2.4 Analyse des discontinuités de raccordement (critère de continuité des prévisions)

La capacité des stratégies de raccordement à générer des transitions réalistes entre deux membres est mesurée par l’amplitude des différences de température entre les échéances 48h et 49h pour la première jonction, et les échéances de 96h et 97h pour la deuxième jonction. La figure 9.5 présente la distribution de ces différences de température pour les trois stratégies. Les différences obtenues avec les membres IFS-EPS qui, par construction, sont des scénarios cohérents, sont prises comme référence. Comme attendu, les différences les plus importantes sont obtenues avec la stratégie RN, tandis que les discontinuités observées avec les affectations NN et HU sont beaucoup plus faibles et proches de celles observées dans les prévisions IFS-

EPS. Ces résultats indiquent donc que les raccordements obtenus avec les méthodes NN et HU sont en mesure de fournir des prévisions visuellement réalistes. Il convient également de noter que des discontinuités plus faibles sont observées pour le raccordement à 96h, en particulier pour la stratégie RN. Ceci est lié au fait que ce second raccordement n'est effectué que pour les membres ARPEGE-EPS sélectionnés à 48h, tandis que les membres IFS-EPS sélectionnés à 48h contribuent à réduire les écarts de température moyens.

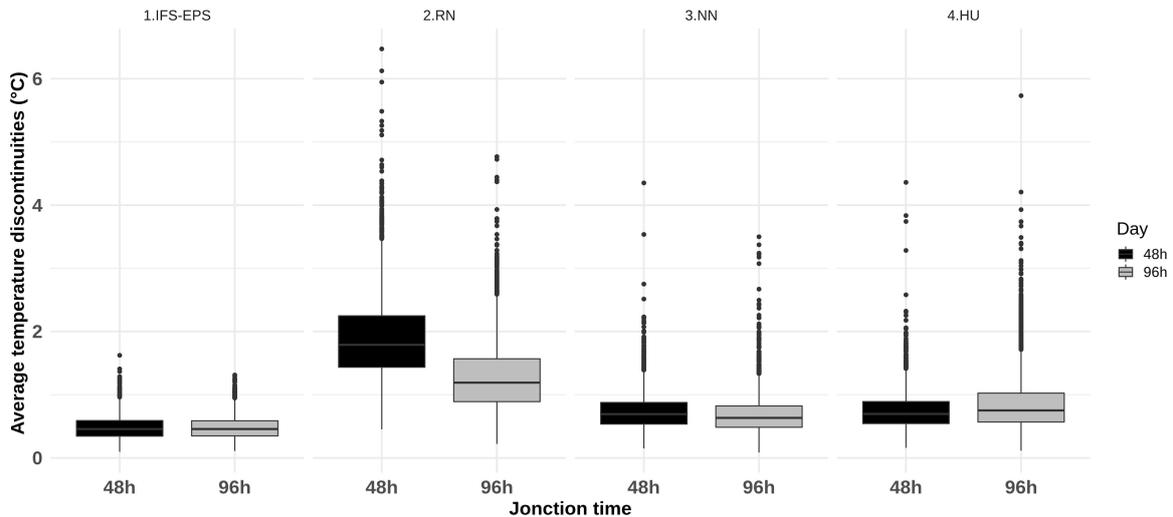


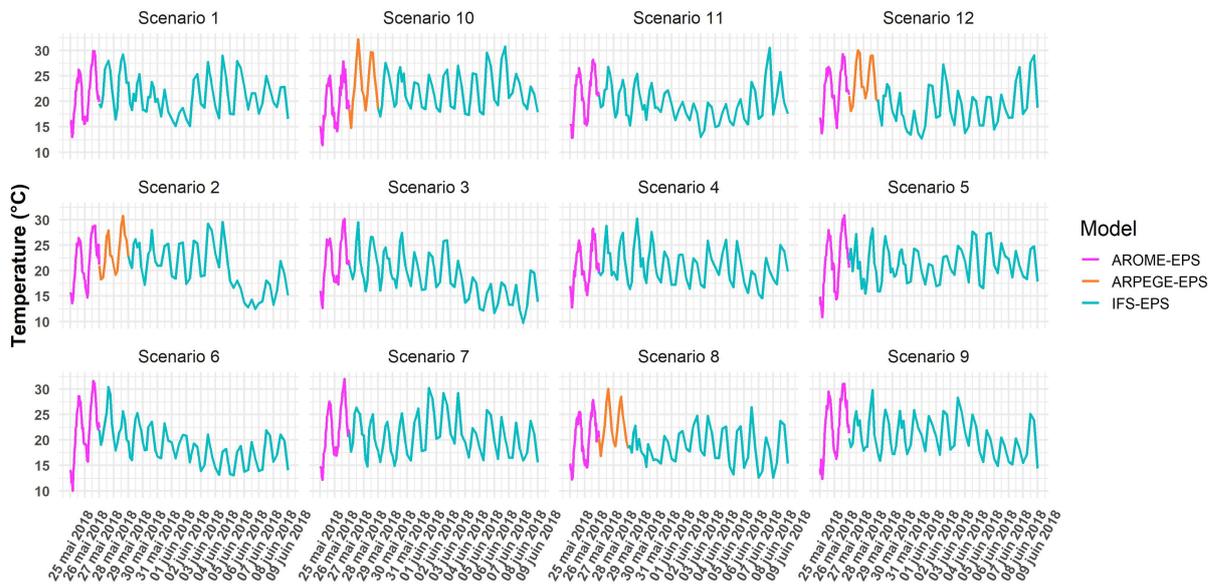
FIGURE 9.5 – Distributions (sous forme de boîtes à moustaches) des différences absolues de température prévue (moyennées sur tous les membres de l'ensemble) entre les échéances 48h et 49h (en noir) et les échéances 96h et 97h (en gris), calculées pour 11 sites sur la période 06/01/2018-31/05/2019. Ces distributions sont présentées, de gauche à droite, pour les membres IFS-EPS et pour les membres des ensembles cohérents RN, NN et HU.

9.2.5 Exemple représentatif d'un ensemble de scénarios cohérents

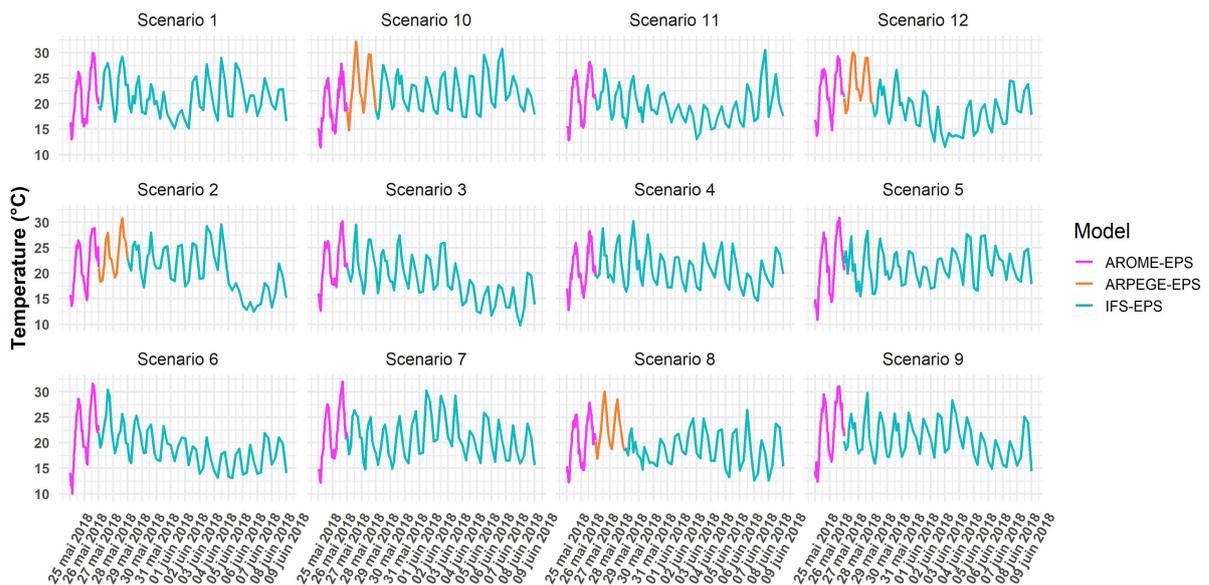
Une visualisation d'un ensemble de scénarios cohérents obtenu avec la stratégie HU est montré sur la Figure 9.6. On peut remarquer que le saut entre les trajectoires des différents modèles est minimisé au moment de la jointure. La structure cohérente et la dynamique des modèles sont préservées, et l'affectation hongroise permet d'assurer l'unicité des membres de cet ensemble.

9.3 La calibration : étape incontournable dans la construction des scénarios cohérents

L'étape de calibration des prévisions est une procédure coûteuse mais qui est apparue essentielle pour la construction des scénarios cohérents. Pour l'illustrer, nous présentons dans



(a)



(b)

FIGURE 9.6 – Ensemble de scénarios cohérents construit avec (a) la méthode HU, (b) la méthode NN. Date de début de la prévision 25/05/2018, à Turckheim. Chaque membre cohérent est initialisé par une prévision AROME-EPS (en rose) puis raccordé à un membre ARPEGE-EPS (en orange) ou IFS-EPS (en bleu) à 48h. Un deuxième raccordement à un membre IFS-EPS est effectué à 96h pour certains membres. Les distances entre membres sont calculées sur les 7 dernières heures de la période commune (de 42h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1)

la suite les résultats d'une analyse en composantes principales (ACP) des prévisions de température non corrigées, ainsi qu'un exemple de scénarios cohérents associés.

Exploration statistique des prévisions par ACP L'Analyse en Composantes Principales (ACP) est une procédure statistique qui utilise une transformation orthogonale en un ensemble d'observations de variables éventuellement corrélées dans un ensemble de valeurs de variables linéairement non corrélées appelées composantes principales (CP). Cette méthode fait partie du groupe des méthodes dites factorielles de décomposition sur une base adaptée (les facteurs sur lesquels projeter les données pour des représentations graphiques en dimension réduite) (Wold *et al.*, 1987). Il s'agit de considérer n individus, caractérisés par p variables initiales. Les données sont résumées en 2 dimensions, chaque individu est représenté par la projection du vecteur initial sur le plan défini par les CPs. Les individus sont les membres des 3 systèmes de prévision d'ensemble et le nombre d'individus est donc de $n = 98$ (12 membres AROME-EPS, 35 ARPEGE-EPS et 51 IFS-EPS) et les variables initiales sont des échéances, soit $p = 49$. Les 98 observations qui se présentent dans une dimension 49 sont ensuite projetées dans une nouvelle base de la même dimension p composée des vecteurs CP_1, CP_2, \dots, CP_p . La Figure 9.7(a) montre la projection des individus dans l'espace de la dimension 2. Les prévisions d'AROME-EPS et d'ARPEGE-EPS comportent des biais positifs quel que soit l'échéance par rapport aux prévisions de l'IFS-EPS. La Figure 9.7(b) montre le pourcentage de variabilité totale des données qui est exprimé sur les axes PC1 et PC2. Dans cet exemple, sur l'axe PC1, il y a plus de 70% de variabilité expliquée, ce qui signifie une bonne représentation.

Prévisions d'ensemble "sans couture" : cas pathologique La Figure 9.8(a) représente des prévisions d'ensemble "sans couture" obtenues avec la stratégie NN en utilisant les données brutes analysées dans la Figure 9.7. Il peut être remarqué que le saut de jointure à J+4 est très élevé. En outre, les figures 9.8(b) et (c) montrent les histogrammes des membres choisis pour la jointure à J+2 et J+4 respectivement. Le fort biais de l'IFS-EPS entraîne la sélection du membre le plus proche jusqu'à 10 fois pour la jointure. Cela réduit considérablement la taille effective de l'ensemble, ce qui n'est pas le cas avec les prédictions calibrées (Figure 9.6(b)).

9.4 Conclusion

Les méthodes de jointure de prévisions d'ensemble proposées dans le chapitre 5 ont été appliquées à la température à 2 mètres. Les résultats obtenus montrent que l'affectation hongroise permet de construire des prévisions d'ensemble cohérentes respectant les critères de continuité temporelle (minimisation des sauts de jointure) et d'indépendance des membres (pas de répétition). Par ailleurs, l'étape de calibration des prévisions est apparue comme un préliminaire incontournable à la jointure en raison des biais des différents systèmes.

Les méthodes de jointure proposées sont simples mais ont l'avantage d'être génériques et peuvent donc être appliquées à d'autres paramètres météorologiques, notamment les précipitations. Les précipitations présentent néanmoins une structure spatiale plus complexe que

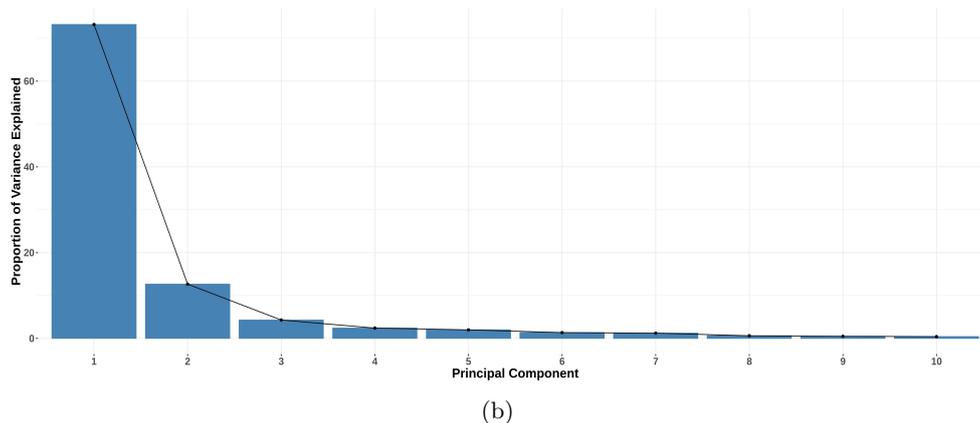
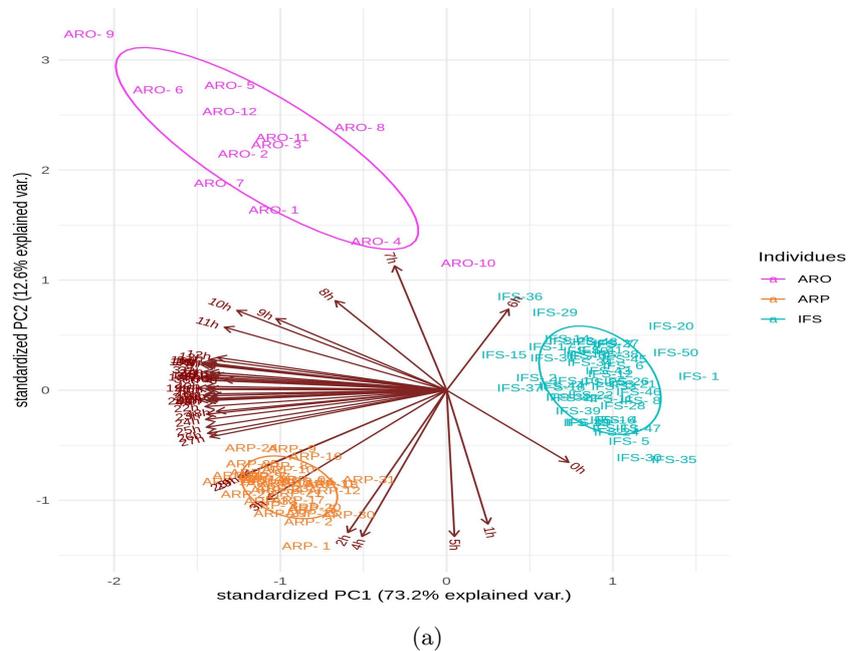
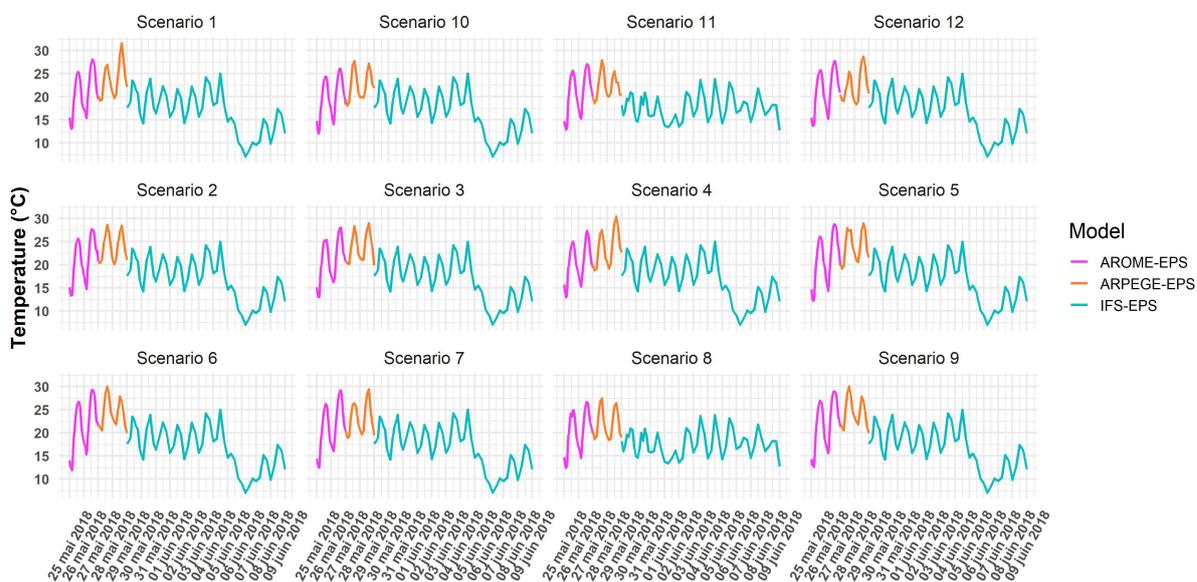


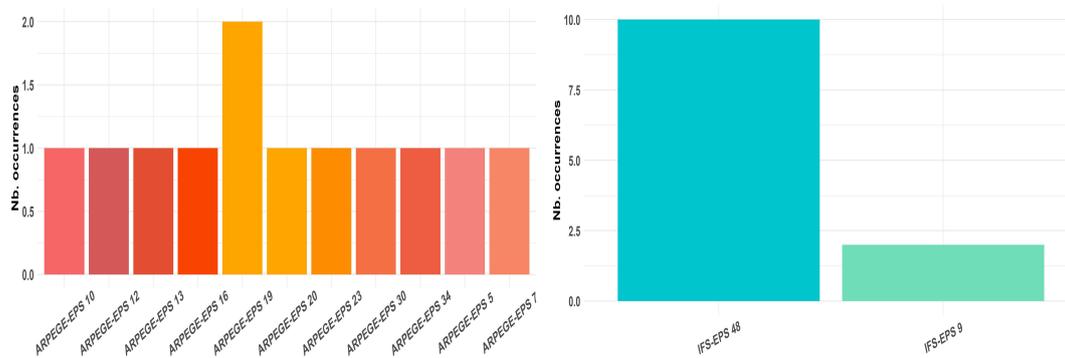
FIGURE 9.7 – Date de début de la prévision 25/05/2018 à Turckheim (a) ACP-Analyse en composante principale pour les 3 modèles AROME-EPS (ARO, en rose), ARPEGE-EPS (ARP, en orange) et IFS-EPS (IFS, en turquoise), (b) PVE - Pourcentage de la variance expliquée.

la température et le choix d'une distance adaptée sera un point important. Développer une procédure de jointure dans un cadre multivarié devra également être envisagé.

Dans le chapitre suivant, la performance de ces prévisions d'ensemble cohérentes est évaluée d'un point de vue météorologique et agronomique.



(a)



(b)

(c)

FIGURE 9.8 – (a) Ensemble de scénarios cohérents construit avec la méthode NN. Date de début de la prévision 25/05/2018, à Turckheim. Chaque membre cohérent est initialisé par une prévision AROME-EPS (en rose) puis raccordé à un membre ARPEGE-EPS (en orange) ou IFS-EPS (en bleu) à 48h. Un deuxième raccordement à un membre IFS-EPS est effectué à 96h pour certains membres. Les distances entre membres sont calculées sur les 7 dernières heures commune (de 42h à 48h) avec la distance dtw (fenêtre sakoechiba band, taille 1). (b) Histogrammes des membres choisis à J+2. (c) Histogrammes des membres choisis à J+4.

Évaluation météorologique et agronomique des prévisions d'ensemble cohérentes

Sommaire

10.1 Vérification météorologique des prévisions d'ensemble cohérentes . .	91
10.1.1 Jointure à 48h	91
10.1.2 Jointure à 24h	95
10.2 Vérification agronomique des prévisions d'ensemble cohérentes	95
10.3 Conclusion	102

Dans le chapitre précédent nous avons analysé les caractéristiques (cohérence temporelle, indépendance des membres) de trois prévisions d'ensemble cohérentes construites avec les approches hongroise, plus proche voisin et aléatoire. Dans cette partie nous nous intéressons aux performances de ces prévisions d'ensemble cohérentes, dans un cadre météorologique puis agronomique. Dans les deux cas le point de comparaison est donné par les performances de la prévision d'ensemble IFS-EPS. Les résultats présentés sont issus de l'article Aleksovska *et al.* 2020 soumis à *Weather and Forecasting* (voir Annexe D).

10.1 Vérification météorologique des prévisions d'ensemble cohérentes

Les performances probabilistes des ensembles obtenus avec les différentes stratégies de raccordement sont évaluées par rapport à la performance d'IFS-EPS. Comme ces ensembles ne comptent que 12 membres, les scores d'IFS-EPS sont calculés en utilisant un échantillon de 12 membres tirés aléatoirement parmi les 50 membres disponibles.

10.1.1 Jointure à 48h

Brier (performance globale). Le score de Brier pour l'évènement température supérieure au seuil $s_\epsilon = 15^\circ\text{C}$ est présenté sur la Figure 10.1. Ces résultats indiquent un impact positif de

CRPS								
2*Period	D-D+1		D+1-D+2		D+2-D+4			
	IFS-EPS	RN,NN,HU	IFS-EPS	RN,NN,HU	IFS-EPS	RN	NN	HU
summer 2018	0.800	0.732	0.835	0.849	0.923	0.943	0.975	0.954
autumn 2018	0.859	0.745	0.888	0.906	1.004	1.024	1.092	1.063
winter 2018/2019	0.883	0.786	0.916	0.916	1.034	1.031	1.102	1.077
spring 2019	0.773	0.735	0.827	0.886	0.928	0.935	0.982	0.959
summer 2019	0.810	0.764	0.869	0.903	0.965	0.970	0.995	0.977
autumn 2019	0.735	0.639	0.782	0.776	0.890	0.898	0.928	0.913

CRPS												
2*Period	D+4-D+7				D+7-D+10				D+10-D+15			
	IFS-EPS	RN	NN	HU	IFS-EPS	RN	NN	HU	IFS-EPS	RN	NN	HU
summer 2018	1.292	1.271	1.323	1.275	1.763	1.745	1.772	1.728	2.010	1.990	2.032	1.972
autumn 2018	1.456	1.437	1.489	1.469	1.975	1.940	1.992	1.903	2.393	2.339	2.433	2.284
winter 2018/2019	1.384	1.364	1.454	1.380	1.993	1.959	2.062	1.979	2.122	2.110	2.191	2.100
spring 2019	1.269	1.259	1.297	1.262	1.829	1.815	1.848	1.809	2.170	2.147	2.211	2.153
summer 2019	1.377	1.364	1.392	1.364	2.029	2.022	2.065	1.999	2.467	2.420	2.501	2.458
autumn 2019	1.247	1.223	1.259	1.228	1.661	1.638	1.678	1.643	1.923	1.901	1.991	1.886

TABLE 10.1 – CRPS moyen des différentes prévisions d’ensemble cohérentes et d’IFS-EPS pour différentes saisons et échéances (D-D+1 représente les échéances 0-24h, D+1-D+2 les échéances 24-48h etc.).

la jointure sur les 24 premières heures d’échéance, lorsque les prévisions AROME-EPS sont utilisées. Pour des échéances plus longues l’ensemble RN a un impact globalement neutre tandis que l’ensemble NN est significativement moins bon, en raison de sa taille effective plus petite. L’ensemble HU présente quant à lui des performances légèrement moins bonnes que celles d’IFS-EPS sur les échéances 2 à 4 jours, lorsque les prévisions ARPEGE-EPS sont utilisées.

CRPS (performance globale). Le CRPS présenté sur la Figure 10.2 fournit des conclusions similaires à celles du score de Brier. Sur les 24 premières heures de prévision, la jointure aux prévisions AROME-EPS conduit à une amélioration moyenne du CRPS de 9%. Pour des échéances plus lointaines, l’ensemble NN est significativement moins bon qu’IFS-EPS et que les ensembles RN et HU, en raison de sa taille effective plus réduite. L’ensemble RN obtient les meilleurs résultats pour les échéances 2 à 4 jours, probablement parce que les membres IFS-EPS, qui ont de meilleures performances que ceux d’ARPEGE-EPS pour ces échéances (Figure 9.2), sont plus nombreux avec l’affectation RN. Entre 5 et 15 jours d’échéance, les ensembles RN et HU ont des performances similaires, avec des valeurs de CRPS qui ne sont pas statistiquement différentes de celles d’IFS-EPS. Le tableau 10.1 montre les performances des prévisions cohérentes et celles d’IFS-EPS (calibré) résumées par saison et par intervalles d’échéance. Les mêmes conclusions sont observées, avec des CRPS très proches pour les méthodes RN et HU, et une moins bonne performance de la stratégie NN sur toute la période.

Courbe ROC et l’aire sous la courbe AUC (discrimination). La courbe ROC est utilisée pour évaluer la capacité d’un système de prévision à distinguer entre les événements observés et non observés. Sur la Figure 10.3 nous présentons les courbes ROC pour les différentes prévisions d’ensemble, calculée pour les échéances de 0 à 48h et 0 à 360h pour l’évènement température est supérieure à 15°C. Sur les premières 48h d’échéance les prévisions d’ensemble cohérentes conduisent à une légère amélioration de la courbe ROC. En moyenne

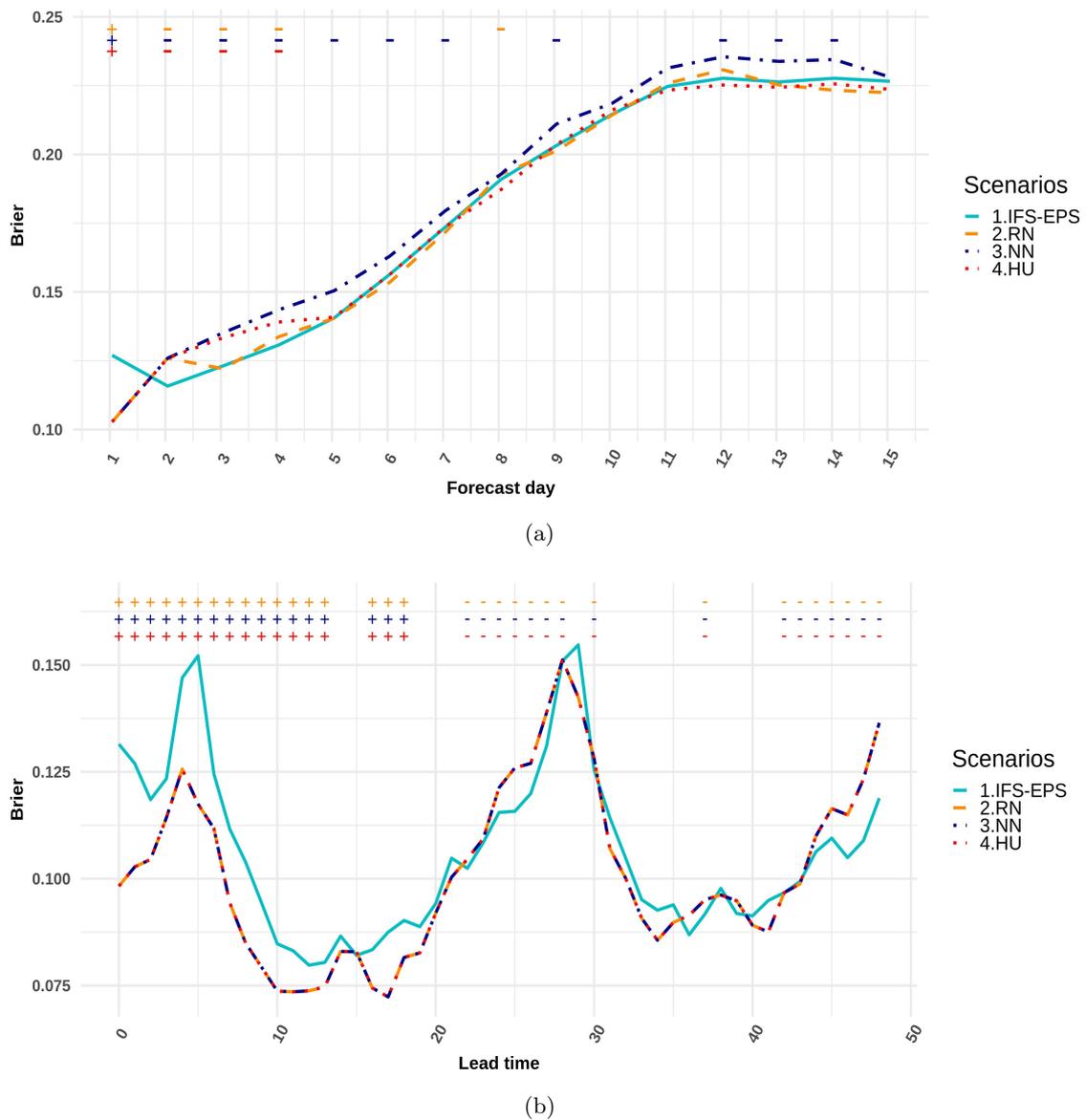
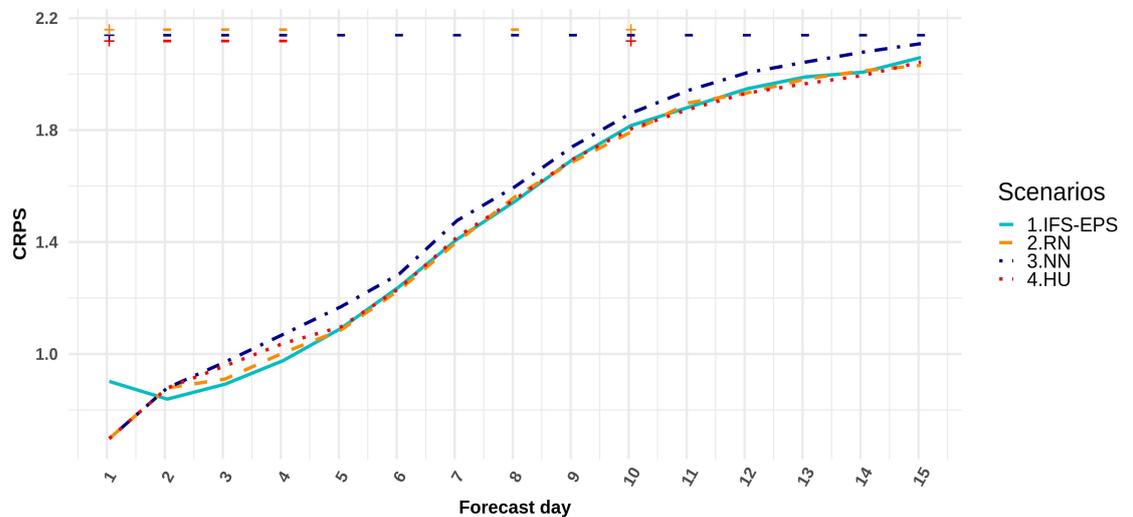
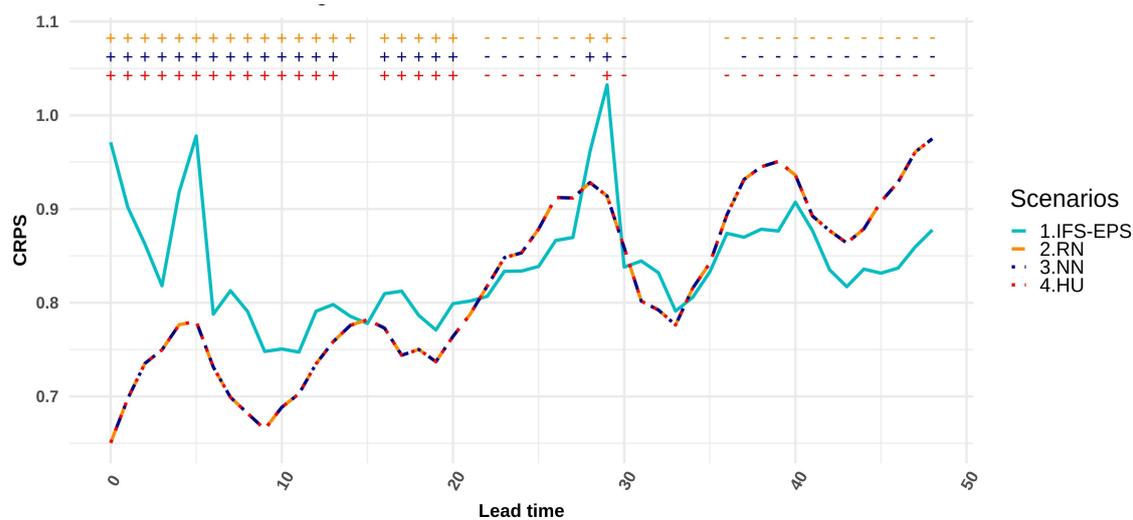


FIGURE 10.1 – Évolution du score de Brier des prévisions d'ensemble de température à 2 mètres en fonction des échéances, calculé pour 11 sites du tableau 7.1 sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision pour l'évènement $\epsilon := \mathbb{T} \geq 15^\circ\text{C}$. Les résultats d'IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) Une valeur par jour (valide à 24h, 48h, etc.), (b) Zoom sur les premières 48h d'échéance. Les signes + (resp. -) indiquent que la performance d'un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d'IFS-EPS, selon le test de Wilcoxon.



(a)



(b)

FIGURE 10.2 – Évolution du CRPS des prévisions d’ensemble de température à 2 mètres en fonction des échéances, calculé pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l’échéance de prévision. Les résultats d’IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) Une valeur par jour (valide à 24h, 48h, etc.), (b) Zoom sur les premières 48h d’échéance. Les signes + (resp. -) indiquent que la performance d’un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d’IFS-EPS, selon le test de Wilcoxon.

sur l'ensemble sur les 15 jours de prévisions les ensembles RN et HU ont des performances similaires à celles d'IFS-EPS tandis que l'ensemble NN est légèrement moins bon.

L'aire sous la courbe ROC (Figure 10.4) confirme la meilleure performance des prévisions cohérentes sur les deux premiers jours de prévision. Au-delà les ensembles RN et HU ont des performances similaires à celles d'IFS-EPS tandis que l'ensemble NN est moins bon.

Diagramme de fiabilité (fiabilité). La jointure améliore également la fiabilité des prévisions (Figure 10.5). Au cours des 48 premières heures, les ensembles cohérents améliorent la fiabilité des probabilités supérieures à 0,6. En moyenne sur les 15 jours de prévision, les ensembles RN et HU sont très proches d'IFS-EPS, tandis que la fiabilité de l'ensemble NN est un peu moins bonne, en particulier pour les probabilités faibles. Des résultats similaires sont obtenus avec les seuils 5 et 10°C.

10.1.2 Jointure à 24h

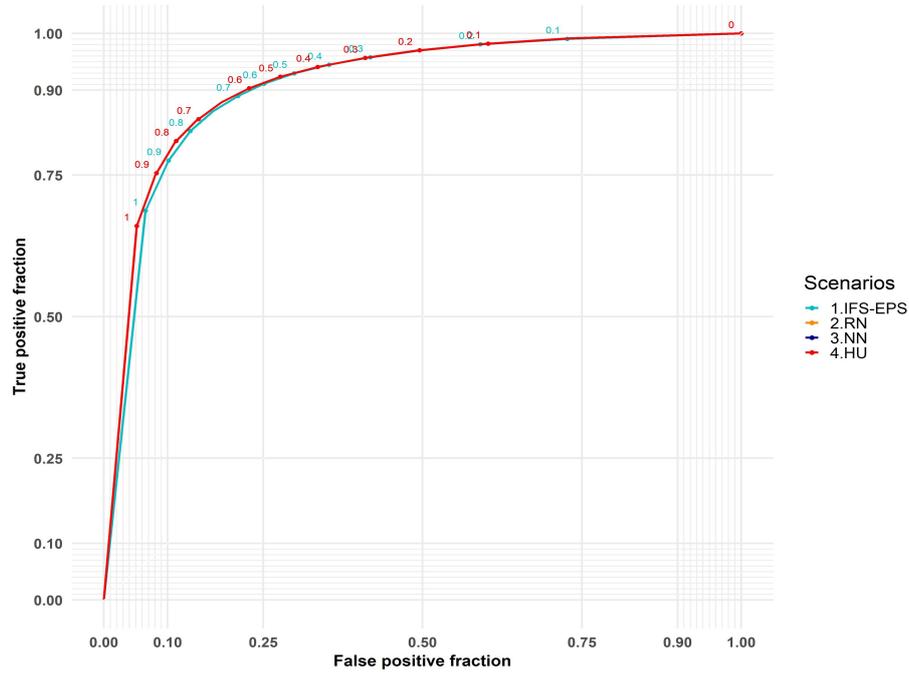
Les résultats précédents ont montré que l'apport des membres AROME-EPS se limite essentiellement aux premières 24 heures d'échéance. Une solution alternative pourrait donc être de décaler le premier raccordement à l'échéance 24h (au lieu de 48h). La Figure 10.6 indique que ce raccordement plus précoce conduit à des améliorations de CRPS notables de l'ensemble HU entre 2 et 4 jours d'échéance. Ces conclusions sont confirmées par l'aire sous la courbe ROC (Figure 10.7) et les diagrammes de fiabilité (Figure 10.8). Dans la suite du manuscrit les ensembles NN et HU utilisés sont obtenus avec la configuration suivante : premier raccordement à 24h et deuxième raccordement à 72h.

Le tableau 10.2 présente le pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le raccordement à 24h, avec les trois stratégies RN, NN et HU, pour différentes saisons. Les affectations basées sur la distance entre membres permettent une sélection plus équilibrée et certaines variations saisonnières sont observées. En automne, par exemple la répartition des membres ARPEGE-EPS obtenus avec la stratégie NN est significativement plus grande que celle des membres IFS-EPS. En revanche pour les autres saisons, le comportement similaire est observé comme dans le cas de la jointure à 48h montre sur le tableau 9.1.

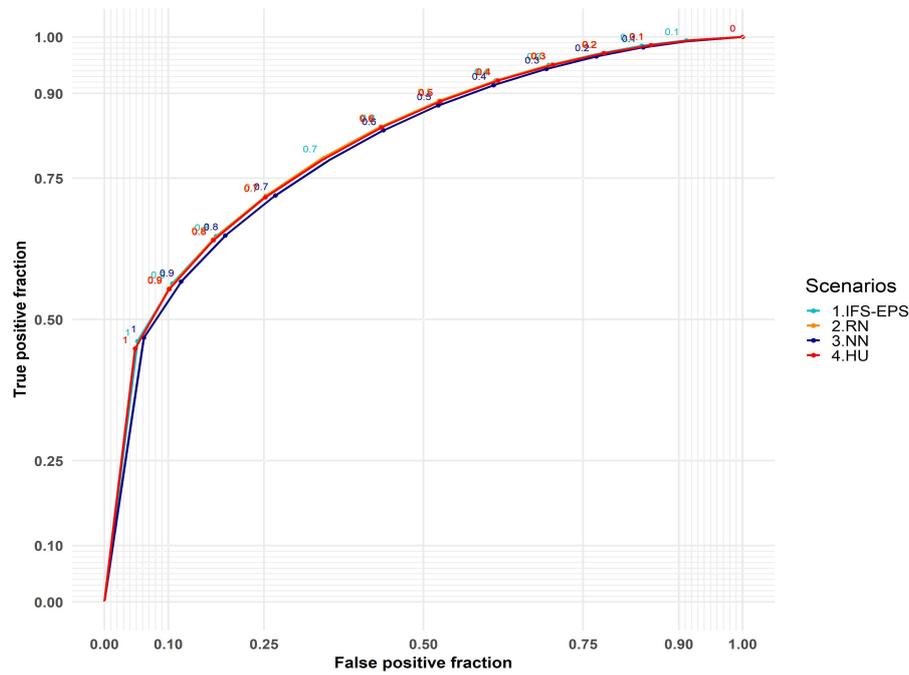
10.2 Vérification agronomique des prévisions d'ensemble cohérentes

Nous présentons dans cette partie l'impact des prévisions d'ensemble calibrées et cohérentes sur les dates de traitement prévues par le modèle EVA.

Biais (précision). Le biais de la date moyenne de traitement de l'ensemble est présenté sur la Figure 10.9. Pour les prévisions à court terme correspondant à des intervalles de temps



(a)



(b)

FIGURE 10.3 – Courbes ROC des prévisions d'ensemble de température à 2 mètres calculées pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l'échéance de prévision pour l'évènement $\epsilon := T \geq 15^\circ\text{C}$. Les résultats d'IFS-EPS calibrées et des trois ensembles de scénarios cohérents sont superposés. (a) ROC calculée sur les échéances 0 à 48h ; (b) ROC calculée sur les échéances 0 à 360h.

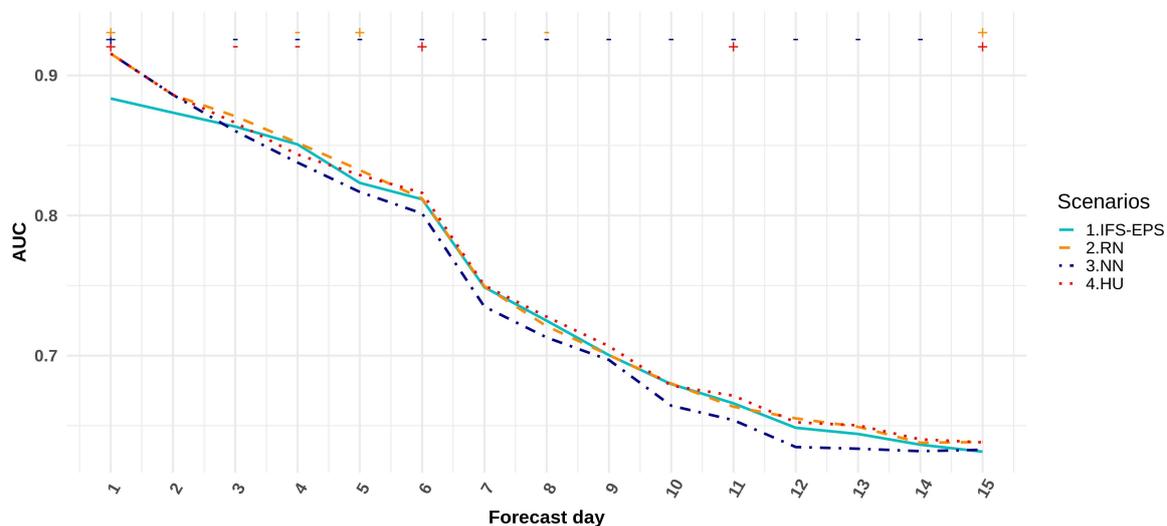
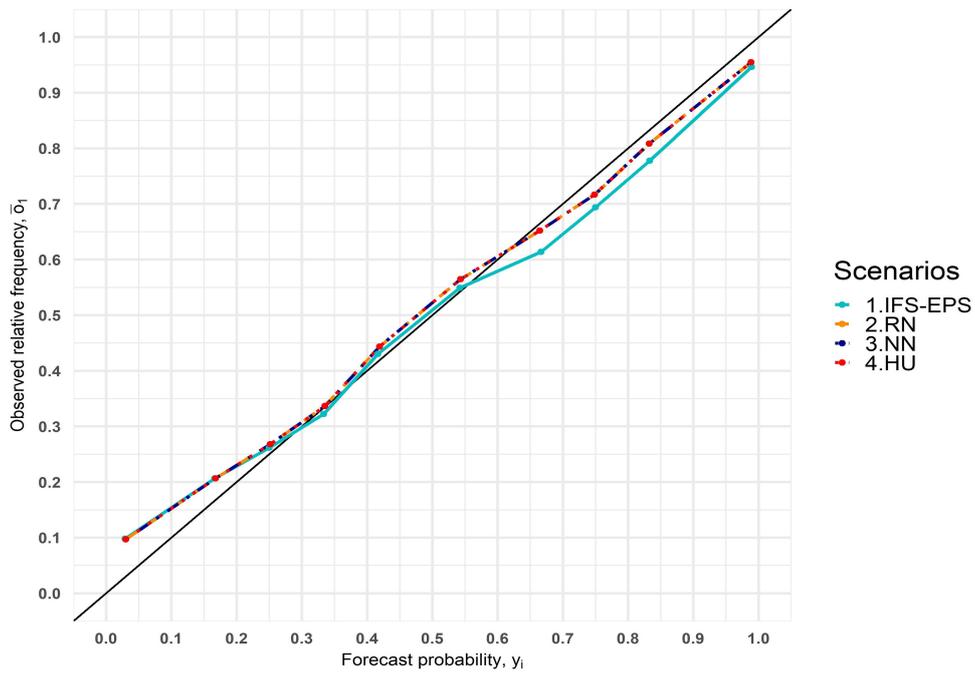


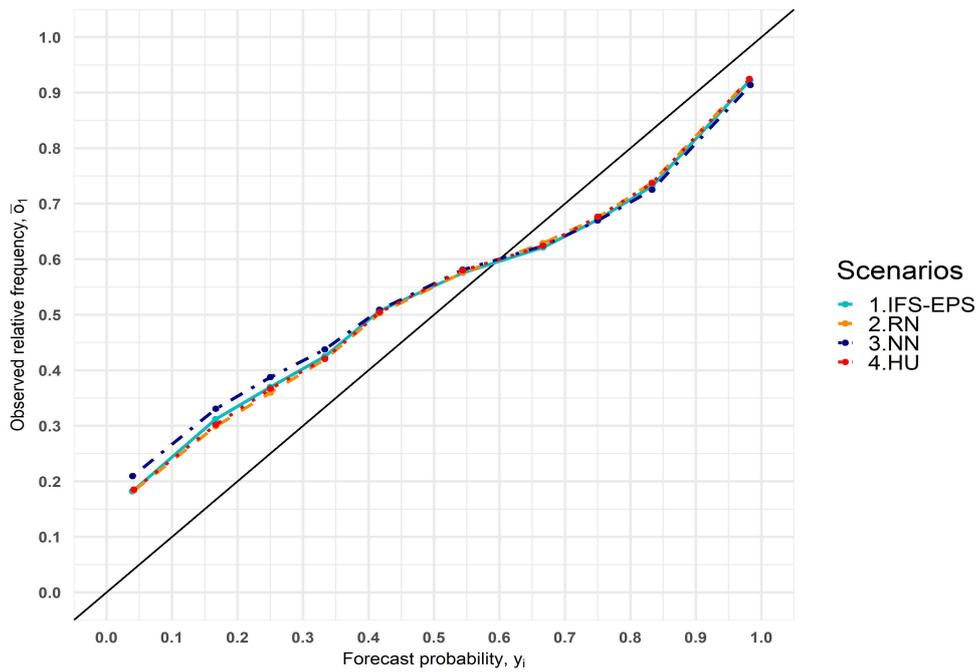
FIGURE 10.4 – Évolution de l’aire sous la courbe ROC (AUC) pour les prévisions de température à 2 mètres supérieures 15°C, calculée pour 11 sites sur la période 06/01/2018-31/05/2019, en fonction de l’échéance de prévision (une valeur par jour). Les résultats d’IFS-EPS et des trois ensembles de scénarios cohérents sont superposés. Les signes + (resp. -) indiquent que la performance d’un ensemble scénarios cohérents est statistiquement meilleure (resp. moins bonne) que la performance d’IFS-EPS, selon le test de Wilcoxon.

2*Time Period	ARPEGE-EPS			IFS-EPS		
	RN	NN	HU	RN	NN	HU
2018-06-01-2018-08-31	40,7 %	46.7 %	45.2 %	59,3 %	53.3 %	54.8 %
2018-09-01-2018-11-30	40,7 %	64.1 %	47.1 %	59,3 %	35.9 %	52.9 %
2018-12-01-2019-02-28	40,7 %	51.9 %	51.3 %	59,3 %	48.1 %	48.7 %
2019-03-01-2019-05-31	40,7 %	45.6 %	45.2 %	59,3 %	54.4 %	54.8 %

TABLE 10.2 – Pourcentage de membres ARPEGE-EPS et IFS-EPS sélectionnés pour le recordement à 24h avec les différentes stratégies d’affectation, pour quatre saisons. Les valeurs théoriques pour la sélection aléatoire sont de 40,7% pour ARPEGE-EPS et 59,3% pour IFS-EPS, en raison de leur taille respective.



(a)



(b)

FIGURE 10.5 – Diagramme de fiabilité pour les prévisions de température à 2 mètres supérieures à 15°C , calculé pour 11 sites sur la période 06/01/2018-31/05/2019. (a) calculé sur les échéances 0 à 48h, (b) calculé sur les échéances 0 à 360h. Les résultats d’IFS-EPS et des trois ensembles scénarios cohérents sont superposés.

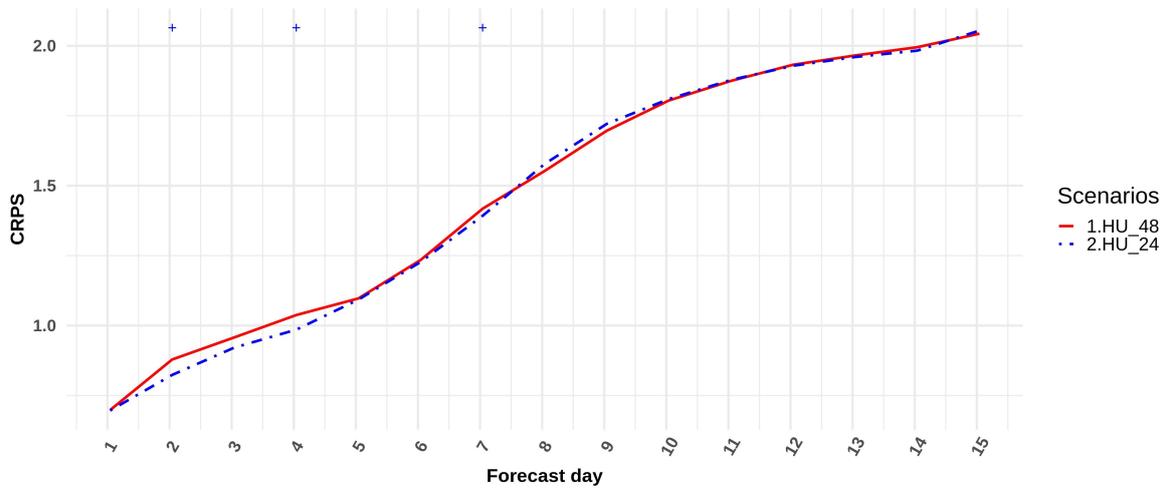


FIGURE 10.6 – Évolution du CRPS des prévisions de température à 2 mètres, calculé pour 11 sites sur la période 01/06/2018-31/05/2019, en fonction de l'échéance de prévision. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue). Les signes + (resp. -) indiquent que la performance du raccordement à 24h est statistiquement meilleure (resp. moins bonne) que celle du raccordement à 48h, selon le test de Wilcoxon.

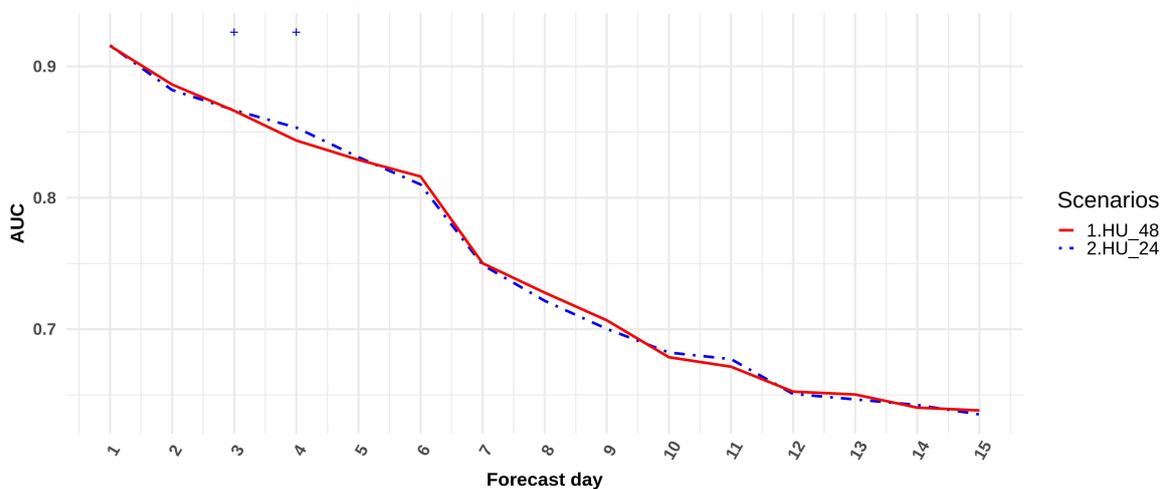
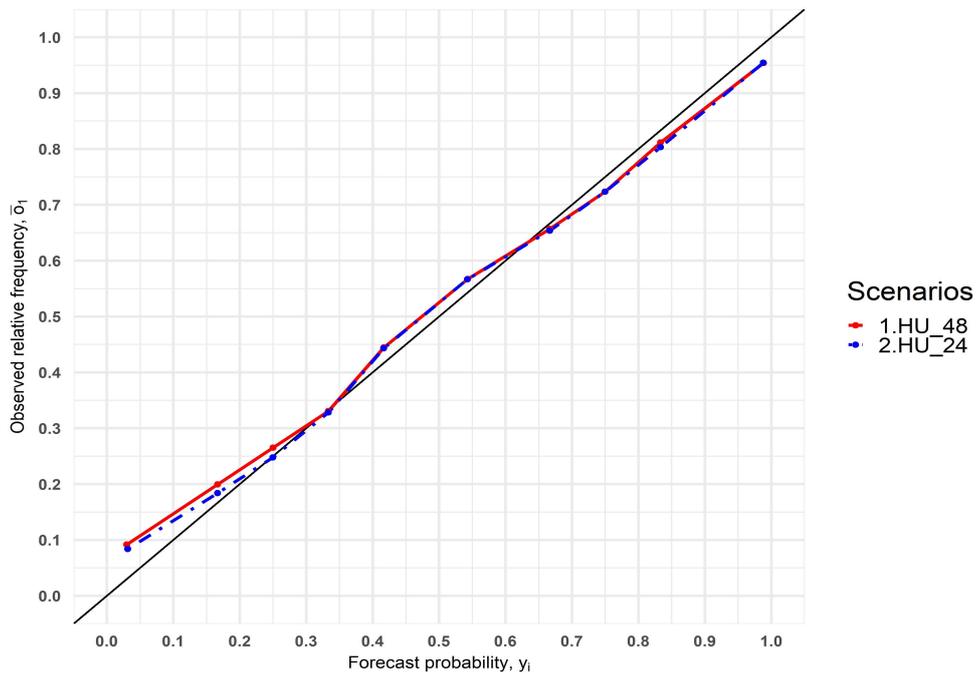
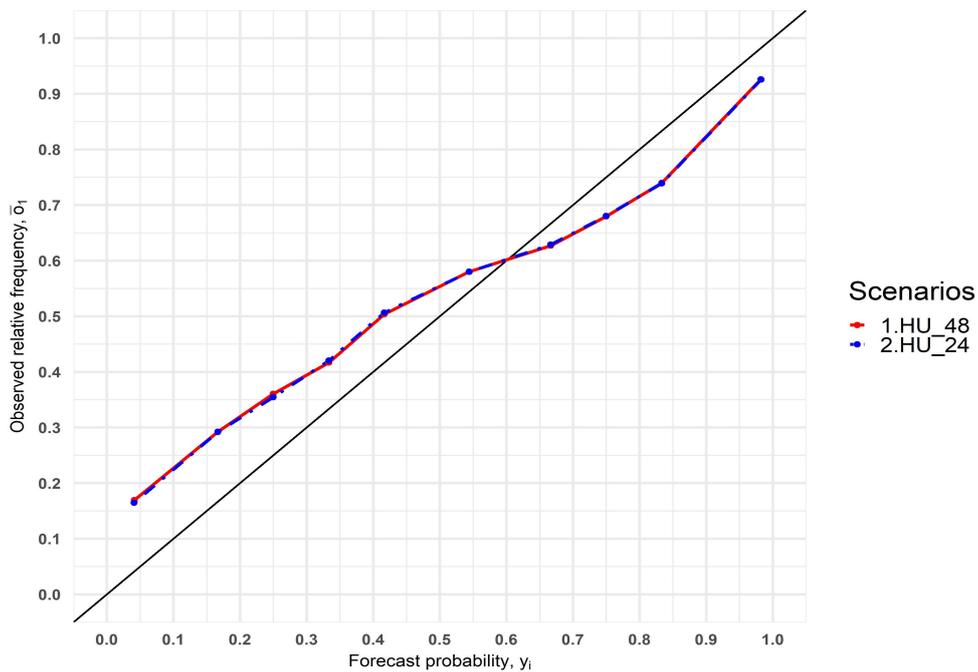


FIGURE 10.7 – Évolution de l'aire sous la courbe ROC (AUC) pour les prévisions de température de 2 mètres supérieures à 15°C, calculée pour les 11 sites sur la période 06/01/2018-31/05/2019, en fonction du jour de début de la prévision. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue). Les signes + (resp. -) indiquent que la performance de la fusion sur 24h est statistiquement meilleure (resp. pire) que la performance du IFS-EPS, selon le test de Wilcoxon.



(a)



(b)

FIGURE 10.8 – Diagramme de fiabilité pour les prévisions de température à 2 mètres au-dessus de 15°C, calculé pour les 11 sites sur la période 06/01/2018-31/05/2019 pour les jointure à 24h et 48h. (a) échéances 0-48h, (b) échéances 0-360h. Les résultats sont donnés pour un ensemble HU avec un premier raccordement à 24h (ligne pointillée) et à 48h (ligne continue).

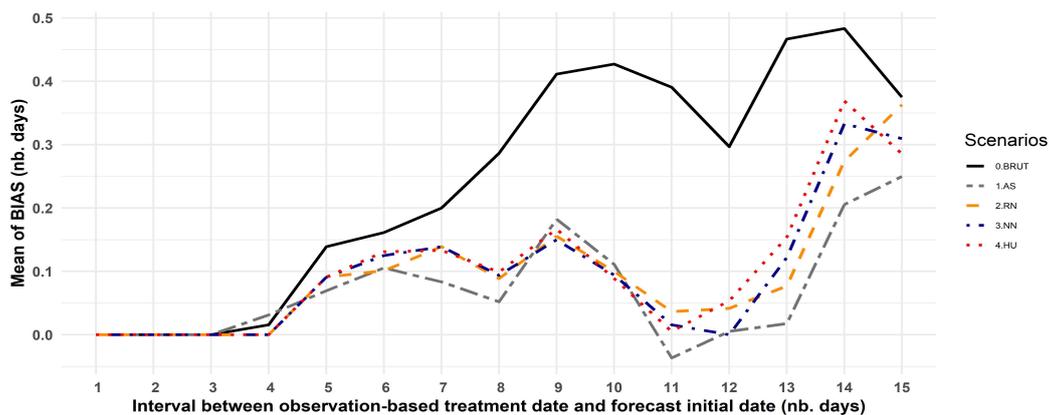


FIGURE 10.9 – Biais de la date moyenne de traitement prévue par l’ensemble de simulations EVA, en fonction de l’intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les scores sont moyennés sur les 8 sites et les années 2018 et 2019. Les biais moyens sont calculés à partir de simulations d’EVA d’ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d’ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).

compris entre 1 et 3 jours, le biais est nul quelle que soit la prévision de température utilisée. Pour les intervalles de temps entre 4 et 7 jours, le biais augmente légèrement vers des valeurs positives (ce qui signifie que la date de traitement est légèrement trop tardive) et il y a un petit avantage à utiliser des prévisions de température calibrées. Ce biais positif est en accord avec le biais de température négatif des prévisions brutes de l’IFS-EPS, comme le montre la Figure 9.1. L’impact positif de la calibration devient plus important pour les prévisions entre une et deux semaines avant la date de traitement, avec des valeurs de biais qui restent proches de zéro. Les scénarios cohérents permettent de prolonger le biais nul jusqu’à 4 jours. Au-delà leur performance est proche ou légèrement moins bonne que celle des prévisions IFS-EPS calibrées.

CRPS (performance globale). L’impact de la procédure de calibration est également évident sur le CRPS (Figure 10.10), qui est significativement meilleur pour les prévisions de la deuxième semaine de près de 40%. En revanche, il est plus difficile de tirer des conclusions sur la valeur et le classement des prévisions cohérentes, par rapport à l’IFS-EPS calibré.

Couverture des observations (précision). La Figure 10.11 montre le pourcentage de dates de traitement observées se situant dans l’intervalle des dates prévues par l’ensemble. Conformément aux résultats précédents, la prévision est presque parfaite pour les échéances courtes et elle est peu sensible à la prévision météorologique utilisée jusqu’à une semaine d’échéance. Pour des échéances plus longues, la couverture des observations diminue de manière significative jusqu’à moins de 70 %, ce qui est inférieur à la valeur théorique de 85 % pour un modèle parfait et un ensemble de 12 membres, tandis que les prévisions calibrées et

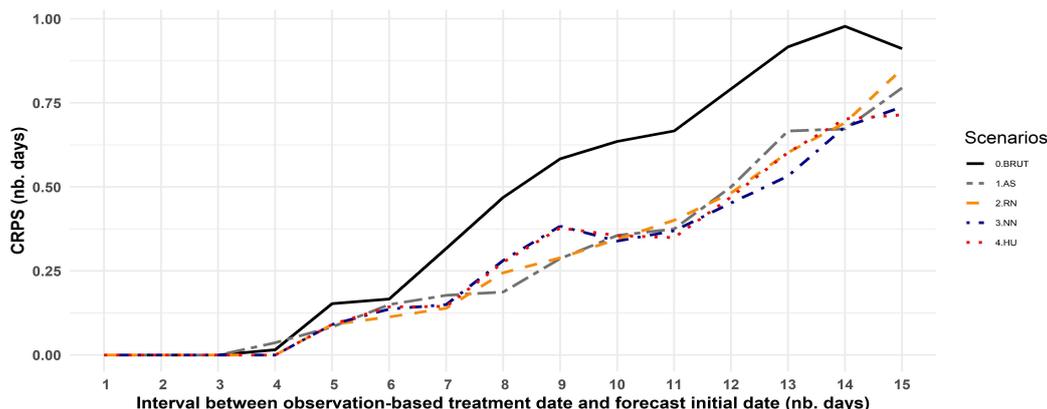


FIGURE 10.10 – CRPS des dates de traitement prévues par l’ensemble de simulations EVA, en fonction de l’intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les scores sont calculés en moyenne sur les 8 sites et les années 2018 et 2019. Les CRPS moyens sont calculés à partir de simulations d’EVA d’ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d’ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).

les ensembles de scénarios cohérents permettent de maintenir des valeurs autour de 95 % ou plus jusqu’à 11 jours. La couverture des observations sur les échéances 7 à 15 jours est légèrement plus importante pour les prévisions HU, avec une valeur moyenne de 94,5%, par rapport aux valeurs comprises entre 91 et 92% pour les prévisions calibrées et les ensembles RN et NN.

10.3 Conclusion

L’évaluation des stratégies de construction des prévisions d’ensemble cohérentes a été réalisée sur la base de trois critères principaux : la continuité temporelle des prévisions, l’indépendance des membres et les performances probabilistes. Les résultats de ce chapitre et du précédent indiquent que, d’un point de vue météorologique, l’ensemble HU est le seul à satisfaire tous ces critères d’évaluation. En particulier ses performances sont supérieures (sur les premières 24h) ou égales (au-delà de 24h) à celles d’IFS-EPS.

D’un point de vue agronomique, l’impact des deux post-traitements, calibration et jointure, des prévisions météorologiques est plutôt limité jusqu’à une semaine avant la date du traitement. Ce comportement pourrait être spécifique au modèle EVA, qui n’est peut-être pas le choix le plus approprié pour mettre en évidence le potentiel des prévisions à haute résolution aux courtes échéances. Il est possible que les règles d’évolution basées sur l’accumulation des températures quotidiennes et les effets de seuil contribuent à réduire l’impact de prévisions plus précises à court terme. En revanche, le bénéfice de la calibration est clairement visible entre 7 et 15 jours d’échéance, ce qui pourrait aider les agriculteurs à mieux anticiper la pré-

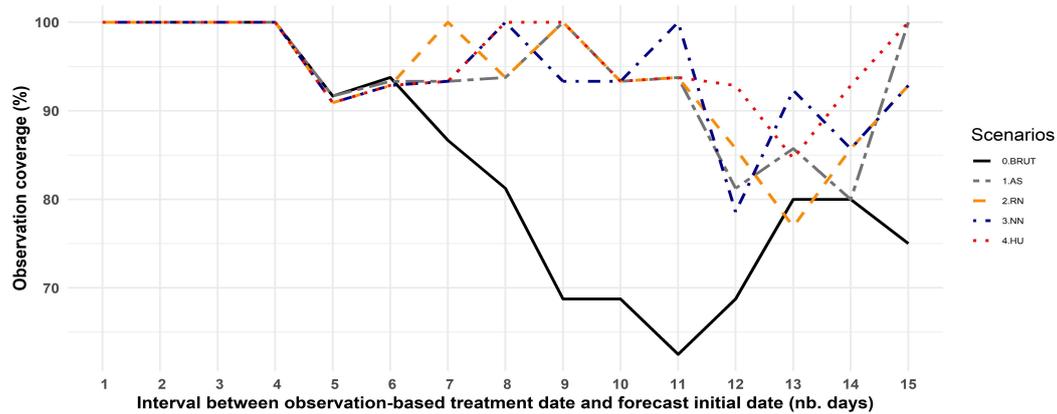


FIGURE 10.11 – Évolution de la couverture des observations, en fonction de l’intervalle de temps (en jours) entre la date observée et la date de début de la prévision. Les pourcentages de couverture d’observation sont calculés en moyenne sur les 8 sites et les années 2018 et 2019. Les couvertures des observations sont calculés à partir de simulations d’EVA d’ensemble utilisant les prévisions IFS-EPS (0.BRUT, en noir), les prévisions calibrées (1.AS, en gris) et les prévisions d’ensemble cohérentes obtenues avec le tirage aléatoire (2.RN, en jaune), le plus proche voisin (3.NN, en bleu) et la stratégie hongroise (4.HU, en rouge).

paration des traitements. L’apport des prévisions cohérentes intégrant des simulations haute résolution à courte échéance devra être examiné plus en détail, en considérant notamment un échantillon d’évaluation plus important, et éventuellement un modèle agronomique plus sensible aux variations météorologiques à courte échéance.

Conclusion, discussion et perspectives

Conclusion

Le monde agricole est fortement dépendant des conditions météorologiques passées et futures qui impactent les productions et leur gestion par les agriculteurs. L'objectif de mon travail était d'examiner le potentiel de la prévision météorologique probabiliste, et plus précisément des prévisions d'ensemble, dans les outils d'aide à la décision en protection des cultures dans l'objectif d'optimiser le positionnement des applications de traitements phytosanitaires. Ce travail a été décomposé en trois étapes, dont nous rappelons les principaux résultats.

Q1 : Quel est l'apport et le potentiel de la prévision d'ensemble sur la quantification de l'incertitude associée à la prévision de la date d'un traitement phytosanitaire par rapport aux approches existantes ?

Les incertitudes météorologiques sont actuellement prises en compte en utilisant plusieurs séries d'observations passées, dans une approche dite "fréquentielle". Cette approche permet de représenter la variabilité climatique inter-annuelle, mais cette dernière peut être assez différente de l'incertitude "du jour", prenant en compte la situation atmosphérique du moment. Nous avons montré que l'incertitude des données fréquentielles est généralement plus importante que celle des prévisions d'ensemble, en particulier aux courtes échéances. Pour les simulations agronomiques, cela se traduit par une incertitude plus grande des dates de traitement prévues avec les données fréquentielles. Cela se vérifie pour le cas d'étude du ver de la grappe en vigne dépendant uniquement de la température horaire ou pour la maladie de la septoriose du blé qui est pilotée par les températures minimales et maximales quotidiennes et les précipitations. Les scores de performance calculés sur les années 2018 et 2019 ont par ailleurs montré un apport significatif de la prévision d'ensemble, avec en particulier une forte réduction du biais pour des échéances d'une à deux semaines avant le traitement.

Q2 : Comment faire la jointure entre plusieurs systèmes de prévision d'ensemble couvrant des échéances spatio-temporelles différentes ?

En protection des cultures, l'anticipation des traitements est essentielle et nécessite de disposer de séries temporelles de prévisions météorologiques sur des échéances longues, ce qui rend l'utilisation des prévisions IFS-EPS pertinente. Nous avons néanmoins montré que les prévisions d'ensemble AROME-EPS et ARPEGE-EPS, utilisant des résolutions spatiales plus élevées qu'IFS-EPS mais sur des échéances plus courtes, peuvent fournir un complément d'information intéressant. Nous avons donc proposé plusieurs stratégies afin de construire des prévisions "sans couture" combinant ces trois systèmes de prévision d'ensemble. Après une étape préalable de calibration des prévisions d'ensemble, la jointure a consisté à déterminer les meilleurs raccordements entre membres, en s'appuyant sur une mesure de distance et des

algorithmes d'affectation. L'application s'est concentrée sur les prévisions de température, et nous avons alors montré que la stratégie "hongroise" est la seule permettant de répondre au besoin de continuité temporelle et d'unicité des prévisions d'ensemble.

Q3 : Quelle est la qualité et l'impact de la prévision d'ensemble "sans couture" d'un point de vue météorologique et de son impact sur les prévisions des OAD ?

L'évaluation météorologique des prévisions d'ensemble "sans couture" indique une amélioration significative sur les premières 24h d'échéance (par rapport à IFS-EPS), lorsque les prévisions AROME-EPS sont utilisées. Dans l'application agronomique EVA, qui simule le développement du ver de la grappe en vigne, la calibration des prévisions de température a un impact positif significatif sur les distributions de dates prévues, en particulier pour des échéances de une à deux semaines avant le traitement. Cependant, le cas d'étude EVA ne semble pas assez sensible aux données météorologiques à courte échéance pour conclure sur l'apport des prévisions "sans couture".

Discussion et perspectives

Discussion

Dans le cadre de l'évaluation agronomique, les prévisions d'ensemble cohérentes n'ont pas montré de différences significatives par rapport aux prévisions calibrées IFS-EPS. Il est possible que le modèle EVA ne soit pas le choix le plus approprié pour l'évaluation des impacts des modèles à échéances courtes et donc des scénarios cohérents, en raison des processus de latence, qui impliquent un décalage de la variabilité aux courtes échéances sur le stade de cumul d'œuf considéré. Ce stade dépend également de la capacité des femelles imagos à pondre des œufs, et donc d'un coefficient de ponte lié à leur âge. Afin de mieux évaluer l'impact des stratégies de jointure, il sera intéressant de considérer d'autres modèles agronomiques dépendant de la température, comme les modèles de stade de végétation (Canal, 2014). Il sera également intéressant d'évaluer les modèles qui dépendent de variables moins prévisibles (par exemple, les précipitations) et qui sont plus sensibles à l'incertitude météorologique à court terme (le modèle SeptoLIS ou les modèles de bilan hydrique comme WaLIS en vigne (Roux *et al.*, 2014) par exemple).

La méthode de jointure proposée est générique et pourrait facilement être appliquée à d'autres paramètres et d'autres échéances. Sur les échéances longues, en particulier, il serait assez intéressant d'évaluer l'intérêt de cette méthode pour introduire un raccordement supplémentaire aux prévisions saisonnières (White *et al.*, 2017) pour une gestion à plus long terme des pratiques agricoles (Canal, 2014). Cette méthode pourrait également bénéficier à d'autres applications pour lesquelles il est important de disposer de prévisions cohérentes dans le temps, notamment l'hydrologie et la production d'énergie.

Nous pouvons par ailleurs envisager certaines améliorations à la méthode de jointure proposée. Par exemple, l'apport du raccordement à ARPEGE-EPS n'est pas totalement évident,

malgré une résolution spatiale plus fine qu'IFS-EPS. Il devra être réévalué avec la future version opérationnelle d'ARPEGE-EPS à 5km de résolution (prévue en 2021). Une autre piste d'amélioration consisterait à modifier l'initialisation actuelle qui n'utilise que les membres AROME-EPS. Une possibilité serait d'effectuer une classification de tous les membres disponibles en N classes (N étant la taille de l'ensemble cohérent souhaité) et de sélectionner un membre par classe. Enfin, il serait également possible de construire la prévision cohérente non plus comme un raccordement de membres mais comme une agrégation d'ensembles (Baudin, 2015 ; Thorey, 2017 ; Zamo, 2016).

Pour ce travail les incertitudes liées au modèle agronomique n'ont pas été prises en compte. Cependant, l'incertitude de la modélisation agronomique peut également jouer un rôle important. Par exemple, les paramètres du modèle EVA ont été fixés aux valeurs choisies par les experts, mais il existe une incertitude sur ces valeurs, avec quelques informations sur la gamme de valeurs possibles. Sur ce point, des simulations d'ensemble qui combinent à la fois l'incertitude des prévisions météorologiques et l'incertitude du modèle ont été effectuées dans trois stages menés en parallèle de la thèse. Une première analyse préliminaire a été faite dans le cadre de modèle EVA, sur l'analyse de variabilité des sorties agronomiques dans différents contextes climatiques. Un travail d'analyse de sensibilité aux paramètres du modèle et aux paramètres météorologiques a ensuite été effectué dans le cas du modèle SeptoLIS en montrant en particulier l'interaction entre les deux sources d'incertitude (Voir l'annexe E.1). Le dernier stage a porté sur l'étude de la propagation de l'incertitude des prévisions d'ensemble dans des modèles de bilan hydrique pour la vigne et le maïs (Voir l'annexe E.2).

Enfin, toutes les évaluations agronomiques effectuées ont utilisé une date de traitement de référence simulée, et donc potentiellement entâchée d'erreur de modélisation. Cette erreur a été négligée, mais devrait être prise en compte dans le calcul des scores.

Perspectives

L'extension de la construction des prévisions d'ensemble "sans couture" à un contexte multivarié est une étape future importante, car la plupart des applications agronomiques nécessitent plusieurs variables d'entrée météorologiques. Il existe au moins deux façons de procéder, soit en utilisant les stratégies de jointure univariées actuelles sur la variable météorologique la plus importante pour l'application, soit en utilisant une mesure de distance multivariée, telle que la distance de Mahalanobis, pour le raccordement. Cet aspect fera l'objet de travaux futurs.

Une autre question importante est de savoir s'il est plus efficace de calibrer les prévisions de température d'entrée ou plutôt de calibrer directement le produit final. Dans le cas présent, la calibration de la température est l'option privilégiée car elle facilite la construction des prévisions d'ensemble cohérentes en réduisant les discontinuités de raccordement. Par ailleurs, la calibration des dates de traitement est, à notre connaissance, un domaine inexploré qui peut présenter plusieurs défis. Par exemple, il n'y a aucune garantie que les dates de traitement suivent une distribution paramétrique simple et l'échantillon historique de dates observées

disponible pour la calibration est plutôt petit car il n’y a que 3 ou 4 générations du ravageur de la vigne par an.

Les résultats d’analyse de sensibilité ont montré l’importance de prendre en compte l’incertitude des conditions météorologiques et du modèle agronomique. Dans les travaux futurs le développement de simulations d’ensemble intégrant ces deux sources d’incertitude pourra être envisagé.

Pour terminer, ces travaux de la thèse pourront être valorisés auprès de la communauté agricole. Nous mentionnons que des études pour des plateformes de démonstration sont en cours afin de montrer le potentiel des simulations probabilistes (projet METEOPREC, 2019-2021). Un travail important d’ergonomie restera à faire pour proposer ces informations probabilistes sous une forme utilisable par les agriculteurs ou leurs conseillers. Par ailleurs, l’outil opérationnel d’aide à la décision Taméo (voir Annexe A) utilise actuellement les prévisions déterministes mais pourrait être étendu pour intégrer les travaux de la thèse et fournir en complément des prévisions probabilistes. En général, l’accueil de ces travaux par les communautés météorologique et agronomique a été positif. Les travaux de la thèse ont répondu à des questions importantes et ont ouvert des pistes de valorisation concrètes ainsi que différentes collaborations.

Plateformes disponibles actuellement utilisées incluant les deux cas d'étude de la thèse

A.1 EPICURE

Le Système d'Information EPICure proposé par l'Institut Français de la vigne (IFV) s'appuie sur différents outils pour accompagner au mieux les viticulteurs et leurs conseillers dans la prise de décision. Différents modèles de prévision des risques épidémiques en viticulture dont le modèle EVA y sont diffusés en s'appuyant actuellement sur données météo issues de stations météo et spatialisées (produit Antilope pour les données de pluie issues des radars). Des prévisions déterministes sont utilisées, ainsi que des scénarios extrêmes issus des scénarios probabilistes (approche quantile). Cette plateforme comporte également un système de collecte des observations des maladies et ravageurs des réseaux de parcelles sur plusieurs vignobles français, permettant une validation continue des modèles.

A.2 Taméo

Cette plateforme opérée conjointement par Arvalis et Météo-France intègre différents modèles pour prédire le stade de culture, le risque d'apparition des maladies, dont la septoriose avec le modèle Septolis, et repère les meilleures périodes d'intervention pour fertiliser, désherber et protéger la culture. Proposée tout d'abord sur la culture du blé tendre d'hiver, elle vise à intégrer d'autres espèces. La partie observation, selon le paramètre considéré, est basée soit directement sur des analyses du modèle AROME ou des flux radiatifs satellite soit sur la fusion de données entre observations aux stations et analyse modèle ou lame d'eau radar. Ces données sont disponibles à l'échelle de la parcelle agricole (précision kilométrique). La partie prévision utilise les prévisions déterministes (post-traitées et/ou expertisées par les prévisionnistes) issues des modèles AROME, ARPEGE et IFS. Les prévisions sont utilisées pour la prédiction de la date de traitement. L'aspect probabiliste n'apparaît qu'au travers de l'affichage des probabilités, de gel et de pluie, aux pas de temps quotidien et horaire, pour mieux prendre en compte les conditions d'application des traitements et notamment les aspects réglementaires (périodes d'application autorisées).

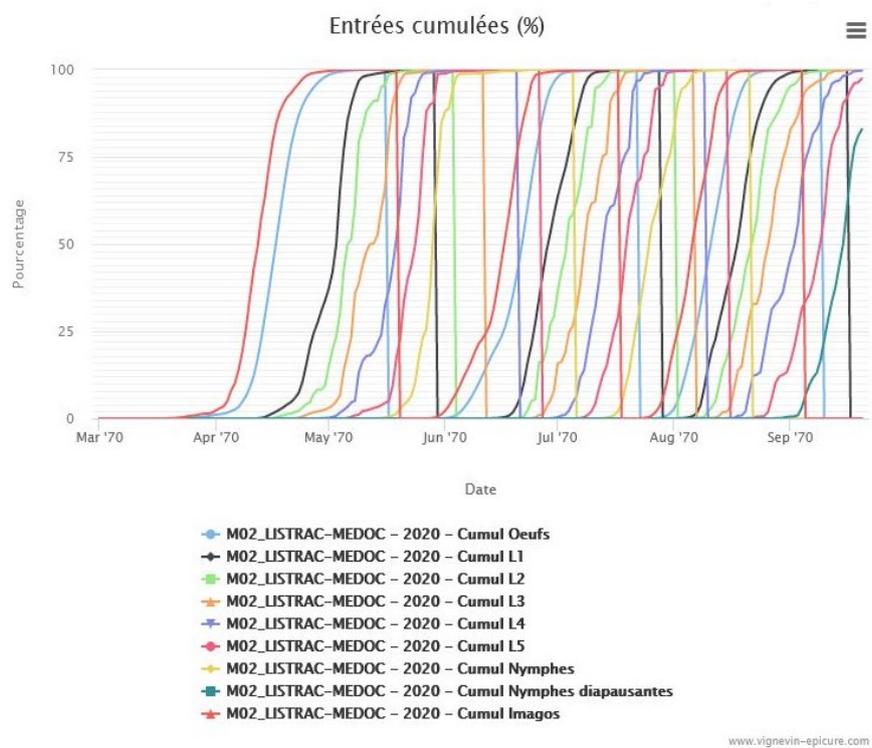


FIGURE A.1 – Exemple des sorties de EPICURE



MALADIES

Pour intervenir au bon moment

Visualisez l'évolution des maladies sur chacune de vos parcelles et optimisez leur protection



Blé tendre

Pilotez les 5 maladies principales : piétin verse, rouille jaune et brune, septoriose (premier traitement et renouvellement) et fusariose

PARCELLES	PIÉ TIN VERSE	ROUILLE JAUNE	SEPTORIOSE		ROUILLE BRUNE	FUSARIOSE
			1 ^{er} Traitement	Renouvellement		
CP Arvas - BERGAMO 23/10/2018	⊗	✓	✓	⊗	✓	⊗
CP Ekass - ARLEQUIN 19/10/2018	⊗	✓	✓	✓	✓	⊗
CP Ev Ombade - ASSOTT 18/10/2018	⊗	✓	13 avril	⊗	✓	⊗
CP Etale /Rhône - RUBINO 28/10/2018	⊗	✓	29 avril	⊗	✓	⊗
CP Home - FRUCTOOR 18/10/2018	⊗	✓	02 Mai	⊗	17 Mai	⊗
CP Petrelles Haies - AREZZO 28/10/2018	⊗	✓	28 avril	⊗	14 Mai	⊗
CP Saurur - CALABRO 24/10/2018	⊗	✓	21 avril	⊗	✓	⊗



Blé dur

MALADIE	03/04	04/04	AUGOURYER	06/04	07/04	08/04	09/04	10/04
Rouille brune	✓	✓	✓	✓	✓	✓	✓	✓
Fusariose	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗
STADES	1 noeur		2 noeurs					
Pluie hebdoA vent	5,2 mm	0,0 mm	0,0 / 0,0 mm	3,3 mm	2,4 mm	0,0 mm	1,7 mm	1,9 mm

- Identifiez les maladies du moment
- Classez les parcelles les plus à risque
- Anticipez les chantiers grâce aux prévisions de date de traitement
- Suivez le risque hebdomadaire et quotidien
- Prévoyez vos stocks en fonction du risque agronomique

FIGURE A.2 – Exemple des sorties de Taméo

**Article vulgarisé - Végéphyll - 12e
Conférence Internationale sur les
Maladies des Plantes (apparu en
décembre 2018)**



Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures

Ivana Aleksovska, François Brun, Laure Raynaud, Robert Faivre, Marc
Raynal, Olivier Deudon

► **To cite this version:**

Ivana Aleksovska, François Brun, Laure Raynaud, Robert Faivre, Marc Raynal, et al.. Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures. 12. Conférence internationale sur les maladies des plantes (Végéphy), Dec 2018, Tours, France. 10 p. hal-01958293

HAL Id: hal-01958293

<https://hal.archives-ouvertes.fr/hal-01958293>

Submitted on 2 Jun 2020

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.

Végéphy1 – 12e CONFÉRENCE INTERNATIONALE SUR LES MALADIES DES PLANTES
TOURS – 10, 11 ET 12 DÉCEMBRE 2018

**PRENDRE EN COMPTE L'INCERTITUDE DES PRÉVISIONS MÉTÉOROLOGIQUES DANS LES
OAD UTILISÉES POUR GERER LES MALADIES ET RAVAGEURS DES CULTURES.**

I. ALEKSOVSKA^(1,2,3), F. BRUN⁽¹⁾, L. RAYNAUD⁽²⁾, R. FAIVRE⁽³⁾, M. RAYNAL⁽⁴⁾, O. DEUDON⁽⁵⁾

⁽¹⁾ ACTA, Castanet Tolosan, France, ivana.aleksovska@acta.asso.fr

⁽²⁾ Météo-France, Toulouse, France

⁽³⁾ INRA, Castanet Tolosan, France

⁽⁴⁾ IFV, UMT SEVEN, Villenave d'Ornon, France

⁽⁵⁾ Arvalis, Boigneville, France

RÉSUMÉ

Les maladies et ravageurs des cultures dépendent fortement des conditions météorologiques. Différents outils d'aide à la décision sont proposés pour prendre en compte ces conditions météorologiques dans la modélisation des dynamiques épidémiques et accompagner les agriculteurs dans leurs prises de décision. Néanmoins, l'écoulement atmosphérique est un phénomène chaotique et les prévisions météorologiques restent incertaines. Dans ce contexte, les centres de prévision météorologique, dont Météo-France, mettent en œuvre des systèmes de prévision probabiliste qui permettent d'accéder à une estimation de l'incertitude des conditions météorologiques prévues. Nous proposons de montrer l'intérêt d'utiliser ces prévisions probabilistes au travers d'outils d'aide à la décision et d'une évaluation des risques associés à cette décision pour la protection des cultures, avec des applications sur la vigne et le blé.

Mots-clés : incertitude, prévisions météorologiques, OAD, septoriose du blé, ver de la grappe.

ABSTRACT

**ACCOUNTING FOR THE UNCERTAINTY OF WEATHER FORECASTS IN DECISION SUPPORT
SYSTEMS USED TO MANAGE CROP DISEASES AND PESTS.**

Crop diseases and pests are highly dependent on weather conditions. Various decision support systems are proposed to take into account these meteorological conditions in the modeling of epidemic dynamics and to assist farmers in their decision-making. Nevertheless, the atmospheric flow is a chaotic phenomenon and weather forecasts remain uncertain. In recent years, weather prediction centers, including Météo-France, have implemented probabilistic prediction systems that provide an estimate of the uncertainty of the weather forecasts. We propose to demonstrate the potential of using these probabilistic forecasts through decision support tools used in an operational way for the protection of crops with applications to the wine and wheat.

Keywords: uncertainty, numerical weather predictions, DSS, wheat septoria, worms of grape vine.

INTRODUCTION

Les maladies et ravageurs des cultures évoluent en fonction des conditions météorologiques. Si des traitements phytosanitaires permettent de maîtriser certaines de ces maladies ou de ces ravageurs, il est néanmoins nécessaire de prendre en compte la situation sanitaire réelle en cours de campagne afin de prendre une décision d'intervention ou non, puis le cas échéant de bien positionner le traitement en fonction du mode d'action des produits sur les cibles visées. En fonction des années, des épidémies apparaissent plus ou moins en avance ou sont plus ou moins importantes. Par exemple, sur la septoriose du blé en région Centre, l'incidence médiane régionale de la maladie sur l'étage foliaire F3 a dépassé le seuil de 20% le 3 mai en 2016 contre le 22 mai en 2013 d'après les données du réseau du Bulletin de Santé du Végétal (Michel *et al*, 2015).

Différents outils d'aide à la décision sont proposés dans la modélisation des dynamiques épidémiques pour accompagner les agriculteurs dans leurs prises de décision, comme le modèle intégré à l'outil Septo-LIS® pour le positionnement du premier traitement sur la septoriose du blé (Gouache et Couleaud, 2009) ou le modèle EVA sur le vers de la grappe (adapté de Chavent, 1983). Ils sont complémentaires des observations de terrain dans la mesure où ils permettent d'anticiper la situation et d'accéder à des variables difficiles à observer sur le terrain. Par ailleurs, les nouvelles solutions de lutte de type biocontrôle présentent des niveaux d'efficacité moindre qui nécessitent en contrepartie plus de précision dans leur positionnement.

Ces modèles s'appuient largement sur les prévisions météorologiques. Néanmoins, ces prévisions restent souvent incertaines. En effet, l'atmosphère est un système chaotique et la mise en œuvre de prévisions météorologiques de qualité constitue un défi scientifique important car il existe de nombreuses sources d'incertitude, liées notamment à l'estimation de l'état initial de l'atmosphère et à la représentation précise des différents processus en jeu. Dans ce contexte, les techniques de prévision probabiliste, mises en œuvre sous la forme d'ensembles de prévisions perturbées, sont proposées par plusieurs centres de prévision dans le monde, dont Météo-France, pour fournir une estimation des incertitudes de prévision. A notre connaissance, ces prévisions d'ensemble ne sont pas encore utilisées dans des applications agricoles.

Notre objectif est de présenter le principe de la prévision d'ensemble météorologique ainsi que son potentiel pour l'application au domaine de la protection des cultures.

MATERIEL ET METHODE

PRINCIPE DE LA PREVISION NUMERIQUE DU TEMPS

Depuis la seconde moitié du XXème siècle, des systèmes de prévision numérique du temps sont utilisés pour simuler quotidiennement l'évolution temporelle de plusieurs paramètres météorologiques (par exemple le vent, la pression, la température, l'humidité) sur des échéances de quelques heures à quelques jours. Ces prévisions s'appuient (Figure 1) notamment sur : 1) la connaissance des lois qui gouvernent l'évolution au cours du temps de l'atmosphère, qui forment le cœur du modèle de prévision ; 2) la connaissance de l'état actuel de l'atmosphère, également appelé état initial.



Figure 1 : Principe de la prévision numérique du temps.
Principle of numerical weather forecast.

Depuis la mise en place des premiers modèles, les prévisions ont connu de fortes améliorations, notamment grâce au développement des réseaux d'observations (en particulier les satellites), à une meilleure compréhension et représentation des processus dynamiques et physiques dans les modèles, et à l'augmentation des ressources informatiques qui permet d'améliorer la finesse des grilles de calcul. Néanmoins, les conditions pour une prévision météorologique parfaite n'existent pas à cause d'erreurs dans l'estimation de l'état initial et d'approximations dans le développement du modèle, qui est une représentation approchée du fonctionnement réel de l'atmosphère.

APPROCHE PROBABILISTE DE LA PREVISION NUMERIQUE DU TEMPS

En raison du caractère chaotique de l'atmosphère, ces erreurs s'amplifient au cours de la prévision. Les travaux de Lorenz (Lorenz, 1963) montrent notamment que des incertitudes d'amplitude infime sur l'état initial augmentent rapidement et peuvent rendre la prévision complètement fautive au bout d'un certain temps. La prévision déterministe n'a alors de sens que si l'on considère l'incertitude qui lui est associée.

Afin d'estimer l'incertitude des prévisions, une approche probabiliste de la prévision devient nécessaire. Il ne s'agit plus de prévoir l'état le plus probable de l'atmosphère mais la probabilité qu'un événement particulier se produise. La prévision probabiliste (appelée également prévision d'ensemble) repose actuellement sur la mise en œuvre d'ensembles de prévisions afin d'échantillonner la distribution de probabilité de l'état de l'atmosphère. En pratique, cela consiste à exécuter le modèle de prévision plusieurs fois afin de tenir compte des différentes sources d'incertitude : chaque prévision utilise des conditions initiales très légèrement perturbées et éventuellement des configurations du modèle légèrement différentes (Figure 2).

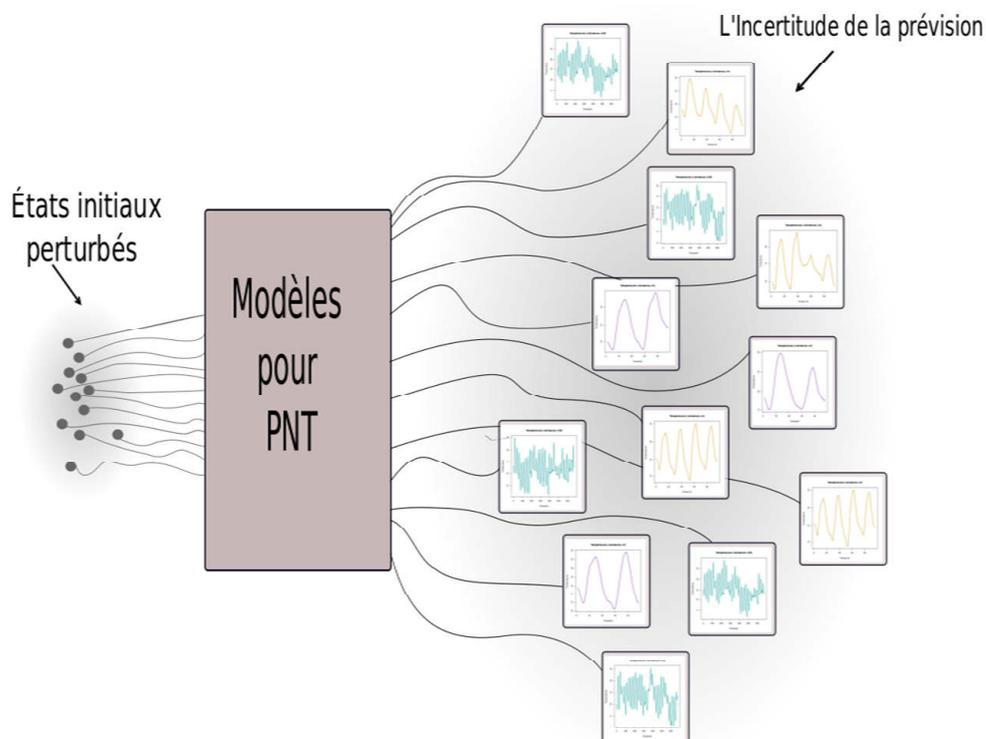


Figure 2 : Principe de la prévision d'ensemble pour les prévisions météorologiques.
Principle of ensemble prediction systems.

Cette technique de prévision d'ensemble est utilisée depuis les années 1990 (Bouttier et Buizza, 2018) pour fournir des probabilités d'événements mais aussi un éventail de scénarios météorologiques, dont la dispersion renseigne sur la prévisibilité des phénomènes simulés.

Elle est aujourd'hui devenue un outil clé pour la prévision météorologique et pour plusieurs applications (prévision des crues, gestion du réseau électrique, routage des avions par exemple).

SYSTEMES DE PREVISION OPERATIONNELS A METEO-FRANCE

Deux types de modèle sont généralement mis en œuvre : les modèles globaux, qui couvrent tout le globe, et les modèles régionaux (ou à aire limitée) couvrant seulement un domaine géographique restreint (typiquement un pays et son voisinage). Outre leur couverture géographique ces modèles diffèrent également par leurs grilles de calcul et leurs échéances. Les modèles régionaux utilisent des maillages plus fins et des échéances plus courtes que les modèles globaux et permettent ainsi de simuler plus précisément les phénomènes de petite échelle.

Météo-France développe et utilise un modèle global, ARPEGE, et un modèle régional, AROME, pour ses prévisions opérationnelles. Le modèle IFS du Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT) est également utilisé par les services de production de Météo France.

Météo-France réalise également des prévisions probabilistes, à partir de versions ensemblistes des modèles précités. On les note dans la suite ARPEGE-EPS, AROME-EPS et IFS-EPS. Ces modèles ont 35, 12 et 51 prévisions respectivement.

ILLUSTRATION DES PREVISIONS PROBABILISTES

Un exemple de prévision probabiliste de la température est représenté sur la figure 3 pour les systèmes AROME-EPS, ARPEGE-EPS et IFS-EPS (avec leurs échéances respectives de 2, 4 et 15 jours).

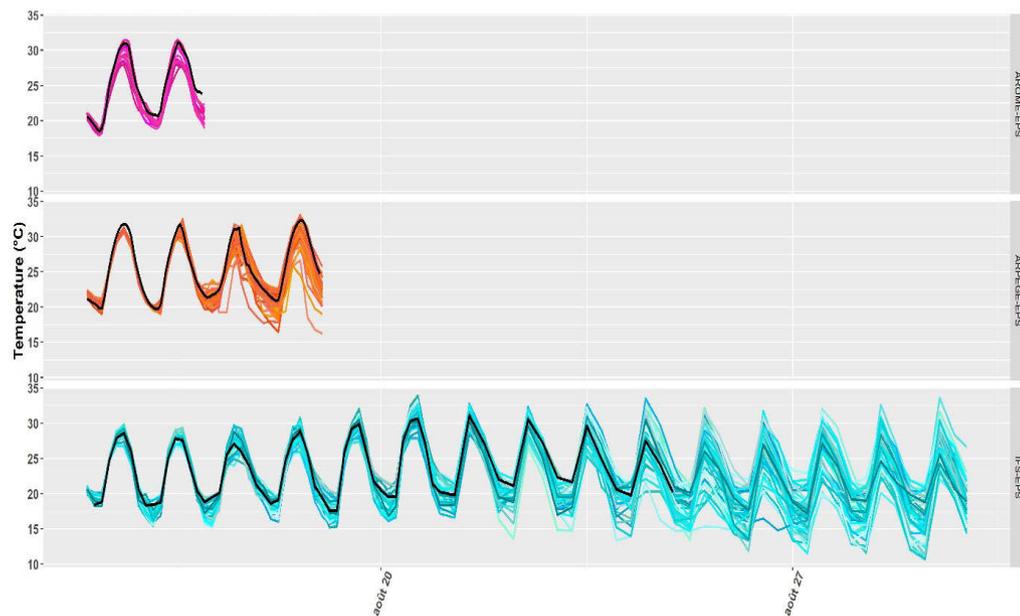


Figure 3 : Exemple de prévision ensembliste pour la température à partir des modèles AROME-EPS (haut), ARPEGE-EPS (milieu) et IFS-EPS (bas) ainsi que la prévision déterministe présentée en noir sur les trois modèles. Chaque courbe correspond à une prévision de l'ensemble. On remarque que la variabilité des scénarios augmente avec l'échéance, ce qui traduit une augmentation de l'incertitude.

Example of ensemble forecast for temperature from models AROME-EPS (top), ARPEGE-EPS (middle) and IFS-EPS (bottom) as well as the deterministic forecast presented with the black line on the three models. Each line corresponds to a forecast of the ensemble. It can be noticed that the variability of the scenarios increases with lead time, which reflects an increase of the uncertainty.

Cette information peut également être synthétisée sous la forme de distributions de probabilité (figure 4).

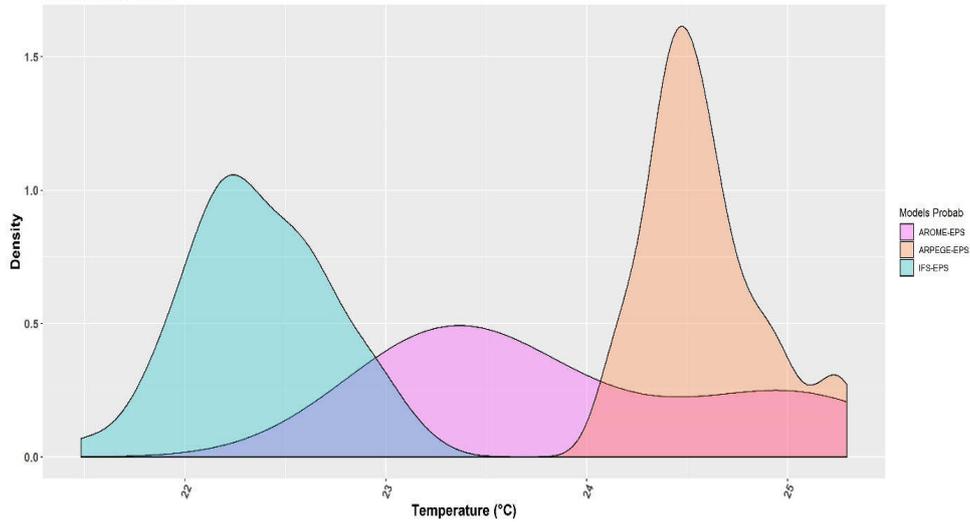


Figure 4 : Fonction de densité de probabilité des températures journalières moyennes aux échéances J+2 à Pic Saint Loup (34). Début des prévisions : 15/08/2018.
Probability density function of the average daily temperatures at lead times $D + 2$ in Pic Saint Loup (34). Initial time of the forecast: 15/08/2018.

La variabilité des scénarios météorologiques pourrait également être représentée via des histogrammes. La Figure 5 donne un exemple de cette représentation en prenant en compte le cumul quotidien de la pluie.

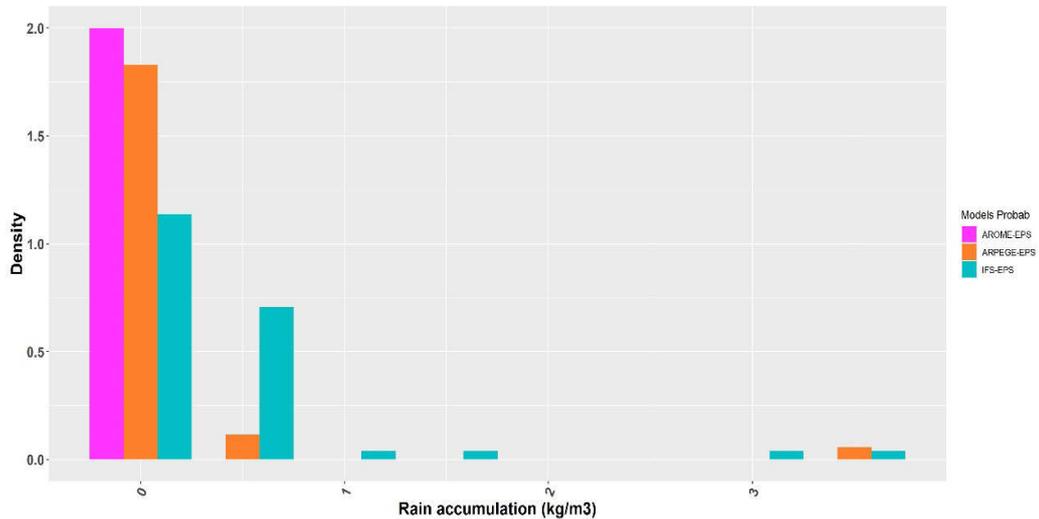


Figure 5 : Histogramme des précipitations journalières moyennes aux échéances J+2 à Pic Saint Loup (34). Début des prévisions : 06/09/2018.

Histogram of the average daily precipitation at maturity $D + 2$ in Pic Saint Loup. Initial time of the forecast : 06/09/2018.

RESULTATS

APPLICATION A LA PREVISION DES RISQUES DE MALADIE ET DE RAVAGEURS

Les prévisions météorologiques sont très souvent des données essentielles pour les modèles de dynamique de maladies et ravageurs utilisés dans les outils d'aide à la décision pour anticiper le déclenchement et le positionnement des traitements. Nous proposons d'illustrer le potentiel des prévisions météorologiques d'ensemble sur deux modèles utilisés de manière opérationnelle en protection des cultures : le ver de la grappe de vigne et la septoriose du blé. Il s'agit de propager au travers des modèles agronomiques les incertitudes météorologiques décrites par les trois systèmes de prévision d'ensemble présentés précédemment (Figure 6).

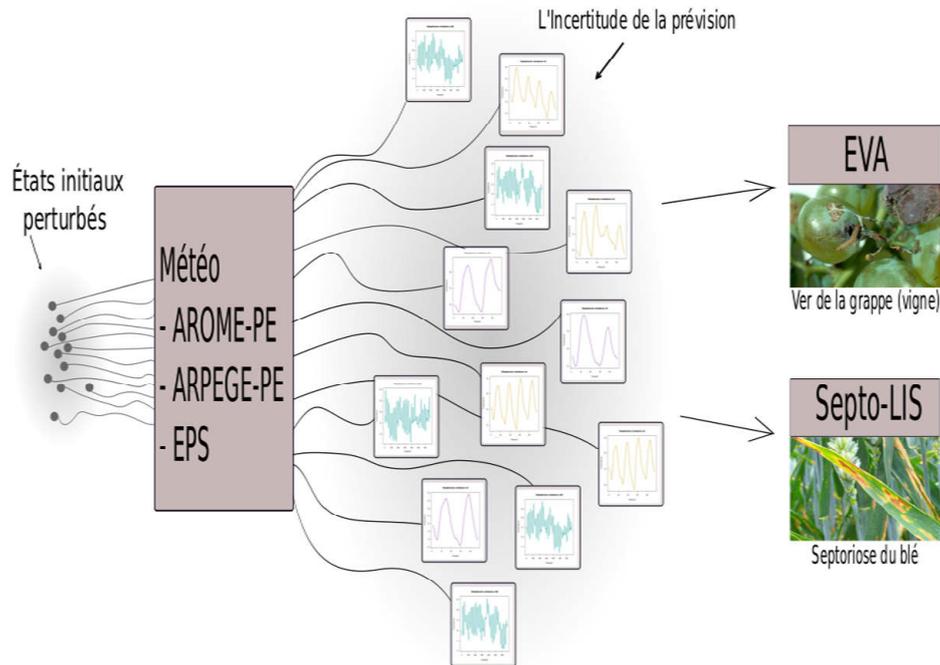


Figure 6 : Propagation des incertitudes des prévisions météorologiques pour prévoir les interventions en septoriose du blé et en ver de la grappe en vigne.
Propagation of the uncertainties of the weather forecasts to predict the interventions in septoria of wheat and worm of the grape in vine.

POSITIONNEMENT DES TRAITEMENTS DE LUTTE CONTRE LE VERS DE LA GRAPPE EN VIGNE.

Les vers de la grappe sont des insectes ravageurs de la vigne. Deux espèces de lépidoptères, la Cochylis (*Eupoecilia ambiguella*) et l'Eudémis (*Lobesia botrana*) pondent leurs œufs sur les inflorescences et les baies de raisin, dont les chenilles se nourrissent pour assurer leur développement, générant ainsi des dégâts quantitatifs et qualitatifs. L'insecte évolue en 4 stades au cours de son cycle : imagos, œufs, larves (L1, L2, L3, L4 et L5), chrysalides (Nymphé et Nymphé diapausante) et peut avoir plusieurs cycles de reproduction. De plus, le nombre de cycles est difficile à prévoir et est très dépendant des conditions météorologiques. Afin de lutter contre ces ravageurs, les viticulteurs disposent de plusieurs moyens de lutte : traitements phytosanitaires chimiques, outils de biocontrôle (trichogrammes et toxines de *Bacillus thuringiensis*) ou par confusion sexuelle.

Pour bien positionner le traitement, des modèles mathématiques sont utilisés en complément d'observations faites sur le terrain pour qualifier la présence effective du ravageur. Ainsi ces modèles estiment la dynamique des différents stades de la population des vers sur toute la saison. Dans le modèle EVA, utilisé ici sur l'Eudémis, le développement de l'insecte dépend uniquement de la température (Chavent, 1983), qui est renseignée de manière

horaire. Sur la 2^e et 3^e génération, le modèle est utilisé pour optimiser le positionnement des traitements phytosanitaires.

Sur la base du modèle existant et de ces seuils, pour la date donnée, nous avons ainsi utilisé les prévisions météorologiques d'ensemble des températures horaires fournies par les 3 systèmes (Figure 7 A). Pour cette simulation, nous avons utilisé les observations météorologiques disponibles jusqu'au 14 août 2018 compris puis les prévisions ensemblistes au-delà, pour réaliser la propagation des incertitudes météorologiques. Ainsi, en sortie, à chaque date de prévision plusieurs simulations du modèle EVA pour la proportion d'œufs sont obtenues (résumées ici sous la forme de boîtes à moustaches, Figure 7 B). La Figure 7A présente les prévisions déterministe et ensembliste de température journalière moyenne pour les 15 et 16 août 2018. La distribution de chaque prévision d'ensemble est représentée sous la forme de box plot. La Figure 7B présente les sorties d'EVA correspondantes, pour la proportion du stade œuf.

On passe ainsi d'une distribution de température à une distribution pour la population d'un stade. On voit également un lien entre la distribution des prévisions météorologiques en entrée et la distribution des prédictions agronomiques en sortie : le modèle météorologique ayant la plus grande incertitude conduit à la plus grande variabilité des sorties agronomiques sur ce cas.

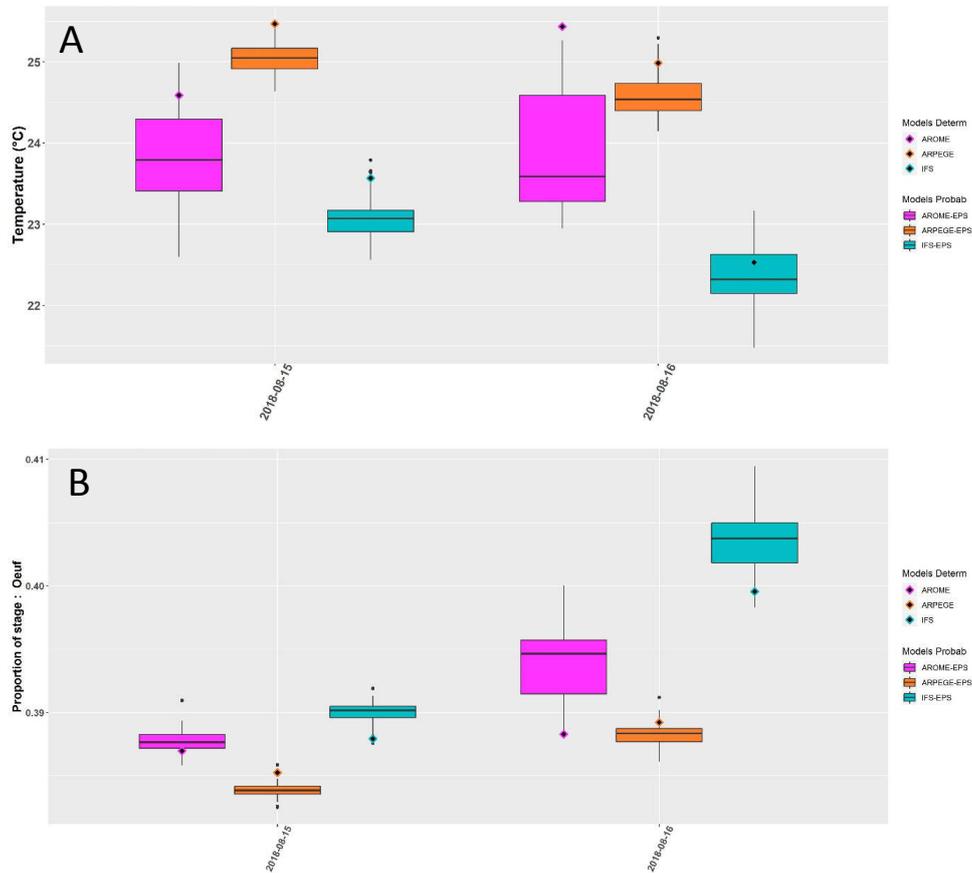


Figure 7 :

A. Prévisions météorologiques d'ensemble des températures moyennes fournies par les 3 systèmes pour les 15 et 16 août 2018, B. proportion d'œufs correspondante à Pic Saint Loup (34). Boxplot : modèles probabilistes, Point : modèles déterministes. Début des prévisions : 15/08/2018.

- A. Overall weather forecasts of average temperatures provided by the 3 systems for August 15 and 16, 2018, B. corresponding proportion of the stage egg in Pic Saint Loup. Boxplot : Probabilistic models, Point : Deterministic models. Initial time of the forecast 15/08/2018.

Sur le stade L4, en appliquant un seuil de décision de 2% (traitement chimiques) et 15% (biocontrôle) (Figure 8 A) nous obtenons alors une distribution des dates préconisées pour appliquer le traitement, au lieu d'une date unique. Avec le seuil de 2% l'incertitude est réduite (Figure 8B), alors que pour le seuil de 15% l'incertitude est plus grande (Figure 8 C). Cela est lié au fait que le seuil 15% est atteint avec des prévisions d'échéance plus longue, pour lesquelles l'incertitude est plus forte. Sur la Figure 8, nous avons utilisé les prévisions issues du modèle IFS-EPS pour la visualisation car son échéance est la plus lointaine. Les prévisions d'AROME-EPS et d'ARPEGE-EPS ne permettaient pas de franchir le seuil nécessaire.

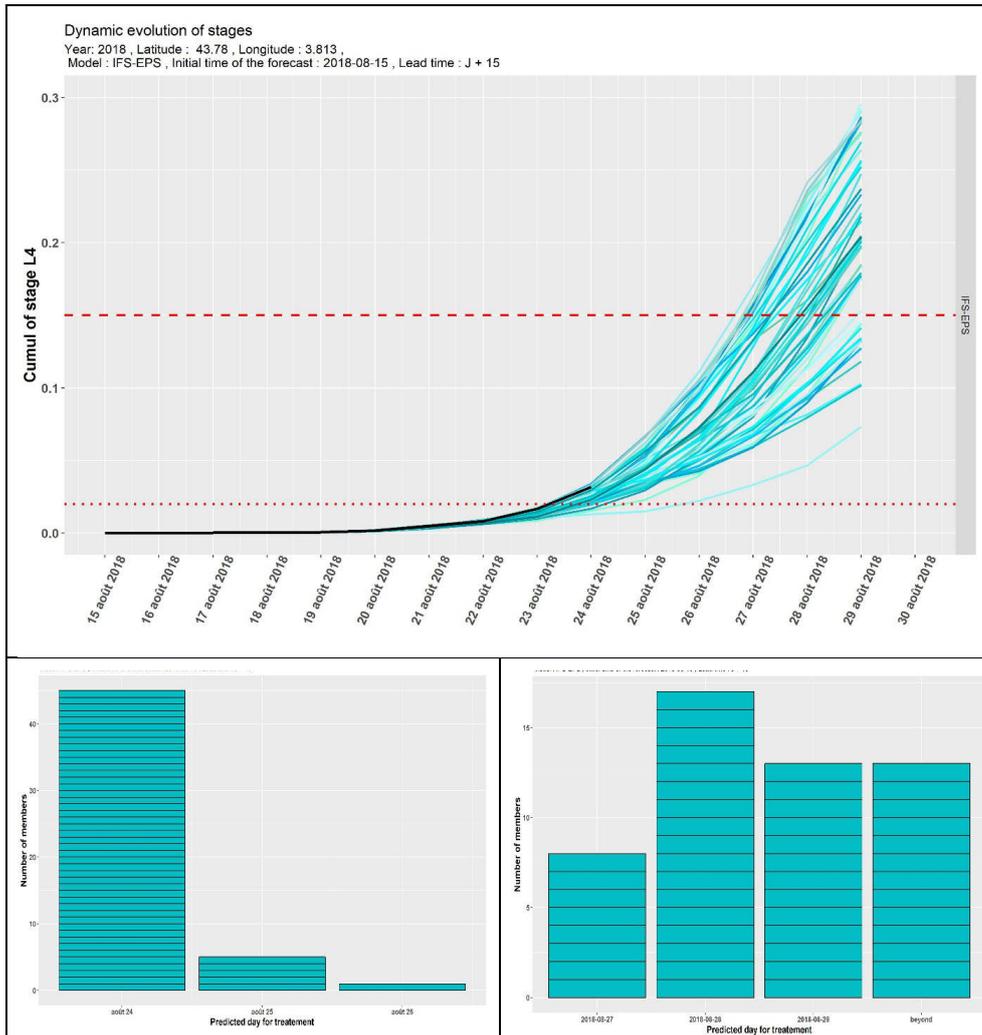


Figure 8 : A. Résultats des simulations d'ensemble pour le stade de larve L4 obtenus en utilisant des prévisions allant jusqu'à J+15 et franchissement du seuil de 2 % (ligne pointillée) et 15 % (ligne tiretée). B. Histogrammes des dates de déclenchement du traitement 2%. C. Histogrammes des dates de déclenchement du traitement de 15%.

A. All simulation results for the larval stage L4 obtained using forecasts up to D+15 and crossing the 2% threshold (dotted line) and 15% threshold (dashed line). B. Histograms of treatment initiation dates 2%. C. Histograms of treatment initiation dates 15%.

Sur la base de discussions avec des experts (IFV et INRA), les 2 seuils d'intérêt (2 et 15%) sont pertinents pour déclencher le positionnement du traitement sur le cumul d'œuf. Un traitement chimique doit être placé lorsque le cumul d'œufs atteint 2% ; un traitement de biocontrôle sera quant à lui positionné au seuil de 15% du cumul d'œufs en raison de la plus faible rémanence du produit. Néanmoins, nous illustrons ces résultats sur la base du stade larve 4 (Figure 8), faute de disponibilité des prévisions IFS-EPS pour des dates plus anciennes (travail en cours).

DECLENCHEMENT ET POSITIONNEMENT DU PREMIER TRAITEMENT DE LUTTE CONTRE LA SEPTORIOSE DU BLE.

La septoriose du blé est une maladie majeure en France. ARVALIS-Institut du végétal a développé un modèle épidémiologique de développement de la septoriose sur blé intégré à un outil, Septo-LIS®, permettant de positionner le 1er traitement contre cette maladie (Gouache et Couleaud, 2009). L'objectif est de protéger la surface verte des deux dernières feuilles. Ce premier traitement constitue ainsi un compromis entre l'attente de l'émergence de ces deux feuilles et le contrôle précoce de la maladie. Septo-LIS® décrit de manière couplée la dynamique de la maladie et celle de la plante. Il s'agit d'un modèle biologique décrivant notamment les contaminations reçues par les différents étages foliaires, en prenant en compte les entrées météorologiques journalières, notamment les précipitations et les températures. Une règle de décision basée sur la somme des contaminations des feuilles permet ensuite de prévoir la date de déclenchement du 1er traitement.

Nous avons procédé de manière similaire au modèle du ver de la grappe. Cependant, Septo-LIS® nécessite plusieurs entrées météorologiques, notamment la pluie, pour laquelle les prévisions sont souvent plus incertaines. De plus, le modèle se base sur des données météorologiques journalières contrairement au modèle EVA qui nécessite des données horaires.

Pour cette étude, nous avons utilisé les prévisions d'ARPEGE-EPS sur les quatre premiers jours et pour les jours suivants, la médiane des séries climatiques observées sur le site les années précédentes. En fait, nous n'avons pas à notre disposition les prévisions du modèle européen IFS-EPS sur 14 jours et nous n'avons pas procédé à des tirages aléatoires dans les séquences climatiques pour obtenir plus de variabilité afin de limiter nos temps de calcul.

Sur la figure 9, on observe la répartition des dates de traitement T1 obtenues par le modèle sur les 35 membres d'ARPEGE-EPS pour chaque date à partir des prévisions prises en compte.

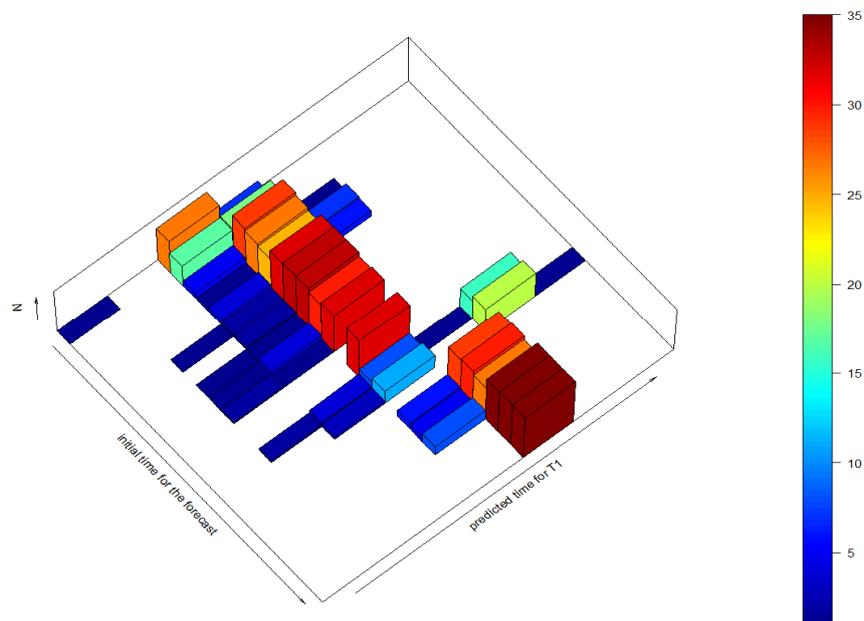


Figure 9 : Histogramme des dates T1 simulées par le modèle Septo-LIS® à partir de scénarios climatiques issus du modèle ARPEGE-EPS. En abscisse x, dates de début de prise en compte des prévisions du modèle ARPEGE-EPS (du 10/04 au 30/04) et en ordonnées y, dates simulées du traitement T1 (du 25/04 au 01/05).

Sur la figure 9, nous pouvons remarquer principalement 3 phases. Pendant la première (du 10/04 au 21/04), nous pouvons constater une dispersion des dates de prévision se concentrant toutefois vers la date du 28 avril. Du 25 avril au 30 avril, la dispersion est moindre et se concentre sur la date du 29 avril. Pendant la phase intermédiaire, la dispersion est beaucoup plus importante (dates T1 du 26 avril au 1^{er} mai). De ces premiers résultats, nous ne pouvons tirer que quelques conclusions partielles qui méritent d'être affinées. L'utilisation de prévisions ensemblistes amène à une dispersion des dates T1 quel que soit la date de début de la prise en compte de ces ensembles mais cette dispersion est moindre quand on se rapproche de la date à prédire. La différence du maximum en la première phase et la dernière phase s'explique de notre point de vue par l'utilisation d'une prévision identique au-delà des 4 premiers jours. On peut remarquer toutefois qu'une dispersion subsiste même aux premières dates de l'expérience. Quant à la phase intermédiaire, il semble que les prévisions météorologiques ensemblistes soient assez variables conduisant à une grande variabilité des scénarios climatiques. Il pourrait s'agir d'une période de conditions météorologiques instables où les modèles ensemblistes rendent mieux compte de ces incertitudes. Afin d'affiner cette analyse, il faudrait reconduire cette étude soit avec les prévisions du modèle IFS-EPS, soit en générant de l'aléa sur la météo au-delà des 4 jours par des tirages aléatoires dans les archives météorologiques du site considéré.

DISCUSSION

Les prévisions d'ensemble en météorologie sont encore récentes et peu connues par le monde agricole. Pourtant, les conditions météorologiques sont déterminantes dans les dynamiques épidémiques et il existe de nombreux modèles utilisés comme outil d'aide à la décision mobilisant des données météorologiques. Aussi, il est important d'évaluer les apports de ces prévisions probabilistes dans le domaine de la protection des cultures.

Ici, nous nous sommes focalisés sur les prévisions météorologiques comme unique source d'incertitude dans l'utilisation de modèles d'aide à la décision, ce qui est une approximation forte. En effet, les modèles biologiques sont loin d'être parfaits et il s'agira également de prendre en compte leurs incertitudes, et éventuellement de combiner des observations en cours de saison pour corriger et réduire ces incertitudes (Brun *et al.*, 2012). Concernant les entrées météorologiques, il ne faudra pas négliger l'utilisation des informations issues de stations connectées chez les agriculteurs tout en gardant à l'esprit de potentielles incertitudes sur les données observées.

Ce passage à une approche probabiliste pourrait être intuitif pour le monde agricole, habitué à vivre avec ces aléas, consultant très fréquemment plusieurs sources de prévisions météorologiques pour prendre des décisions. Mais, l'information sur l'incertitude peut également pousser certains décideurs à se positionner trop précocement dès que le risque devient visible. Aussi, il sera nécessaire de les accompagner pour bien interpréter cette information plus riche. Il s'agira également de penser à l'ergonomie des sorties des modèles probabilistes pour qu'ils soient facilement interprétables.

Dans cet article nous présenterons des résultats préliminaires des travaux de la thèse en utilisant tous les scénarios météorologiques et en les couplant aux modèles agronomiques pour les simulations numériques. Nous précisons que la problématique principale de la thèse est de faire une stratégie de sélection/couplage de ces scénarios disponibles pour mieux représenter l'incertitude sur les sorties des modèles agronomiques tout en essayant de réduire le coût de calcul (travail en cours). Ces résultats pourront être généralisés à différentes situations climatiques et d'autres OAD.

REMERCIEMENTS

Ces travaux sont menés dans le cadre de la thèse d'Ivana Aleksovska "Améliorer les prévisions à court et moyen termes des modèles agronomiques en prenant mieux en compte l'incertitude des prévisions météorologiques", financée par l'institut de convergence DigitAg (Agence Nationale de la Recherche au titre du programme d'Investissements d'Avenir portant la référence ANR-16-CONV-0004), le projet "METEOPREC - Les apports de la météo de précision au service des agriculteurs" (Casdar, AAP RT 2018), l'Acta, Arvalis et l'IFV via l'UMT SEVEN (Santé des Ecosystèmes Viticoles Economes en intraNts) et RMT Modélisation et Analyse de données.

BIBLIOGRAPHIE

Bouttier F. et R. Buizza. (2018). La prévision d'ensemble. *La Météorologie*, (100) :pp. 29_36.

Brun F., Makowski D. et Piroux F., 2012. Comment combiner réseau d'observations, expertise et modélisation pour élaborer le Bulletin de Santé du Végétal ? Actes AFPP CIMA - 10e conférence internationale sur les maladies des plantes tours 3, 4 et 5 décembre 2012.

Chavent F. Etude de la dynamique des populations d'Eudémis de la vigne (*Lobesia botrana* Den. et S hi.). Adaptation d'un modèle d'évolution aux conditions climatiques de Provence, Mémoire de fin d'étude. PhD thesis, E.N.S.S.A.A. Dijon, 1983.

Gouache D., Couleaud G. (2009). Le positionnement des traitements fongicides : enjeu pour la septoriose et intérêt du modèle « Septo-LIS® ». AFPP, 9ème Conférence Internationale sur les Maladies des Plantes, Tours, 8-9 décembre 2009. pp 455-461.

Lorenz Edward N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130-141.

Michel L., Decarrier A., Franche M., Bochu V., Benredjem K., Hugerot G., Simonneau D., Piroux F., Gourdain E., Veslot J., Makowski D., Brun F. (2015). Un outil d'analyse des dynamiques épidémiologiques pour le Bulletin de Santé du Végétal : application à la septoriose du blé. AFPP –Onzième Conférence Internationale sur les Maladies des Plantes. Tours – 7 au 9 décembre 2015.

**Article crop protection - version en
cours de révision après la
soumission (soumis en décembre
2019)**

***Manuscript**

[Click here to view linked References](#)

Title

Accounting for the uncertainty of weather forecasts in decision support systems used to manage crop diseases and pests.

Authors

Ivana Aleksovska (a,b,c,*), François Brun (a), Laure Raynaud (b), Robert Faivre (c), Marc Raynal (d), Olivier Deudon (e)

Affiliation

(a) ACTA – French Technical Institute, UMR1248 AGIR, 31326 Castanet-Tolosan, France

(b) Météo-France, CNRM/GMAP/RECYF, Toulouse, France

(c) INRAE, UR875 MIAT, Castanet-Tolosan, France

(d) IFV, UMT SEVEN, Villenave d'Ornon, France

(e) Arvalis, Boigneville, France

Abstract

Crop diseases and pests are highly dependent on weather conditions. Various decision support systems are proposed to take into account these meteorological conditions in the modelling of epidemic dynamics of pests and diseases and to assist farmers in their decision-making for phytosanitary treatment. Nevertheless, the atmospheric flow is chaotic and weather forecasts remain uncertain. In recent years, several meteorological services have developed probabilistic prediction systems that provide an estimate of weather forecasts uncertainty. Using these probabilistic forecasts within the framework of disease forecasting through decision support tools provides a distribution of treatment dates instead of a single date with deterministic approach. The potential of this probabilistic information is here illustrated for the management and protection of crops with application to the vine grape berry moth and *Septoria tritici* blotch of wheat. Agronomy simulations are performed using observed-weather data and long-range probabilistic forecast starting on different initial dates. The treatment date associated to the observed-weather, is then used as a reference treatment date. The results shown that if the initial date of the forecast is far from the reference treatment date, the uncertainty about the treatment dates is higher. In contrast, getting closer to the date of reference treatment date, the uncertainty about the treatment dates becomes progressively lower. The range of uncertainty can vary from few days to dozen days depending of the spatial location and initial date of forecast. It is therefore necessary to use precocious simulations that allow for medium-term anticipation and then use daily updated forecast to refine the information of uncertainty of treatment days.

Highlights

- Probabilistic approach of weather forecasting is used to handle the uncertainty for decision making.
- Decision support systems are used for pests and diseases management.
- Use of probabilistic weather forecasting leads to a distribution of dates for treatment.
- The uncertainty about treatment date decreases as we approach to the observation-based treatment date.

Keywords: probabilistic weather forecasts, disease forecasting, phytosanitary treatment, decision support tools, *Eupoecilia ambiguella*, *Mycosphaerella graminicola*

1) Introduction

Agriculture is one of the most weather-dependent economic activities, in particular because crop diseases and pests develop according to weather conditions (Chaulagain, et al., 2019), (Moyer, et al., 2016), (Pietravalle, et al., 2003), (Te Beest, et al., 2008). Phytosanitary treatments make it possible to control some of these diseases or pests, but they need to be applied according to different norms, criteria and only when required (Cooke, et al., 2006). It is necessary to consider the current and future disease situation in order to take a decision on whether or not to apply a treatment (Rossi, et al., 2012). In addition, properly positioning the treatment date according to the product type is another essential aspect. For instance, new biocontrol solutions have lower levels of efficiency, and thus may require more precision in their positioning (Shipp, et al., 1999). For this purpose, different Decision Support Tools (DST) are developed for the modelling of epidemic dynamics to support farmers in their decision-making (Pertot, et al., 2017). They are

complementary to field observations in that they provide an opportunity to anticipate the situation and to access variables that are difficult to observe. Most of these tools require accurate weather forecasts from the next hour to several days.

Numerical Weather Prediction (NWP) models are based on a set of equations that govern the evolution of the atmospheric dynamics (Abbe 1901), (Bjerknes 1999). Efficient and accurate numerical methods are used to resolve these equations on a grid-based discretization of the atmosphere (Richardson 2007) for different time ranges. In addition, the initial state of the atmosphere is computed from the observations available through in situ (e.g., ground and sea-based measurement stations, radiosondes) and satellite data using complex algorithms for data assimilation (Courtier, et al., 1994) (Courtier, et al., 1998). Due to the chaotic nature of the atmosphere, errors in the modelling and in the initial state, even small, increase during the forecast and can lead to large forecast uncertainties after a certain time (Lorenz, 1963). Hence, deterministic forecasting that provides a single forecast only makes sense if one considers the associated uncertainty. For that purpose, a probabilistic approach becomes necessary and aims at predicting the probability distribution of the atmospheric state (instead of a single value). Probabilistic prediction currently relies on the implementation of Ensemble Prediction Systems (EPSs) that provide an ensemble of perturbed forecasts. Each perturbed forecast is designed to account for the different sources of uncertainty and thus uses slightly perturbed initial conditions and possibly different model configurations (Leutbecher, et al., 2008). This ensemble prediction technique has been used operationally since the 1990s to provide not only event probabilities but also a range of possible weather scenarios. It has now become a major tool for weather forecasting, and it can be used in several downstream applications (Fundel, et al., 2019) including agriculture (Christ, et al., 2015), power grid management (Pinson, et al., 2009), aircraft routing (Cheung, et al., 2015). Deterministic and probabilistic forecasts are provided by either global or limited-area models. Global models cover the entire globe with a horizontal grid spacing between 10 and 60 Km and are generally used for medium-range forecasting of 3-15 days ahead. Limited-area models, with a horizontal grid spacing of 1-4km, run over smaller domains (typically a country and its neighbourhood) and for shorter timescales (1 or 2 days). These regional models are able to better resolve some aspects of local weather such as the location and intensity of thunderstorms. In this work we use mainly IFS (Integrated Forecast System), the global deterministic model of the European Centre for Medium-Range Weather Forecasts (ECMWF) and IFS-EPS the corresponding probabilistic model (Palmer, 2019). Météo-France, the French national weather service, also develops and uses a global model, ARPEGE, and a convective-scale model, AROME, for its operational forecasts, along with the corresponding probabilistic versions AROME-EPS (Raynaud, et al., 2016) and ARPEGE-EPS (Descamps, et al., 2015). However, their short lead times are less appropriate to anticipate the need for a phytosanitary treatment several days in advance.

It is advised to agronomists who are interested in avoiding risks related to their decision to use a phytosanitary treatment to take full advantage of probabilistic information instead of using, for example, the mean of the outputs from agronomic models (Carlson, 1970). Ignoring uncertainty can lead to a wrong decision and act only on deterministic information, farmers can spray the treatment too late and then have an average loss (Rossing, et al., 1994). Probabilistic forecasts have already been used in agriculture for seasonal forecasting or climate change impact studies (Zhao, et al., 2019) (Ogut, et al., 2018), but significantly less for shorter time scales (Christ, et al., 2015), (Aleksavska, et al., 2018), (Raposo, et al., 1993). As a result of ensemble forecast, a distribution of treatment dates will be proposed to agronomists instead of a single date. However, the final decision should be made by the agronomists taking into account the risks and profitability associated with their decision using the total economic cost (the cost of phytosanitary treatments and the cost of yield loss) (Rossing, et al., 1994).

The goal of this paper is to demonstrate the potential of ensemble weather prediction for application to crop protection, based on two DSTs widely used in France: EVA for grape berry moth and Septo-LIS® for wheat septoria.

In French vineyard, there are two species of insects from the lepidopteran order that are at the source of grape berry moth symptoms: *Eupoecilia ambiguella* and *Lobesia botrana* (Thiery, et al., 2013). Both lay their eggs on inflorescences and grape berries, on which the larvae feed to ensure their development, and so generate quantitative and qualitative damages (Chavent, 1983). This article will focus on the *Eupoecilia ambiguella*. This insect evolves in four successive development stages during its cycle and can have several reproduction cycles, generally 3 and can go up to 5 per year, but the 2nd and 3rd cycles are the most critical with regard to damages. The number and the precise timing of cycle are very dependent on weather conditions and remain difficult to predict. In order to control grape berry moth, winegrowers have three types of treatments: chemical products, biocontrol tools (trichograms and *Bacillus thuringiensis* toxins) (Thiery, et al., 2018) and sexual confusion. The correct positioning of the treatment depends on the mode of

action of its active ingredient in response to the target (eggs, larvae, etc.) Mathematical models are used to estimate the dynamics of the grape berry moth population in order to optimize the timing of phytosanitary treatments.

Septoria tritici blotch (STB) of winter wheat (*Triticum aestivum* L.) is one of the most important disease worldwide (Eyal, 1999) in Europe (Ghaffary, 2011) (Miedaner, et al., 2013) (Kollers, et al., 2013) (Suffert, et al., 2011) as in France (Ghaffary, et al., 2011). The disease was first described in France in (Sprague, 1938) it is a foliar disease of wheat caused by the fungus *Mycosphaerella graminicola* (Fuckel) Schröter (anamorph *Septoria tritici*), causing average yield losses of about 1.5 t/ha in the absence of fungicides (Jørgensen, et al., 2008). The propagation of weather uncertainties within disease development models should allow for a better positioning of optimal treatment dates.

Both diseases are economically very important, and treatments are practically essential. In order to control the use of phytosanitary treatments, it is necessary to choose the date on which it should be applied. This is the purpose of this research.

The paper is organized as follows. Section 1 presents the general context of our study and the approach we used to take into account the uncertainty of the weather forecasts in crop protection management. Section 2 combines all the material and methods used for the purpose of this article, with a description of the NWP models and the DST we considered. In section 3 we illustrate the simulation plan and finally section 4 provides a summary on the uncertainty of the treatment dates. Conclusions and future works are given in section 5.

II) Materials and methods

1. Ensemble weather forecasts

For this work, we used IFS and IFS-EPS for deterministic and probabilistic forecasts respectively and ground-based weather stations from Météo-France for observations. The IFS model's horizontal resolution is 9km. The deterministic forecasts are computed twice a day (at 00 and 12 UTC) with a forecast range of 10 days. The probabilistic version uses 50 perturbed forecasts. It is run twice a day (at 00 and 12 UTC) up to 15 days (with a horizontal resolution of 18km) and a 32-day forecast is run once a week (with a horizontal resolution of 36km). An example of deterministic and probabilistic forecasts is given in Figure 1 for temperature and in Figure 2 for precipitation.

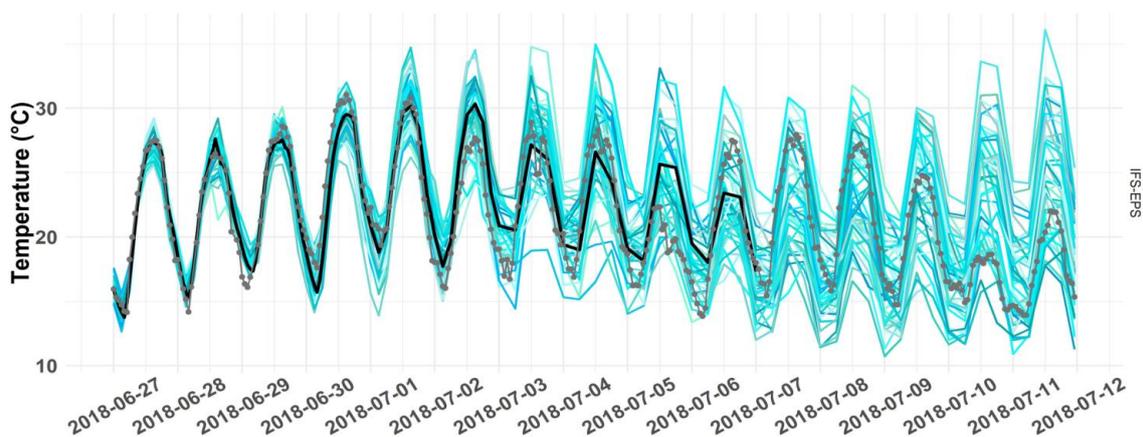


Figure 1: IFS and IFS-EPS weather forecasts for hourly temperature. Deterministic (black), ensemble forecasts (color, each line represents a perturbed forecast) and the observation (grey) of temperature at the location of Belval-sous-Chatillon. Initial time of the forecast is 27/06/2018.

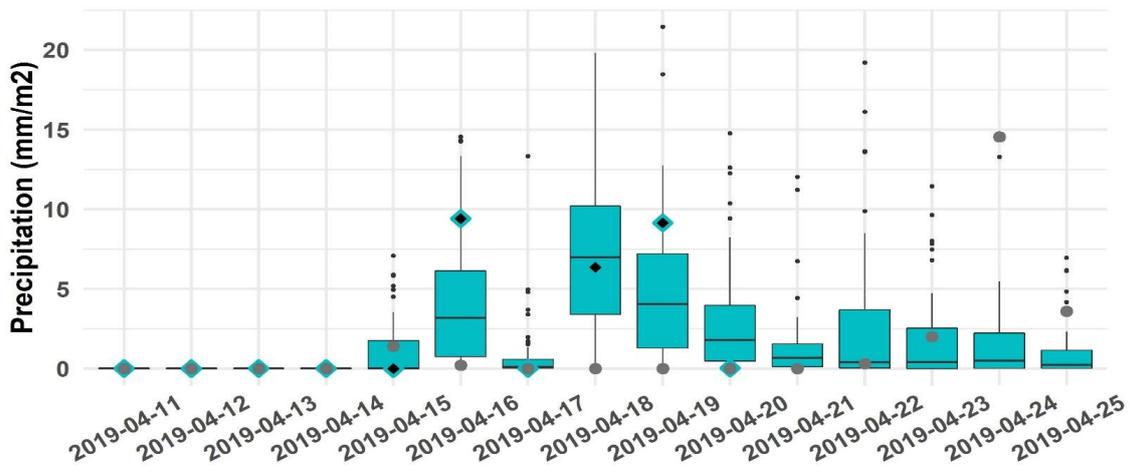


Figure 2: IFS and IFS-EPS weather forecasts for daily precipitation. Deterministic (black), ensemble forecasts (box plot) and the observations (grey) of precipitation at the location of La Chapelle Saint Sauveur. Initial time of the forecast is 11/04/2019.

2. Two DST used for decision making in crop protection

2.1. EVA to optimize the date of treatment for grape berry moth of vine

We use in this work the EVA model described by (Chavent, 1983) and used by the IFV-French Institute of Vine and Wine- Institut Francais de la Vigne et du Vin. EVA DST is a phenological model based on thermal conditions for development cycle. From the accumulation of hourly temperature, the model determines a development rate for each of the stages. This then allows for the calculation of the proportion and cumulation of the individuals in each stage over time. When it reaches a certain level, a fraction of the population moves to the next stage. This fraction, corresponding to a group of individuals, evolves in the next stage with the corresponding rate of development. The timing of treatment is based on a decision threshold of development. On the basis of discussions with experts, two thresholds are relevant to trigger the positioning of the treatment on egg accumulation: 2% egg accumulation for chemical treatment and 15% for a biocontrol treatment due to the lower persistence of the product. To perform a simulation with EVA for year N, the model must be initialized on September 20 of the previous year. It needs then a complete hourly temperature time series as an input to the model from the initialisation date to an ending date. This temperature time series are provided by observations from the initial date to the starting date of the forecast, and from weather forecasts from the starting date to the ending date of the forecast.

2.2. Septo-LIS® to optimize the date of treatment for Wheat Septoria

ARVALIS - Institut du végétal, French applied agricultural research organization dedicated to arable crops, has developed an epidemiological model (Gouache, et al., 2013) for the development of septoria on wheat integrated into a tool named Septo-LIS®, to position the first treatment against this disease (Gouache, et al., 2009). Septo-LIS® describes in a coupled way the dynamics of the disease and that of the plant. It is a biological model describing in particular the contamination received by the different leaf stages, taking into account daily meteorological inputs of precipitation and minimum and maximum temperatures. A decision rule based on the sum of the contamination of the upper leaves then makes it possible to predict the optimal date on which the first treatment will be initiated. To perform a simulation with Septo-LIS® for a given year, the model is initialized at the sowing date, and then needs daily weather time series of precipitation, maximal and minimal temperature from this date to an ending date. As in the previous case, the temperature time series are given by observations and then forecasts.

3. Simulation plan and strategy to quantify the uncertainty

A schematic representation of the propagation of weather forecast uncertainty through agronomic models is given on Figure 3. We conducted simulations for different locations in France (Figure 4), corresponding to eight of the most famous French vineyards, where the grape berry moth causes significant damages (left). In addition, we took into account different locations, eighteen in total, where wheat septoria is the most frequent (right). Meteorological variability is examined by taking into account all the sites throughout the French territory. For the purpose of this article, the years 2018 and 2019 are used for simulations.

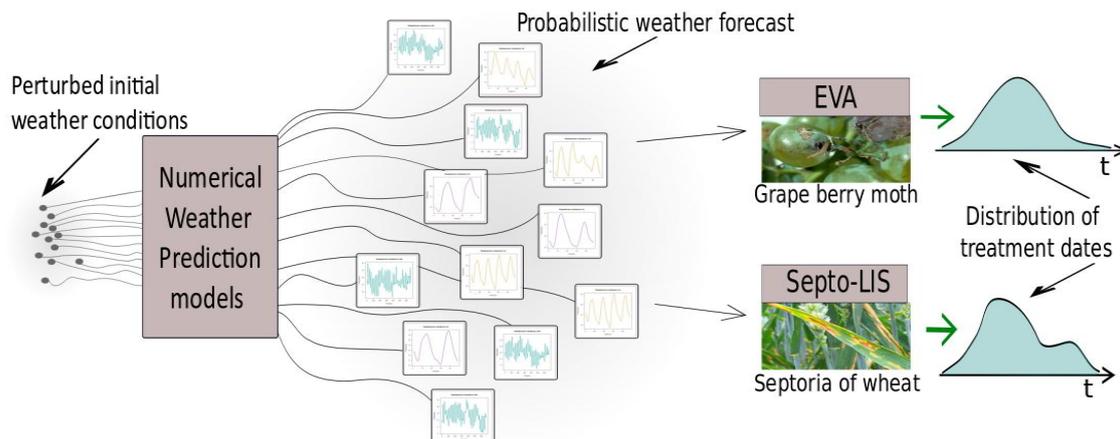


Figure 3: Ensemble prediction and propagation of weather forecast uncertainties through the both cases wheat septoria and grape berry moth.

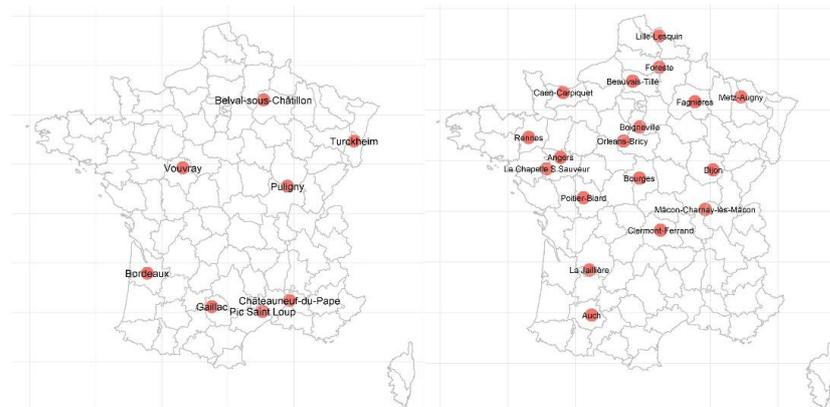


Figure 4: Left: French vineyards to be considered for EVA simulations, Right: Cereals fields in France to be considered for Septo-LIS® model simulations.

In order to summarize the uncertainty on the processing dates and propose a generalization on its evaluation over time, we were interested in calculating different statistics such as the minimum of the treatment dates obtained with the IFS-EPS weather forecasts (that correspond to the precocious treatment day), the maximum (that correspond to the latest treatment day), the median date as well as the standard deviation. Different quantiles were used to measure the inter-quantile distance for different locations and the years under consideration. The criteria often used are IC50 (the difference between the 75% and 25% quantiles), IC80 (the difference between the 90% and 10% quantiles) and IC90 (the difference between the 95% and 5% quantiles). These criteria are expressed in number of days. If the uncertainty

over the processing dates is spread over several days, the IC value will be higher. For the purpose of this article we chose to present the evaluation of IC90 in time considering all locations. In the following, we use the therm observation-based treatment date or reference treatment date for the treatment date obtained using weather observed data. Statistical studies such as "In how many cases did the reference treatment date fall into the distribution of dates associated with probabilistic prediction" will be used to complete our uncertainty analyses.

4. Results

We illustrate here the propagation of uncertainty in weather forecasting through the models of the decision support tools for the two case studies before demonstrating to the general results.

4.1. Grape berry moth

1. Weather uncertainty affects population dynamic, with an impact that depends on the dynamic timing

Using the IFS-EPS forecasts from Figure 1, the ensemble outputs from the EVA model are given on the Figure 5 for the proportion of the egg stage. This set of 50 weather forecasts gives a set of agronomic forecasts that predicts several evolutions of the population dynamics. Temperature uncertainty spreads to the grape berry moth population prediction, especially at long lead times. We also note that the value predicted by EVA with the observed weather data is included in the ensemble predictions.

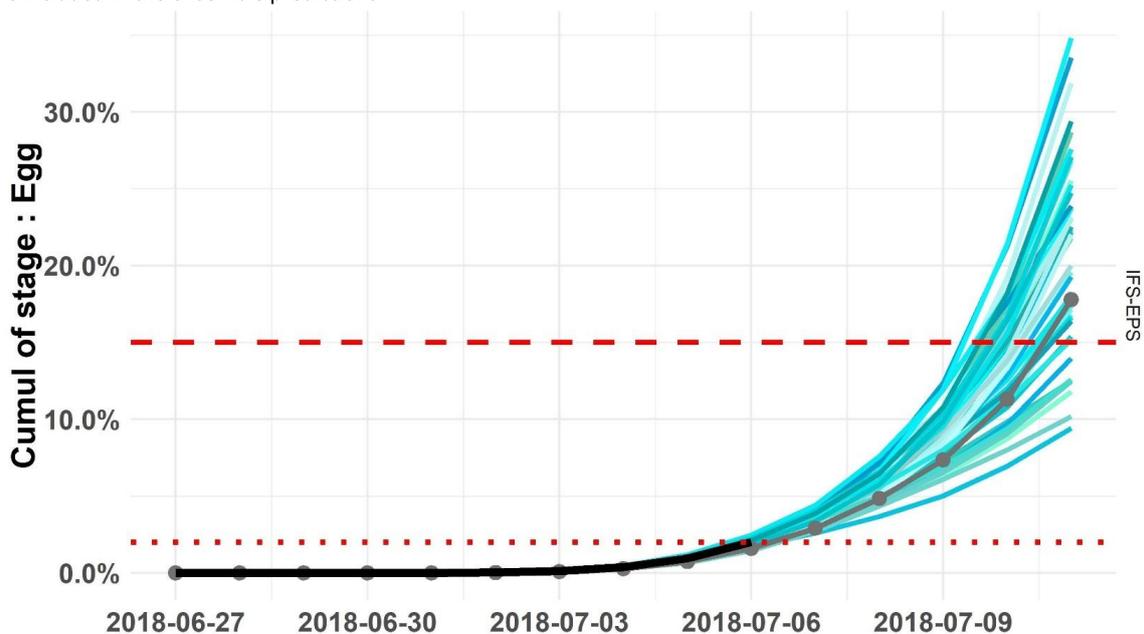


Figure 5: Total population of the egg stage at the location of Belval-sous-Chatillon forecasted over the period 27 June -11 July 2018, using forecasts from the IFS-EPS (color solid lines), IFS (black solid line) and an observation-based simulation (grey solid line). Forecast is initialized on 27 June 2018. Decision thresholds of 2 % (dotted line) and 15 % (dashed line) are highlighted.

2. From an ensemble of population dynamic to a distribution of treatment date

Applying a decision threshold on the egg stage of 2% (chemical treatment) and 15% (biocontrol), we obtain a distribution of the recommended dates to apply the treatment, rather than a single date. The treatment date associated to the simulation results obtained with the observed weather data is considered as a reference. We can notice on Figure 5 that the uncertainty about crossing the threshold of 2% is low, since the majority of members, including the deterministic forecast, cross the critical thresholds on the same day. On the contrary, for the 15% threshold the uncertainty is larger. This is related to the fact that the 15% threshold is reached for longer range forecasts for which the uncertainty is higher. The corresponding distribution of treatment dates (Figure 6) indicates that two treatment dates

10 and 11 July are almost equally likely, while the observed date is 11 July. In addition, seven forecasts do not reach the threshold at the end of the simulation period (they are mentioned as « beyond »). Note also that we cannot obtain a predicted treatment date associated to the deterministic forecast scenario because its maximum lead time is too short to reach the threshold of 15%. Simulations with more recent forecasts, i.e. with initial dates closer to the treatment date, may be useful to reduce the uncertainty on the two possible date.

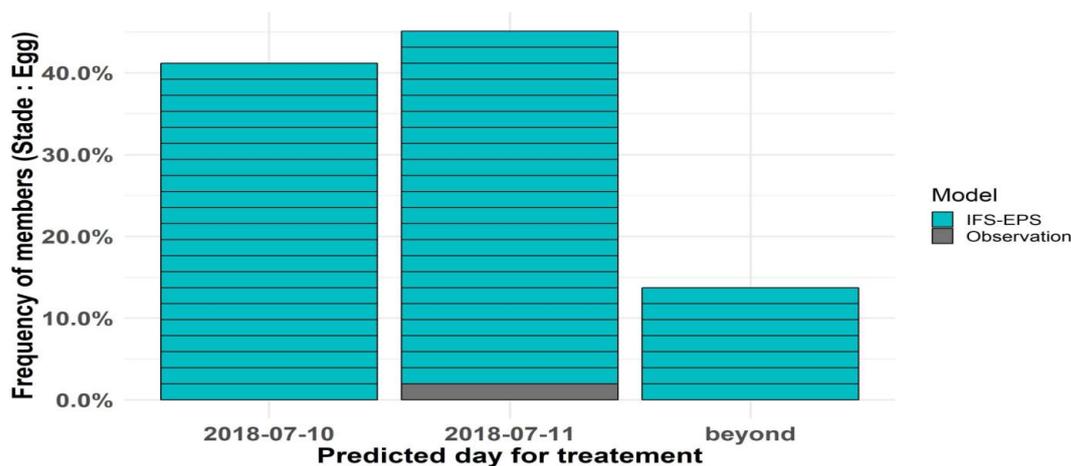


Figure 6: Histogram of treatment initiation dates derived from the EVA simulations coupled with the IFS-EPS, using the threshold 15% at the location of Belval-sous-Chatillon.

4.2. Wheat septoria, illustration of the variability according to location and years

1. Weather uncertainty leads to a distribution of recommended dates for treatment

As in the previous case, using ensemble forecasts provide a distribution of the recommended dates to apply the first treatment, rather than a single date (Figure 7). The uncertainty in the outputs of this model is introduced by temperature but a significant part of the uncertainty can also come from the precipitation, whose localization, timing and intensity are difficult to predict, especially for the long lead times where the uncertainty is larger. Even if the uncertainty is high, the long-range simulations using IFS-EPS is useful in order to predict the earliest date T1, as in Figure 7 the 17/04/2019 for a simulation starting on 11/04/2019. This date could be considered for early preparations for phytosanitary treatment but not for decision of treatment because the reference date predicted by the weather observed data is a week later. In fact, in this case the associated probability for the earliest date is very low and could be considered as a first information that can be updated each date to precise the range.

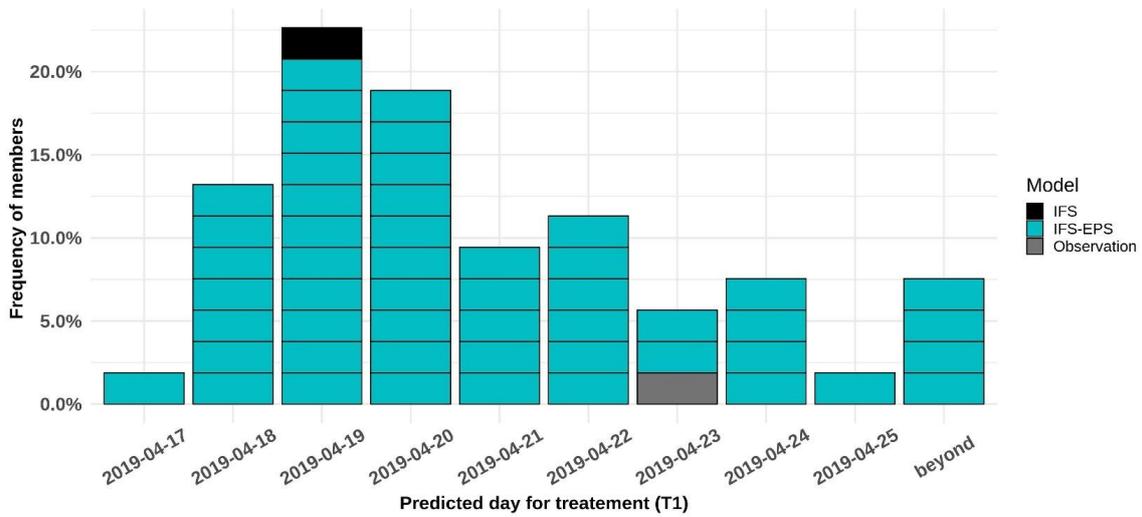
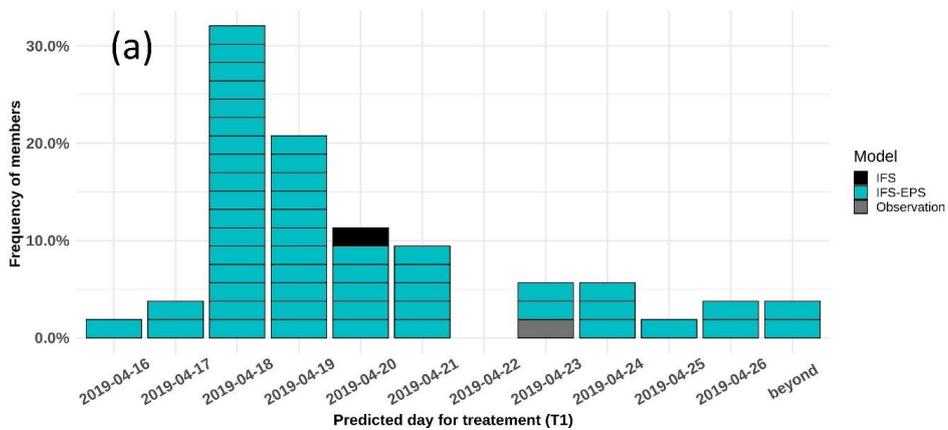


Figure 7: Histogram of treatment initiation dates derived from the Septo-LIS[®] simulations coupled with the IFS-EPS, initial time of the forecast is 11/04/2019 at the location of La Chapelle Saint Sauveur.

As the uncertainty in the output can evolve in time, new simulations should be run closer to the treatment date in order to refine the range of the possible date for optimal treatment. As an illustration, the treatment dates derived from forecasts initialized on four different dates are presented on Figure 8. As we get closer to the observation-based treatment date, the uncertainty about date T1 is getting more and more reduced. It is interesting to notice that deterministic and ensemble simulations started on 14/04 and beyond forecast a most probable date in agreement with the observed date. This thus allows for an anticipation of one week.



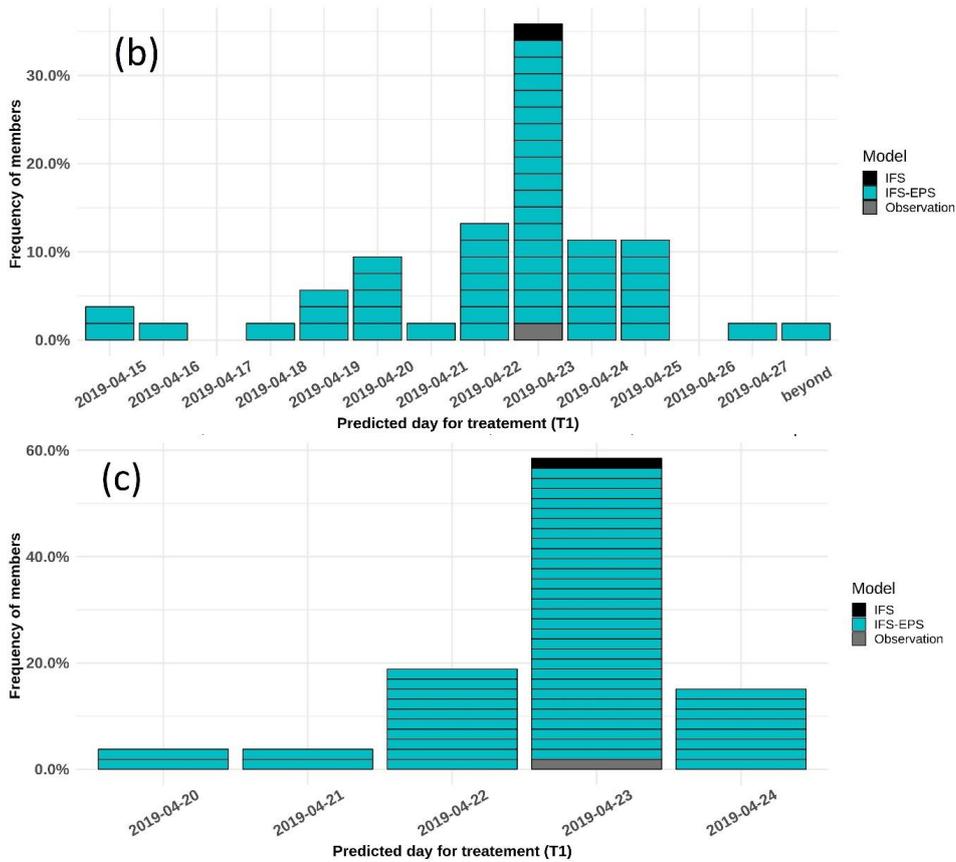
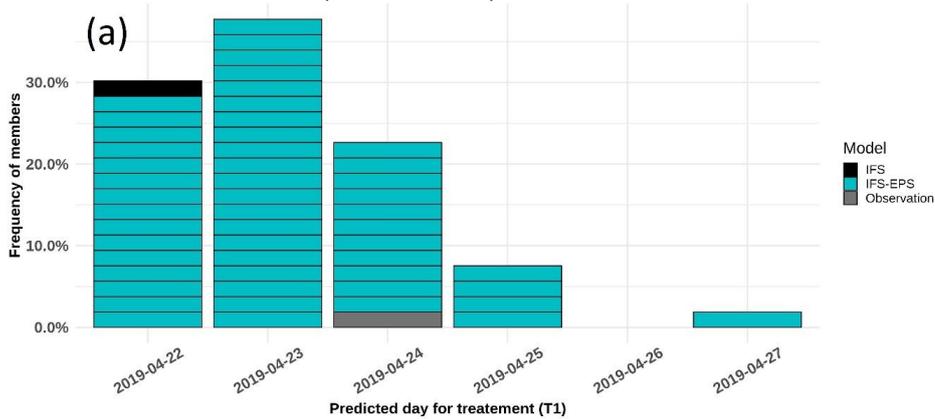


Figure 8: Histograms of treatment initiation dates derived from the Septo-LIS[®] simulations coupled with model IFS-EPS, IFS and observed weather data. Initial time of the forecast is: (a) 12/04/2019, (b) 14/04/2019 and (c) 16/04/2019 at the location of La Chapelle Saint Sauveur.

Uncertainty also depends on the spatial location. As an illustration Figure 9 presents the distributions of dates at two sites, derived from forecasts issued at the same date. At the location of Caen, the distribution of dates is spread over 6 days while at the location of Rennes it is spread over 11 days.



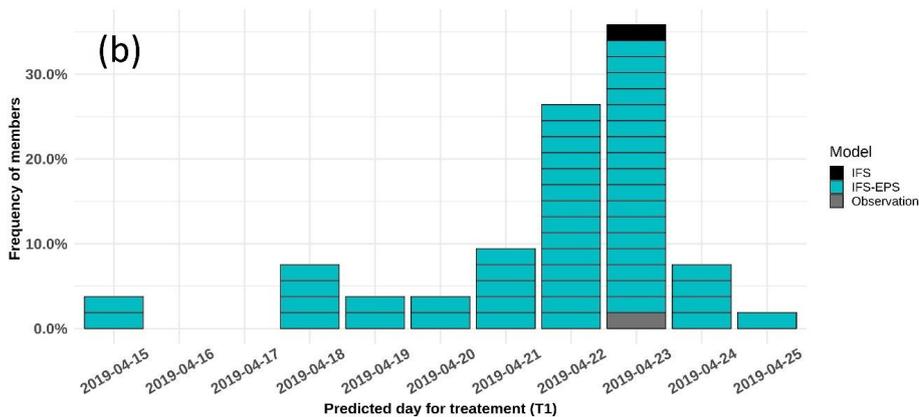


Figure 9: Histograms of treatment initiation dates derived from the Septo-LIS[®] simulations coupled with IFS-EPs initial time of the forecast is 15/04/2019 at the location of (a) Caen Carpiquet, (b) Rennes.

Conclusions: Forecasts with long lead-times are very important in agronomy in order to obtain an estimate well before the treatment date, so that the most appropriate phytosanitary treatment can be prepared. However, if the initial date of forecast is far away from the treatment date the uncertainty about the dates is higher. Hence, a compromise/trade-off must be found between earlier simulations that allow for more anticipation and later simulations with a reduced uncertainty. Similar simulations could be performed with the AROME-EPs and ARPEGE-EPs forecasts, but their shorter lead times would not allow the necessary thresholds to be exceeded and the treatment date is not reached in their corresponding forecast period.

4.3. Summary of uncertainty on treatment date

We have seen how uncertainty varies among year depending on the initial time of the forecast. In order to further investigate the time evolution of this uncertainty we examined inter-percentile range IC90 as a function of the interval between the observation-based treatment date and the forecast starting /initial date.

1. Grape berry moth

Figure 10 shows the IC90 for the years 2018 and 2019. When the initial time of the forecast is 3 to 4 days before the reference date the IC90 is 0. On the other hand, as we get further away from the observation-based treatment date the IC90 increases. When the initial time of the forecast is around one week before the reference treatment date the IC90 indicates that the uncertainty about treatment date is about one day. On the contrary, when the initial time of the forecast is between one and two weeks before the treatment date, then the IC90 can reach the value of two days. We are therefore able to estimate the median of the IC90 when the forecast is about 12, 13 days before the reference treatment date. It should be mentioned that the missing values of IC90 statistics at long lead times are due to the fact that in some cases there exist members that can not reach the threshold (Figure 5), which are interpreted as « beyond » in Figure 6. For instance, in Figure 10, we can see that for 2018, the 75% can be estimated when the initial time of the forecast used for the simulation is up to 11 days before the reference treatment date.

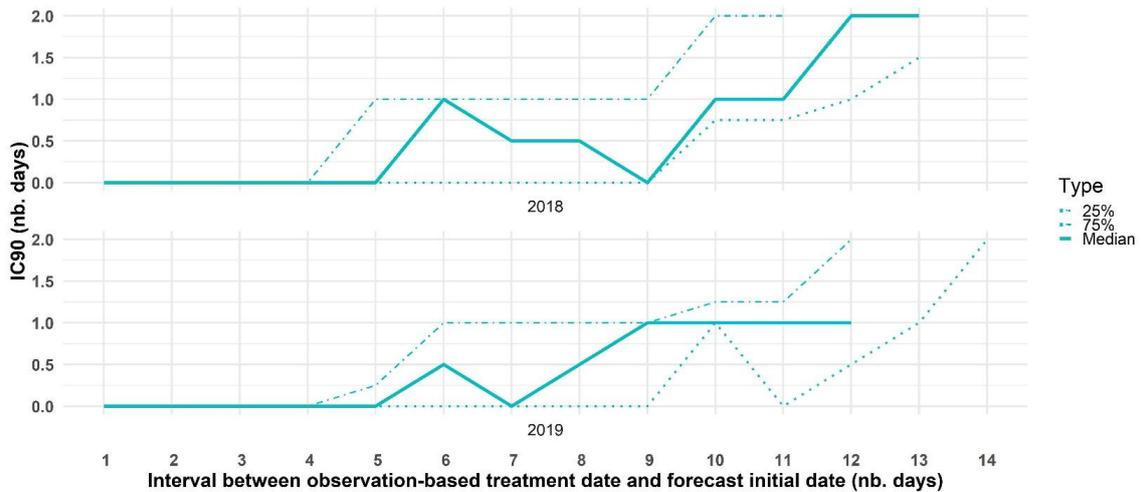


Figure 10: The median, the 25% and the 75% quantile of IC90 of treatment initiation dates (estimated for the 2nd generation with a threshold of 15%) derived from EVA simulations coupled with IFS-EPS for all site locations and for two different years as a function of the number of days between the reference treatment date and the initial time of the forecast.

In addition, in 83.2 % of the cases the reference treatment date falls into the distribution of dates associated to the probabilistic prediction.

2. Wheat septoria

We proceeded in the same way as in the case of EVA. Similarly to EVA, we observe that the value of the IC90 significantly increases when the initial time of the forecast is more than one week earlier than the reference treatment date (Figure 11), to reach a value about one week. We also notice that the IC90 values for Septo-LIS[®] are much larger than for EVA, which may be related to the impact of precipitation forecasts uncertainty. There is also an inter-annual variability, with a larger uncertainty in 2018 at short lead times.

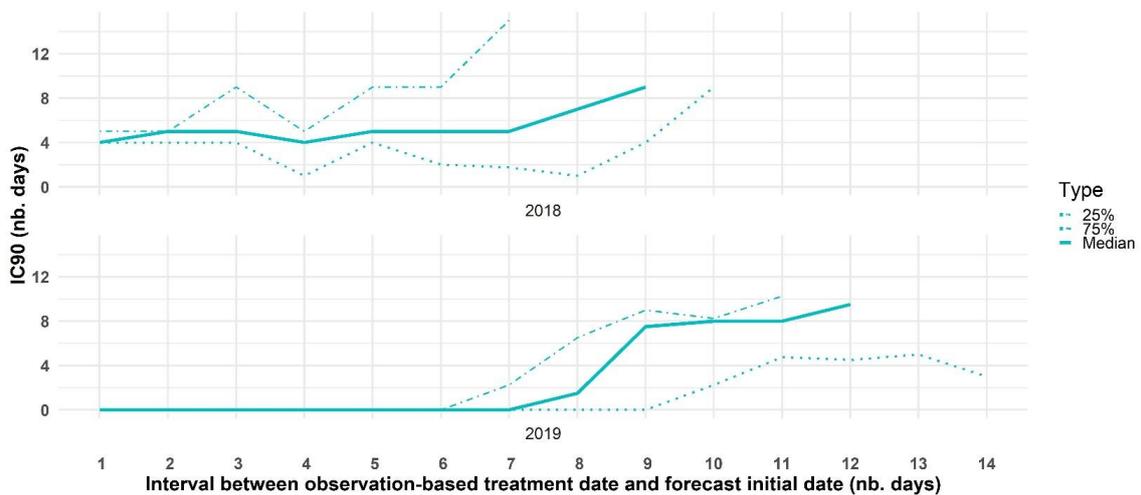


Figure 11: *The median, the 25% and the 75% quantile of IC90 of treatment initiation dates derived from Septo-LIS® simulations coupled with IFS-EPS for all site locations and for two different years as a function of the number of days between the reference treatment date and the initial time of the forecast.*

Furthermore, the reference treatment date belongs to the predicted distribution in 67.24% of the cases. This indicates that the prediction for Septo-LIS®, in addition to being more uncertain than for EVA, is also subject to larger errors.

Although these results are based on a limited number of locations and years, they suggest that the uncertainty on the estimated treatment date is rather small up to almost one week in advance for the two models. In addition, the uncertainty may be larger with the Septo-LIS® model because of its dependence on precipitation forecasts.

5. Discussion

Long-range forecasts are very important in agronomy to obtain an estimate well ahead of the treatment date, in order to be able to prepare the most appropriate phytosanitary treatment. However, if the initial date of the forecast is far from the observation-based treatment date, the uncertainty about the dates is greater. New simulations are carried out every day and thus make it possible to refine the interval of the possible date for optimal treatment. Depending on the results and the operational reality, it therefore remains to be specified until when it is possible to anticipate while providing useful information to agricultural stakeholders.

The uncertainty due to meteorological condition has in general a larger importance compared to other sources of uncertainty such as the model formulation (Rossing, et al., 1994). In this work, we considered that the model equations and tuning parameters are sufficiently accurate and therefore the model uncertainties are not taken into account (de Vries, 1994). However, the uncertainty in the agronomic modelling may also play an important role. For example, in the case of EVA, the model parameters were set to the values chosen by experts, but a range of possible values may give acceptable results. Hence, ensemble simulations that combine both weather forecast uncertainty and model uncertainty may provide a more skillful forecast.

For the purpose of this article raw weather forecasts have been used. It should be noticed that these raw forecasts are often subject to biases and dispersion errors, that can be reduced with appropriate post-processing techniques (Gneiting, 2014). The use of these calibrated forecasts could benefit to the agronomic simulations and will be considered in future works.

Another avenue for improvement may be the combination of the different available weather forecasts. For instance, the IFS-EPS (Palmer, 2019) used in this study could be combined with the French AROME-EPS (Raynaud, et al., 2016) and ARPEGE-EPS (Descamps, et al., 2015) models that provide a finer-scale information on short lead times. The design of temporally consistent scenarios from these three models and their impact on the agronomic simulations will be the subject of a forthcoming study.

Probabilistic forecasting provides a richer, but also more complex information than deterministic forecasting. As shown in this paper, one can derive from the distribution of treatment dates different scenarios, e.g., the earliest, the latest or the most probable treatment date. An optimized processing strategy is thus not straightforward and could involve for instance the risks and benefits associated with each decision. In addition, agronomists and decision-makers are mainly used to deal with a deterministic information. Hence a close collaboration between scientists and end-users should be encouraged to successfully include probabilistic forecasts in decision making and increase farmers' awareness of the benefits of using the DSTs for disease forecasts (Shtienberg, 2013).

6. Conclusions

The agronomic sector, particularly crop diseases and pests, is highly dependent on weather conditions. Decision support tools are proposed to take these weather conditions into account for disease forecasting. They estimate the development of pests and diseases and are very useful for farmers in making decisions regarding phytosanitary treatments. However, weather forecasts remain uncertain and can impact the decision about treatment date. Several weather services have developed probabilistic forecasts that provide an estimate of the forecast uncertainty. If the deterministic approach aimed to predict a single treatment date, the use of these probabilistic predictions through decision support tools makes it possible to obtain a distribution of treatment dates. In this article, we have shown the

potential of this probabilistic information in the two case studies for crop management and crop protection with application to the vine and wheat sectors. The processing date associated with the observed weather was then used as the reference processing date. The results showed that if the initial date if the probabilistic forecast date is far from the reference processing date, then the uncertainty in the processing dates remained significant. On the other hand, agronomic simulations that use weather forecasts with the initial date closer to the reference treatment date can be used to refine this information and reduce the uncertainty in the proposed treatment dates.

Acknowledgements

This work is a part of a Ph.D. thesis that is supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004, for #DigitAg “Digital Agriculture Convergence Lab” (www.hdigitag.fr/en). It is funded too by CASDAR funding of French Ministry of Agriculture for the project METEOPREC “the contribution of accurate forecasting to farmers” and the network Modeling and Data Analysis for Agriculture (www.modelia.org).

This article has greatly benefited from the support of my colleagues Pascal RAYNAUD and Frank SOUVERAIN. I would like to thank them for their collaboration in extracting ECMWF forecast data, which is time-consuming. I would also like to thank Sébastien PRATS for his willingness to help me and his contribution in learning and manipulating queries for data extraction from the Meteo-France forecasts database.

References

- Abbe, Cleveland, 1901. The physical basis of long-range weather forecasts. *Monthly Weather Rev.*, 12, 551-561.
- Aleksovska, I., Brun, F., Raynaud, L., Faivre, R., Raynal, M., et Deudon, O, 2018. Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures. *Végéphyt - 12e conf. int. sur les maladies des plantes - Tours (France)*.
- Bjerknes, Vilhelm, 1999. The problem of weather forecasting as a problem in mechanics and physics. *The Life Cycles of Extratropical Cyclones*. 1-4.
- Carlson, Gerald A, 1970. A decision theoretic approach to crop disease prediction and control. *American J. of Agric. Economics*. 2, 216-223.
- Chaulagain, B., Ian S., James M S., Clyde F., Richard N R., Philippe C. Emile R. , 2019. Weather-based predictive modeling of orange rust of sugarcane in Florida. *Phytopathology*. 1-38.
- Chavent, F, 1983. Etude de la dynamique des populations d'Eudémis de la vigne. Adaptation d'un modèle d'évolution aux conditions climatiques de Provence. *E.N.S.S.A.A. Dijon*.
- Cheung, J., Alan H., Jaap H., Adri M., et Jean-Louis B., 2015. Recommendations on trajectory selection in flight planning based on weather uncertainty. *Proc. 5th SESAR Innovation Days (SID2015), Bologna, Italy*. 1-8.
- Christ, E. H., Peter J. W., Guy D C., Violeta E. T., et Seth A.B., 2015. Using Precipitation Forecasts to Irrigate Cotton. *The J. of Cotton Sci.* 351-358.
- Cooke, B.M, Jones, D.G., Kaye, B., 2016. *The epidemiology of plant diseases*. Vol. 2. Springer.
- Courtier, P., Andersson, E., Heckley W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher M., Pailleux, J., 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly J. of the Royal Meteorol. Society*. 550, 1783-1807.
- Courtier, P., Thépaut J-N, et Hollingsworth A., 1994. A strategy for operational implementation of 4D-Var, using an

- incremental approach. Quarterly J. of the Royal Meteorol. Society 120, 519 ,1367-1387.
- de Vries, FWT Penning, 1994. Evaluation of simulation models in agriculture and biology: conclusions of a workshop.» *Agricul. Systems.* 4, 419-448.
- Descamps, L, Labadie, C., Joly, A., Bazile, E., Arbogast, P., et Cébron P.,2015. PEARP, the Météo-France short-range ensemble prediction system. Quarterly J. of the Royal Meteorol. Society. 690, 1671-1685.
- Drohan, P. J., Bechmannb, M., Budac, A., Djodjicd, F., Doodye, D., Duncanf, J. M., Ihog, A., Jordanh P., Kleinmanc, P. J., McDowell, R., Mellanderj, P.E., Thomask, I. A. and Withersl P.J.A.,2019. A global perspective on phosphorus management decision support in agriculture: Lessons learned and future directions. *J. of environ. quality.* 5, 1218-1233.
- Eyal, Zahir, 1999. The Septoria tritici and Stagonospora nodorum blotch diseases of wheat. *European J. of Plant Pathology.*7, 629-641.
- Fundel, V. J., Fleischhut, N., Herzog, S. M., Göber, M., et Hagedorn, R.,2019. Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. Quarterly J. of the Royal Meteorol. Society.
- Ghaffary, MS T., 2011. Efficacy and mapping of resistance to *Mycosphaerella graminicola* in wheat.
- Ghaffary, MS T., Robert, O., Laurent, V., Lonnet, P., Margalé, E., van der Lee, TA., Visser, R.G., Kema, G.H.,2011. Genetic analysis of resistance to *Septoria tritici blotch* in the French winter wheat cultivars Balance and Apache. *Theor. and Appl. Genetics.* 123, 741-754.
- Gneiting, Tilmann, 2014. Calibration of medium-range weather forecasts.» *European Centre for Medium-Range Weather Forecasts.*
- Gouache, D, et Couleaud, G.,2009. Timing of fungicide treatments against Septoria: stakes and potential of the "SeptoLIS" model. *Association Française de Protection des Plantes, 9ème conf. int. sur les maladies des plantes, Tours, France.*
- Gouache, D., Bensadoun, A., Brun, F., Pagé, C., Makowski, D., et Wallach, D.,2013. Modelling climate change impact on *Septoria tritici blotch* (STB) in France: accounting for climate model and disease model uncertainty. *Agric. and forest meteorol.* 170, 242-252.
- Jørgensen, L. N., Nielsen, GC, Orum, J. E., Jensen, J.E., et Pinnschmidt H.O.,2008. Integrating disease control in winter wheat—optimizing fungicide input. *Outlooks on Pest manag. (Res. Inform.)* 19, 206-213.
- Kollers, S., Rodemann, B., Ling J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M. W., et Röde, M. S., 2013. Genetic architecture of resistance to *Septoria tritici blotch* (*Mycosphaerella graminicola*) in European winter wheat. *Molecular breeding.* 2, 411-423.
- Leutbecher, M., et Palmer, T. N.,2008. Ensemble forecasting. *Journal of Computational Physics* 227. 7, 3515-3539.
- Lorenz, Edward N.,1963. Deterministic nonperiodic flow. *J. of the atmospheric sciences* 20. 2, 130-141.
- Miedaner, T., Zhao, Y., Gowda, M., Longin, C.F., Korzun, V., Ebmeyer, E., Kazman, E., Reif, J.C., 2013. Genetic architecture of resistance to *Septoria tritici blotch* in European wheat. *BMC genomics* 14, 1, 858.
- Moyer, M.M, Gadoury, D.M., Wilcox, W.F, et Seem, R. C.,2016. Weather during critical epidemiological periods and subsequent severity of powdery mildew on grape berries. *Plant disease* 100, 1, 116-124.
- Ogutu, G. EO, Franssen, W. HP, Supit, I., Omondi, P., et Hutjes, R. WA.,2018. Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts. *Agricul. and forest meteorol.* 250, 243-261.

- Palmer, Tim., 2019. The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Quarterly J. of the Royal Meteorol. Society.* 145, 12-24.
- Pertot, I., Caffi, T., Rossi, V., Mugnai, L., Hoffmann, Christoph GND, Grando, M.S., Gary, C., Lafond, D., Duso, C., Thiery, D., Mazzonia, V., Anfora, G., 2017. A critical review of plant protection tools for reducing pesticide use on grapevine and new perspectives for the implementation of IPM in viticulture. *Crop Protection.* 97, 70-84.
- Pietravalle, S, Shaw, M.W., Parker, S.R., et Van Den Bosch, F., 2003. Modeling of relationships between weather and *Septoria tritici* epidemics on winter wheat: a critical approach. *Phytopathology* 93, 10, 1329-1339.
- Pinson, P., Nielsen, H Aa, Madsen, H., et Kariniotakis, G., 2009. Skill forecasting from ensemble predictions of wind power. *Applied Energy.* 86.7-8, 1326-1334.
- Raposo, R., Wilks, D.S., Fry, W.E., 1993. Evaluation of potato late blight forecasts modified to include weather forecasts: a simulation analysis. *Phytopathology* 83, 1, 103-108.
- Raynaud, L., et Bouttier, F., 2016. Comparison of initial perturbation methods for ensemble prediction at convective scale. *Quarterly J. of the Royal Meteorol. Society* 142. 695, 854-866.
- Richardson, L. F., 2007. *Weather prediction by numerical process.* Cambridge University Press.
- Rossi, V., Caffi, T., et Salinari, F., 2012. Helping farmers face the increasing complexity of decision-making for crop protection. *Phytopathologia Mediterranea.* 457-479.
- Rossing, WAH, Daamen, RA, et Hendrix, EMT., 1994. Framework to support decisions on chemical pest control under uncertainty, applied to aphids and brown rust in winter wheat. *Crop Protection* 13. 1 (1994), 25-34.
- Rossing, WAH, Daamen, RA, et Jansen, MJW., 1994. Uncertainty analysis applied to supervised control of aphids and brown rust in winter wheat. Part 1. Quantification of uncertainty in cost-benefit calculations. *Agricul. systems* 44. 2, 419-448.
- Rossing, WAH, Daamen, RA, et Jansen, MJW., 1994. Uncertainty analysis applied to supervised control of aphids and brown rust in winter wheat. Part 2. Relative importance of different components of uncertainty. *Agricul. systems* 44. 4, 449-460.
- Shipp, J L., Clarke, N. D., 1999. Decision tools for integrated pest management. Dans *Integrated Pest and Disease Management in Greenhouse Crops.* 168-182.
- Shtienberg, D., 2013. Will decision-support systems be widely used for the management of plant diseases?. *Annual review of phytopathology* 51. 1-16.
- Sprague, Roderick., 1938. The status of *Septoria graminum*. *Mycologia* 30. 6, 672-678.
- Suffert, F., Sache, I., et Lannou, C., 2011. Early stages of *septoria tritici blotch* epidemics of winter wheat: build-up, overseasoning, and release of primary inoculum. *Plant Pathology (Wiley Online Library)* 60. 2, 166-177.
- Te Beest, DE, Paveley, ND., Shaw, MW, et Van Den Bosch, F., 2008. Disease--weather relationships for powdery mildew and yellow rust on winter wheat. *Phytopathology* 98. 5, 609-617.
- Thiéry, D., Louâpre, P., Muneret, L., Rusch, A., Sentenac, G., Vogelweith, F., Iltis, C. et Moreau, J., 2018. Biological protection against grape berry moths. A review. *Agronomy for sustainable development* 38, 2, 15.
- Thiéry D. Carton Y. Vidal C. Gauthier N. et al, 2013. Histoire de l'installation de quelques ravageurs.
- Zhao, T., Wang, Q. J, Schepen, A., et Griffiths, M., 2019. Ensemble forecasting of monthly and seasonal reference crop evapotranspiration based on global climate model outputs. *Agricul. and forest meteorol.* 264. 114-124.

**Article Weather and forecasting -
version soumise (soumis en février
2021)**

1 **Design and evaluation of calibrated and seamless ensemble weather**
2 **forecasts for crop protection applications**

3 Ivana ALEKSOVSKA * and Laure RAYNAUD

4 *CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France*

5 Robert FAIVRE

6 *INRAE, UR875 MIAT, 31326 Castanet-Tolosan, France*

7 François BRUN

8 *ACTA-French Technical Institute, UMR1248 AGIR, 31326 Castanet-Tolosan, France*

9 Marc RAYNAL

10 *Institut Français de la Vigne et du Vin (IFV), UMT SEVEN, Bordeaux, France*

11 *Corresponding author: ivana.aleksovska@meteo.fr

ABSTRACT

12 Agriculture is a highly weather-dependent activity, climatic conditions impact both directly crop
13 growth and indirectly diseases and pests developments causing yield losses. Weather forecasts are
14 now a major component of various decision-support systems that assist farmers to optimize the
15 positioning of crop protection treatments. However, properly accounting for weather uncertainty
16 in these systems still remains a challenge. In this paper, three global and regional ensemble
17 prediction systems (EPSs), covering different spatio-temporal scales, are coupled to a temperature-
18 driven developmental model for grape vine moth in order to provide probabilistic forecasts of
19 treatment dates. It is first shown that a parametric post-processing of the EPSs significantly
20 improves the prediction of treatment dates. Anticipating the need for phytosanitary treatments also
21 requires seamless weather forecasts from the next hour to sub-seasonal time scales. An approach
22 is presented to design seamless ensemble forecasts from the combination of the three EPSs used.
23 The proposed method is able to leverage the increased performance of high-resolution EPS at short
24 ranges, while ensuring a smooth transition toward larger-scale EPSs for longer ranges. The added
25 value of this seamless integration on agronomic predictions is, however, difficult to assess with the
26 current experimental setup. Additional simulations over a larger number of locations and years
27 may be required.

28 **1. Introduction**

29 Numerical Weather Prediction (NWP) models are now widely used for operational forecasting
30 over a wide range of spatio-temporal scales (Bauer et al. 2015). In addition, the inherent uncertainty
31 of NWP forecasts has motivated the development of probabilistic forecasting based on ensemble
32 prediction systems (EPSs) (Palmer 2019). Typically, a set of perturbed forecasts is designed in
33 order to account for the different sources of error, regarding for instance initial conditions and the
34 model formulation. Originally developed for global models (Molteni et al. 1996; Leutbecher and
35 Palmer 2008), such systems have more recently become popular at the convective scale for short
36 ranges (Bouttier and Raynaud 2018a; Hagelin et al. 2017). Several downstream applications also
37 started to use these probabilistic forecasts (Fundel et al. 2019), including in particular agriculture
38 (Christ et al. 2015), power grid management (Pinson et al. 2009), aircraft routing (Cheung et al.
39 2015), hydrology (Bellier et al. 2017, 2018) and wildfires (Worsnop et al. 2019).

40 In agriculture, pests develop according to weather conditions (Moyer et al. 2016). Crop manage-
41 ment operations such as applications of phytosanitary treatments are common practices to control
42 diseases and pests, but they need to be properly positioned and applied according to some criteria
43 that depend on the development of the pest population (Cooke et al. 2006). The precision in
44 their positioning is even more important for example for biocontrol solutions, which use living
45 organisms to prevent or reduce damages caused by pests, because of their lower levels of efficiency
46 (Shipp and Clarke 1999). For that purpose, different Decision Support Tools (DSTs) have been
47 developed to predict the evolution of pest dynamics and support farmers in their decision making
48 (Pertot et al. 2017). They are complementary to field observations and they provide an opportunity
49 to anticipate the situation and to access variables that are difficult to observe. Most of these tools
50 require skillful extended-range weather forecasts, that are most of the time given by deterministic

51 models (Olatinwo et al. 2011, 2012). Probabilistic forecasts could provide valuable additional
52 information to support decision making, and hence contribute to a sustainable use of treatments
53 through minimal and timely applications.

54 However, raw ensemble weather forecasts may have systematic errors regarding bias and dis-
55 persion, that can have a negative impact on the subsequent DSTs outputs. Several statistical
56 post-processing methods have been proposed to reduce these errors, that basically aim at estimat-
57 ing a statistical relationship between the observations and the predicted probability distributions.
58 Once this relationship is established, it can be used to generate calibrated probabilistic forecasts
59 from ensemble outputs. Post-processing methods can be divided into two categories: paramet-
60 ric and non-parametric approaches. Parametric approaches are based on the assumption that the
61 calibrated predictive distribution of a weather variable is given by a parametric density function,
62 whose parameters depend on ensemble members and are fitted over a training period (Gneiting
63 2014). Standard parametric approaches include NR (Non-homogenous Regression)/EMOS (En-
64 semble Model Output Statistics) methods (Gneiting and Raftery 2005) and BMA (Bayesian Model
65 Averaging) methods (Raftery et al. 2005). On the other hand, non-parametric methods remove any
66 assumption on the variable distribution. One example is the quantile regression forest proposed by
67 Taillardat et al. (2019).

68 Calibrated ensemble members can be generated from the post-processed distributions, with a ran-
69 dom or quantile sampling for instance. However, since the calibration is performed independently
70 at each location and lead time, the spatio-temporal structure of ensemble members is lost. These
71 dependencies, which are crucial to many users, can be recovered with copula methods (Scheffzik
72 2011), that perform a reordering of the sampled values based on a given dependence structure.
73 For instance, the Ensemble Copula Coupling (ECC) approach (Scheffzik et al. 2013) uses the rank

74 structure of the raw ensemble members, while the Schaake shuffle approach (Clark et al. 2004)
75 uses the rank structure of past observations.

76 As mentioned previously, DSTs rely on seamless weather trajectories, from the next hour to several
77 days ahead, in order to make early decisions about crop protection (Calanca et al. 2011). Seamless
78 members can be obtained from an extended-range EPS such as the ECMWF EPS (Pappenberger
79 et al. 2013), but it may also be interesting to integrate higher-resolution EPSs available on shorter
80 time scales for a better representation of small-scale phenomena. To the best of the authors
81 knowledge, seamless probabilistic multi-source forecasting is still an open question. In this paper,
82 the seamless combination of three EPSs with different spatio-temporal scales will be considered as
83 a simple concatenation problem, that aims at matching short-range forecasts from a fine resolution
84 EPS to coarse resolution forecasts from larger-scale EPSs. A similar approach was adopted by
85 Wetterhall and Di Giuseppe (2018) in order to merge sub-seasonal to seasonal ECMWF ensemble
86 forecasts for hydrological applications. Seamless integration methods should address two major
87 challenges : ensuring smooth transitions (i.e. without temporal discontinuities) between the
88 different EPSs considered, and providing enhanced performances with respect to the reference
89 ECMWF EPS, at least at short ranges.

90 The goal of this paper is to examine the impacts of post-processing and seamless EPS merging
91 on both weather and agronomic outputs. Standard EMOS and ECC methods are used to generate
92 calibrated distributions, while different strategies for designing seamless ensemble forecasts are
93 proposed and evaluated in a univariate setting. The outline is as follows. Section 2 describes the
94 NWP models and the agronomic use case, as well as the post-processing and seamless integration
95 methods. Section 3 presents the characteristics and performances of seamless calibrated ensemble
96 forecasts on 2-meter temperature, while the impact on agronomic outputs is discussed in section
97 4. Finally, section 5 provides conclusions and future works.

98 2. Material and Methods

99 *a. Ensemble weather forecasts*

100 For the purpose of this work, three operational EPSs are considered: the convective-scale
101 AROME-EPS developed at Météo-France, that covers a Western Europe domain (Raynaud and
102 Bouttier 2016), the global French ARPEGE-EPS (Descamps et al. 2015) and the global IFS-EPS
103 of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Palmer 2019). These
104 models account for both initial and model uncertainties, and the AROME-EPS uses selected
105 ARPEGE-EPS members as lateral boundary conditions (Bouttier and Raynaud 2018b). The main
106 characteristics of these EPSs are given in Table 1.

107 *b. Ensemble model output statistics*

108 For post-processing ensemble weather forecasts, the NR/EMOS univariate statistical method is
109 used. Let us denote y the weather variable of interest and x_1, \dots, x_N the N forecasts corresponding
110 to the N members of a given EPS. The NR/EMOS method assumes that the predictive distribution
111 of y is a parametric density function p whose parameters depend on ensemble output statistics
112 (Gneiting and Raftery 2005):

$$Y | x_1, \dots, x_N \sim p(Y | x_1, \dots, x_N). \quad (1)$$

The parameter of interest in this study is the 2-meter temperature, which is commonly post-processed using a Gaussian distribution (Gneiting et al. 2005). Its mean μ and variance σ^2 are assumed affine functions of the ensemble mean \bar{x} and the ensemble variance s^2 respectively :

$$\mu = \alpha_0 + \alpha_1 \bar{x}, \quad \sigma^2 = \beta_0^2 + \beta_1^2 s^2,$$

113 where coefficients α_0 , α_1 , β_0 and β_1 are estimated from pairs of past raw forecasts-observations, by
114 minimizing the mean Continuous Ranked Probability Score (CRPS) of the calibrated distribution
115 over a training period, as proposed by Gneiting et al. (2005). In the remainder of the paper, the
116 considered training period is a sliding window of 60 days prior to the EPS starting date (Hemri et al.
117 2014). The observations used are 2-meter temperature measurements from the French real-time
118 weather observation network.

119 A finite N -sample of calibrated members $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$ is then obtained by randomly sampling
120 the post-processed distribution at each lead time. These members are finally re-ordered with the
121 ECC matching procedure according to the raw ensemble template (Scheffzik et al. 2013). This step
122 ensures that the new trajectories present a realistic temporal consistency across lead times. Spatial
123 and multivariate dependencies could also be addressed with the same approach, however they
124 are not considered in this study since the agronomic use case considered only requires univariate
125 forecasts at independent locations.

126 *c. Seamless ensemble forecasts*

127 In this section, a simple design of seamless ensemble members, that performs a concatenation
128 of the systems AROME-EPS, ARPEGE-EPS and IFS-EPS over different spatio-temporal scales, is
129 described. Given the characteristics of these EPS, the seamless forecasts are started at 00 UTC and
130 constructed with the following pragmatic approach: the seamless ensemble is composed of the 12
131 AROME-EPS members only over the forecast period 0-48h, for longer ranges each AROME-EPS
132 member is matched to an ARPEGE-EPS and/or IFS-EPS member according to some assignment
133 rules described below. This configuration is motivated by the higher performance of the AROME-
134 EPS for its timeframe.

135 The simplest assignment is to randomly select, without repetition, members in the target sample,
 136 which contains members of ARPEGE-EPS and IFS-EPS, similarly to (Wetterhall and Di Giuseppe
 137 2018). This method, hereafter denoted random neighbor (RN), will be used as a naive benchmark
 138 for the evaluation of more advanced strategies.

Another approach is to minimize the distance between matched forecasts in order to ensure a smooth transition between lead times. Given an appropriate distance measure d , d_{ij} denotes the distance between the i^{th} member of sample 1 (with size N) and the j^{th} member of sample 2 (the target sample, with size M). The nearest neighbor approach (hereafter denoted NN) assigns to each member of sample 1 the closest member of sample 2:

$$j_i^* = \arg \min_j \{d_{ij}\}, i = 1, \dots, N.$$

139 This method thus performs an individual optimization for each member. One drawback is that it
 140 allows for member replication, several members of sample 1 can be matched to the same member
 141 of sample 2, thus leading to an effective ensemble size smaller than N .

On the other hand, the Hungarian method (hereafter denoted HU) uses the Kuhn-Munkres algorithm (Kuhn 1955) to find the optimal bijective mapping between two samples, by minimizing the total distance of the assignment:

$$j^* = \arg \min_j \left\{ \sum_{i=1}^N d_{ij} \right\}.$$

142 Contrary to the nearest neighbor approach, this method does not allow for member repetitions.

143 Two important aspects of this configuration may be mentioned. First, the size of the seamless
 144 ensemble is constrained by the AROME-EPS size (12 members), secondly the matching procedure
 145 between high-resolution and lower resolution members is applied one or two times, depending
 146 on the selected members at 48h. If an AROME-EPS member is connected at 48h lead time to
 147 an IFS-EPS member, the matching procedure is only applied once for that member because IFS

148 provides forecasts up to 15 days ahead. If it is connected to an ARPEGE-EPS member then the
149 matching procedure will be repeated because ARPEGE provides forecast up to 96h only. This
150 second assignment will then perform the connection between this ARPEGE-EPS member and an
151 IFS-EPS member.

152 Since the agronomic use case considers 2-meter temperature forecasts at independent locations,
153 the merging of EPS members applies to time series. The metric d chosen is then the Dynamic Time
154 Warping (DTW) distance, initially proposed by Berndt and Clifford (1994), since it is well suited
155 for measuring the similarity between two time series. The DTW performs an optimal nonlinear
156 alignment between the two sequences, which can provide a more intuitive similarity measure than
157 the Euclidian distance since it allows for similar shapes to match even if they are out of phase.
158 The type of the warping window, as well as its width, which controls the amount of temporal shift
159 allowed, are tuning parameters. In the following, the Sakoe-Chiba window is used with a width
160 of one hour (Sakoe et al. 1990). The DTW is implemented using the R algorithm proposed in
161 (Giorgino et al. 2009) and applied over a period W prior to the merging time t_{merge} . This matching
162 criterion is thus only based on the information prior to the connection time. Other strategies using
163 information posterior to the connection time could also be considered but they are beyond the scope
164 of this paper.

165 The three assignment procedures RN, NN and HU are applied at $t_{merge} = 48h$ using calibrated
166 AROME-EPS members as sample 1 and calibrated ARPEGE-EPS+IFS-EPS members as sample
167 2. If a second assignment is necessary it is performed at $t_{merge} = 96h$, the ARPEGE-EPS members
168 selected at 48h are used as sample 1 and the remaining IFS-EPS members (i.e., those not chosen at
169 48h) are used as sample 2. Note that all members, including the control members of both models,
170 are considered for the matching.

171 *d. Probabilistic verification scores*

172 The verification of ensemble forecasts is performed with a comprehensive set of standard metrics,
173 including bias and Root Mean Squared Error (RMSE) of the ensemble mean and probabilistic scores
174 presented below. Note that observation errors are not accounted for in the scores computation.

175 (i) *Continuous Ranked Probability Score (CRPS)* The CRPS is well suited to the verification of
176 probabilistic forecasts of a continuous variable (Candille and Talagrand 2005; Candille et al. 2007).
177 Let us denote F the cumulative distribution function (CDF) associated with an ensemble forecast
178 and $\mathbf{1}_{\{y \leq x\}}$ the CDF of a scalar observation y , the CRPS is defined as the quadratic difference
179 between the forecast and observed CDFs:

$$\text{CRPS}(F, y) = \int_{\mathbb{R}} (F(x) - \mathbf{1}_{\{y \leq x\}})^2 dx. \quad (2)$$

180 The CRPS measures the overall performance of an ensemble forecast, the smaller the score, the
181 better the performance.

182 (ii) *Receiver Operating Characteristics (ROC)* The ROC curve is used to determine how well a
183 forecast discriminates between events and non-events (Hanley and McNeil 1982). Given a binary
184 event, the ROC curves plots the hit rate against the false alarm rate using increasing probability
185 thresholds to make the yes/no decision. The area under the ROC curve (denoted hereafter AUC)
186 summarizes the forecast skill. An area of 0.5 indicates no skill, the closer the AUC is to 1 the better
187 the forecast skill.

188 (iii) *Reliability diagram* For a given binary event, the reliability diagram plots the observed
189 frequency of the event against the binned predicted probabilities. A forecast is reliable if the
190 observed and predicted probabilities are consistent, in other words, if the reliability diagram is
191 close to the diagonal.

192 (iv) *Observation coverage* The observation coverage measures the percentage of observations
193 that fall in the predicted range. The expected observation coverage percentage for a perfectly
194 reliable ensemble of size N is calculated as $\frac{N-1}{N+1} \times 100\%$.

195

196 Statistical significance of scores differences between two forecasts is assessed with the Wilcoxon
197 signed-rank test (Wilcoxon et al. 1970). Wilcoxon tests are performed using the significant level
198 $\alpha = 0.05$.

199 *e. The agronomic use case*

200 In the French vineyard, two species of insects of the Lepidoptera's order, *Eupoecilia ambiguella*
201 and *Lobesia botrana*, known as grape vine moth, are responsible for quantitative and qualitative
202 damages (Thiéry et al. 2013). These insects evolve in four successive stages of development
203 and may have several reproduction cycles, usually 3 and up to 5 per year according to climatic
204 conditions and mainly temperatures. Second and third cycles usually are the most critical in terms
205 of damages. In order to control grape vine moth, winegrowers have three types of treatments at
206 their disposal: chemical products, biocontrol tools (trichograms and *Bacillus thuringiensis* toxins)
207 (Thiéry et al. 2018) or sexual confusion. The correct positioning of the treatment depends on
208 the mode of action of its active ingredient in response to the target (eggs, larvae, etc.). Since the
209 number and precise timing of the development cycles are highly dependent on weather conditions
210 and can be difficult to predict, mathematical models are often used in addition to field monitoring
211 to describe their evolution and predict their attainment.

212 In this work we consider the EVA model described (in French) by Chavent (1983), and oper-
213 ationally used by the French Institute for Vine and Wine. EVA is a phenological model based
214 on the thermal conditions of the grape vine moth development cycles. From the accumulation of

215 the hourly temperature, the model calculates the proportion and accumulation of individuals at
216 each stage over time. When it reaches a certain level, a fraction of the population moves on to
217 the next stage. The timing of treatment is based on a developmental decision threshold. Based
218 on discussions with experts, the two thresholds relevant to trigger the positioning of the treatment
219 on eggs accumulation are commonly fixed at 2% for ovicide chemical treatments and 15% for
220 biocontrol treatments (due to the lower persistence of the product).

221 EVA simulations consist of an analysis step and a prediction step. In the analysis step the model
222 is run using time series of observed hourly temperatures from an initialization date to the forecast
223 starting date. This procedure then provides the initial conditions for the prediction. In the prediction
224 step the model is integrated using hourly temperature forecasts. A schematic representation for the
225 analysis and the prediction steps is shown on Figure 1. Ensemble EVA simulations are obtained
226 by running the prediction step with temperature forecasts from EPS members or from the seamless
227 members, while initial conditions are the same for all EVA members. Note that hourly temperature
228 is a direct output from AROME-EPS, while ARPEGE-EPS and IFS-EPS outputs are linearly
229 interpolated to a 1-hour frequency.

230 A probabilistic evaluation of the EVA ensemble forecasts is performed with the same measures
231 used in NWP, applied to the predicted distribution of recommended treatment dates. It is important
232 to note that an observation-based treatment date is calculated by running the EVA model with
233 observed temperature over the entire season. This is the standard operational practice because field
234 observations of the grape vine moth's stages series is often difficult to obtain. In the remainder of
235 the paper results will be presented for the positioning of phytosanitary treatments on the second
236 development cycle using the 15% threshold.

237 *f. Experimental setup*

238 Agronomic simulations and evaluations with the EVA model are performed for the years 2018
239 and 2019 at eight sites that correspond to famous French vineyards where the grape vine moth
240 causes significant damages (Figure 2). It is assumed that there is no spatial dependency between
241 these locations. Initialization date for EVA simulation is set on 20th September of the previous
242 year. From this initialization date to the starting date of the forecast, the model is integrated using
243 temperature measurements from the French real-time meteorological observation network. EVA
244 forecasts are run in the period ranging from 1st April to 19th September (for years 2018 and 2019),
245 when the impact of the grape berry moth is the most significant.

246 Three additional sites corresponding to wheat fields that can be affected by tritiform septoria
247 (STB), a major disease in Europe of winter wheat (Miedaner et al. 2013; Kollers et al. 2013; Suffert
248 et al. 2011; Ghaffary 2011), have been considered for the purpose of meteorological evaluation.
249 These 11 locations provide a wide variety of meteorological conditions in different parts in France.
250 The verification of calibrated and seamless 2-meter temperature ensemble forecasts is performed
251 over a one-year period from the 1st June 2018 to the 31st May 2019, and uses observations from
252 the French real-time meteorological network. Corresponding AROME-EPS, ARPEGE-EPS and
253 IFS-EPS forecasts are extracted from the operational archive at the nearest model gridpoints to the
254 agronomic locations.

255 Seamless forecasts are designed using the DTW distance computed over the last $W = 7$ hours
256 before the merging time, following the performances obtained with different W values between 1
257 and 48 hours.

258 **3. NWP Results**

259 In this section we first present the impact of statistical post-processing of temperature forecasts,
260 and we then discuss the characteristics and probabilistic performances of the different seamless
261 temperature forecasts.

262 *a. Verification of calibrated forecasts*

263 Figure 3 presents the bias and RMSE for 2-meter temperature ensemble mean forecasts over a one
264 year period, before and after post-processing. As expected, the parametric calibration procedure
265 allows for a noticeable improvement of the three EPS performances over the entire forecast range.
266 These calibrated ensembles are used in the remainder of the paper to compute seamless forecasts.

267 *b. Justification of seamless design*

268 We recall that the proposed strategies for building seamless forecasts are based on two main
269 choices: using the AROME-EPS only over the first 48 hours and selecting subsequent ARPEGE-
270 EPS and IFS-EPS members based on their proximity to AROME-EPS members over the last 7
271 hours (corresponding to forecast ranges 42 to 48 hours). We provide here some objective results
272 to justify this configuration.

273 Figure 4 presents the CRPS for the three calibrated EPSs over a one-year period. The AROME-
274 EPS is shown to perform significantly better up to around 20h lead time. From 24h to 48h ranges,
275 the AROME-EPS is better than ARPEGE-EPS and close to or slightly worse than the IFS-EPS.
276 These results thus support the idea of opting for the AROME-EPS over its timeframe. Over the
277 forecast window 48-96h the ARPEGE-EPS performs worse than the IFS-EPS, despite its higher
278 resolution. This thus justifies the choice of considering both ARPEGE-EPS and IFS-EPS members
279 for the connection at 48h, instead of ARPEGE-EPS members only.

280 *c. Characteristics of seamless forecasts*

281 *(i) Origin of selected members at 48h* As a first diagnostic, Table 2 presents the percentage of
282 ARPEGE-EPS and IFS-EPS members selected for the connection at 48h, for the three seamless
283 designs and over different periods. With the random draw, 59.3% of the selected members come
284 from the IFS-EPS, because of the larger size of this ensemble. On the other hand, the distance-
285 based assignments provide more balanced selections and some seasonal variations are observed.
286 In spring and summer the IFS-EPS is more frequently chosen, while the ARPEGE-EPS is slightly
287 preferred in winter. In autumn, both EPSs are almost equally likely. These variations reflect the
288 mean distances between the AROME-EPS members and the ARPEGE-EPS and IFS-EPS members.
289 It has been verified for instance that in winter the ARPEGE-EPS members are on average closer
290 to the AROME-EPS members. Since the ARPEGE-EPS is used as lateral boundary conditions for
291 the AROME-EPS, one could have expected a higher representation of ARPEGE-EPS members in
292 this 48h-selection. However, it is likely that 2-meter temperature mainly results from small-scale
293 physical processes and only slightly depends on the large-scale coupling.

294 It can finally be noticed that the percentages obtained with the NN and HU methods are very
295 similar. In addition, differences in the NN and HU selections, which are only due to the assignment
296 rule (the distance criterion is the same), are rather small with an intersection between NN and HU
297 selected members equal to 87% on average for this 48h-connection.

298 *(ii) Repetition of members for the NN strategy* As explained in section 2c, a major limitation
299 of the nearest neighbor approach is the possible repetition of selected ARPEGE-EPS and IFS-
300 EPS members, thus breaking the assumption of independent ensemble members. In order to
301 examine this aspect, Figure 5 presents the distribution of effective ensemble size (i.e., the number
302 of independent ensemble members) after the first and second merging steps. It appears that most

303 ensembles have between 9 and 12 independent members, while full 12-member ensembles after
304 the first and second connections only represent 20% and 10% of the cases respectively. Ensembles
305 with less than 6 independent members are also observed in a few extreme cases.

306 *d. Performance of seamless forecasts*

307 The different seamless designs are evaluated using two criteria : the temporal discontinuities in
308 temperature time series in the vicinity of the junction time, and the probabilistic performances up
309 to 15 days ahead. Ideally, the best seamless design is the one that optimizes the two criteria, in
310 other words, that minimizes temporal discontinuities and maximizes the performance.

311 *(i) Analysis of average discontinuities* The ability of the seamless design strategies to generate
312 smooth transitions between two EPS members is measured with the amplitude of temperature
313 differences between lead times 48h and 49h for the first assignment, and lead times 96h and 97h
314 for the second assignment. The distributions of these one-hour absolute temperature differences
315 are shown in Figure 6 for the three seamless designs. Differences obtained with the IFS-EPS
316 members, which are seamless by construction, are taken as a reference. As expected, the largest
317 discontinuities are obtained with the RN strategy, while the discontinuities obtained with the NN
318 and HU assignments are much smaller and close to those observed in the seamless IFS-EPS
319 forecasts. These results thus indicate that the mergings obtained with the NN and HU methods
320 are able to provide realistic forecasts. Note also that smaller discontinuities are observed for the
321 96h-connection, especially for the RN strategy. This is because this second assignment is done only
322 for the ARPEGE-EPS members, and the IFS-EPS members selected at 48h contribute to reduce
323 the average temperature differences.

324 (ii) *CRPS* The probabilistic performances of the different seamless ensemble forecasts are as-
325 sessed with respect to the IFS-EPS performance. Since the seamless ensembles have only 12
326 members, the IFS-EPS scores are computed using a random subset of 12 members. As expected
327 from Figure 4, the positive impact of seamless forecasts is the largest in the first 24 hours, when
328 forecasts from the AROME-EPS are used, with an average CRPS improvement of 9% over the
329 first forecast day (Figure 7). For longer lead times, the NN ensemble performs significantly worse
330 because of its smaller effective ensemble size. The RN ensemble performs best for forecast days 2
331 to 4 because the number of IFS-EPS members, which perform better than ARPEGE-EPS members
332 (Figure 4), is larger. From forecast days 5 to 15, RN and HU ensembles have similar performances,
333 with CRPS values not statistically different from the IFS-EPS CRPS. Examination of CRPS at
334 individual locations (not shown) indicates that over the first 2 days using AROME-EPS leads to an
335 improvement at 8 sites (over 11). At longer ranges (from days 7 to 15), the seamless HU and RN
336 forecasts both improve on average the IFS-EPS performance at 7 sites.

337 (iii) *AUC and Reliability* The short-range improvement of seamless forecasts is linked to both
338 increased resolution (Figure 8) and reliability (Figure 9). Over the first 48 hours, seamless
339 forecasts improve the reliability of probabilities higher than 0.6 for the event 2-meter temperature
340 above 15°C. Over the whole 15 days forecast range, RN and HU ensembles are very close to the
341 IFS-EPS, while the NN reliability is slightly worse, especially for small probabilities. Similar
342 conclusions are obtained with thresholds 5 and 10°C.

343

344 Since the added value of using the AROME-EPS is mainly restricted to the first 24 hours, an
345 alternative seamless setup would be to perform the first merging at 24h instead of 48h. Figure
346 10 indicates that this earlier merging leads to noticeable improvements of the HU ensemble

347 performance from day 2 to 4.

348

349 Overall, from a NWP point of view the HU seamless design is the only one to fulfill both
350 smoothness and performance evaluation criteria. This seamless forecast is shown to efficiently
351 combine the higher performance of the convective-scale AROME-EPS at short ranges, while
352 maintaining the performance of the IFS-EPS at longer ranges. In the next section, the impact of
353 calibrated and seamless forecasts is examined on the performance of the agronomic simulations.

354 **4. EVA results**

355 In this section we focus on the development of the second generation of the grape vine moth.
356 The treatment date is determined from the cumulative level of the egg stage using the threshold
357 15%, commonly used by experts for the application of biocontrol treatments. Five ensembles of
358 12 EVA simulations are compared, which are driven by 12 random temperature forecasts from the
359 raw and calibrated IFS-EPS, and by the three 12-member seamless ensembles. In order to evaluate
360 these predictions, scores commonly used for NWP EPS verification are applied to the distributions
361 of treatment dates. These scores are presented as a function of the time interval between the
362 observation-based treatment date and the forecast starting date. According to this approach and
363 considering that the maximum lead time of probabilistic forecasts used for this work is 15 days,
364 the results are provided for time intervals between 1 and 15 days. For long time intervals, the
365 treatment threshold may not be reached by some members at the end of the forecast. For these
366 members the model sets the treatment date to "beyond". In order to perform the scores computation,
367 these "beyond" dates are arbitrarily set up to 17. Finally, given the results of the meteorological
368 evaluation, only seamless forecasts with a first merging at 24h are shown in this section.

369 (i) *Bias* The bias of the ensemble mean treatment date is presented in Figure 11. For short-term
370 predictions corresponding to time intervals between 1 and 3 days, the bias is zero whatever the
371 temperature forecast used. For time intervals between 4 and 7 days, the bias slightly increases
372 toward positive values (which means that the treatment date is slightly too late) and there is a small
373 benefit from using calibrated temperature forecasts. This positive bias is in agreement with the
374 negative temperature bias of raw IFS-EPS forecasts seen in Figure 3. The advantage of calibration
375 becomes larger for forecasts between one and two weeks before the treatment date, with bias values
376 that remain close to zero.

377 (ii) *CRPS* The impact of the calibration procedure is also clearly visible on the CRPS (Figure 12),
378 which is significantly improved for week-2 forecasts by almost 40%. This improvement is observed
379 at all sites except Bordeaux. On the other hand, it is more difficult to conclude about the value and
380 ranking of seamless forecasts, compared to calibrated IFS-EPS. Local CRPS at individual locations
381 (not shown) indicate that the seamless HU week-2 forecasts lead to a slight improvement over the
382 calibrated IFS-EPS at 6 sites (over 8). On the other hand, NN and RN showed an improvement
383 over 5 sites.

384 (iii) *Observation coverage* Finally, Figure 13 shows the percentage of observation-based
385 treatment dates that fall in the predicted range. Consistently with previous results, the prediction
386 is almost perfect for short intervals and hardly sensitive to the meteorological forecast used up to
387 a one-week interval. For longer time intervals, the observation coverage significantly decreases
388 to less than 70%, which is lower than the theoretical value of 85% for a reliable 12-member
389 ensemble, while calibrated and seamless forecasts help maintaining a high level of observation
390 coverage, with values around 95% or more up to 11 days. The mean observation coverage over
391 days 7-15 is slightly larger for the HU forecasts, with a value of 94.5% compared to values between

392 91% and 92% for the calibrated, RN and NN forecasts.

393

394 From an agronomic point of view, the impact of weather forecast improvements is rather limited
395 up to one week before the treatment date. This behaviour may be specific to the EVA model. It
396 is possible that the evolution rules based on the accumulation of daily temperatures and threshold
397 effects contribute to reduce the impact of forecast accuracy and uncertainty at short ranges. On
398 the other hand, the advantage of using calibrated forecasts is clearly visible between 7 and 15 days
399 and this may help winegrowers to better anticipate the scheduling of treatments. It is however not
400 possible to properly conclude on the added value of seamless forecasts with the current experimental
401 setup. Additional simulations covering a larger number of locations and years may be required to
402 detect a more robust signal.

403 **5. Discussion and future works**

404 The goal of this paper was to evaluate the potential of statistical post-processing and seamless
405 integration methods to generate skillful 2-meter temperature ensemble forecasts for application
406 to crop protection. Post-processing is based on the well-known EMOS and ECC methods, while
407 new strategies have been proposed to design seamless ensemble forecasts from the concatenation
408 of several ensemble predictions with different spatio-temporal scales. The seamless methods aim
409 at generating ensemble members from high-resolution AROME-EPS forecasts at short ranges,
410 with a smooth transition towards larger scale EPSs at longer ranges. The proposed approach
411 performs minimum-cost associations between members from the different EPSs, according to a
412 given distance and matching rules.

413 The evaluation is first performed from a meteorological point of view. It is shown that a naive
414 random association of EPS members achieves satisfactory probabilistic performances, but with

415 the inconvenience of generating large discontinuities in the vicinity of the matching time. On
416 the other hand, the distance-based associations lead to smaller discontinuities and achieve good
417 performances with the Hungarian matching algorithm.

418 Ensembles of agronomic forecasts are then issued from the calibrated and seamless weather
419 forecasts. Statistical post-processing strongly benefits to the quality of the treatment date prediction,
420 especially between one and two weeks before the treatment date, which indeed allow a better
421 anticipation of the spraying schedule for the winegrowers. However, the impact of seamless
422 forecasts is more difficult to assess, and may require a more thorough evaluation considering
423 additional locations and years. Agronomic models that depend on less-predictable variables and
424 that are more sensitive to short-range weather uncertainty will also be considered in future studies
425 to better highlight the potential of the seamless ensembles.

426 The proposed method is generic and could be applied to other parameters. Since precipitation
427 play a major role in the development of diseases and pests, calibration and seamless combination
428 of precipitation forecasts is a natural step forward. The calibration of precipitation has already
429 been considered in the work of Scheuerer (2014); Taillardat et al. (2019) for instance. The design
430 of seamless predictions may require specific considerations especially regarding the choice of
431 the distance metric to measure the similarity between precipitation forecasts which are subject to
432 spatio-temporal intermittency.

433 Seamless forecasts are important to several applications, including in particular hydrology and
434 energy production, and concern different time scales, e.g. from nowcasting to NWP or from NWP to
435 seasonal forecasts. It would be straightforward to design seamless forecasts over wider time ranges
436 with the integration of more EPSs and matching steps. For instance, additional connections to S2S
437 forecasts (White et al. 2017) could be interesting for a longer term management of agricultural
438 practices.

439 The smoothness criterion is central in the proposed seamless design strategy, although it may
440 not be crucial for this crop protection application that is mainly sensitive to the cumulative effect
441 of some weather variables. However, it is relatively intuitive and of particular interest for other
442 applications, such as the electricity production from renewable energies, which is very sensitive to
443 rapid changes in the meteorological conditions.

444 Possible avenues for improving the seamless integration method include a relaxation of the
445 ‘AROME-EPS only’ assumption over the first 2 days with, for instance, an initialization of the
446 seamless ensemble with a clustering of all available EPS members. It is also possible to develop a
447 backward matching procedure, starting from all IFS-EPS members at long ranges with backward
448 connections to shorter-range EPSs. This approach would preserve the full ensemble variability at
449 long ranges (that may be the most relevant for crop protection). Another elegant way would be
450 to build seamless blended ensemble forecasts from the aggregation of probabilistic forecasts and
451 Schaake shuffle reconstruction, as proposed by Zamo (2016).

452 The extension of the seamless design to a multivariate context is an important future step since
453 most applications require several meteorological input variables. There are at least two ways of
454 doing this, either using the current univariate matching scheme based on the most important weather
455 variable for the application, or using a multivariate metric such as the Mahalanobis distance. This
456 aspect will be the subject of future works.

457 Finally, an important question is whether it is more efficient to calibrate the input temperature
458 forecasts or rather to calibrate directly the end-product. In the present case, the calibration of
459 temperature is the preferred option because it first improves the construction of seamless forecasts
460 by reducing the concatenation gaps. Secondly the calibration of treatment dates is, to the best
461 of our knowledge, an unexplored area that may face several challenges. For instance, there is no
462 guarantee that treatment dates follow a simple parametric distribution and the historical sample of

463 observation-based treatment dates available for calibration is rather small because there are only 3
464 or 4 generations of grape vine moth per year.

465 *Acknowledgments.* This work is a part of a Ph.D. thesis that is supported by the French National
466 Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004,
467 for DigitAg “Digital Agriculture Convergence Lab” (www.hdigitag.fr/en). It is also supported by
468 the CASDAR-funded project METEOPREC “the contribution of accurate forecasting to farmers”
469 (CASDAR are funded by the French Ministry of Agriculture) and the network Modeling and Data
470 Analysis for Agriculture (www.modelia.org). This article has greatly benefited from the support
471 of my colleague Michael Zamo for the EMOS post-processing. I also thank Frank Souverain,
472 Sébastien Prat and Pascal Raynaud for their help regarding the extraction of weather forecasts.

473 *Data availability statement.* The forecast data used for this work are taken from the Météo-France
474 and ECMWF databases. The agronomic EVA model is the property of the French Institute for
475 Vine and Wine.

476 **References**

477 Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction.
478 *Nature*, **525**, 47–55.

479 Bellier, J., G. Bontron, and I. Zin, 2017: Using meteorological analogues for reordering postpro-
480 cessed precipitation ensembles in hydrological forecasting. *Water Resources Research*, **53** (12),
481 10 085–10 107.

482 Bellier, J., I. Zin, and G. Bontron, 2018: Generating coherent ensemble forecasts after hydrological
483 postprocessing: Adaptations of ecc-based methods. *Water Resources Research*, **54** (8), 5741–
484 5762.

- 485 Berndt, D. J., and J. Clifford, 1994: Using dynamic time warping to find patterns in time series.
486 *KDD workshop*, Seattle, WA, Vol. 10, 359–370.
- 487 Bouttier, F., and L. Raynaud, 2018a: Clustering and selection of boundary conditions for limited-
488 area ensemble prediction. *Quarterly Journal of the Royal Meteorological Society*, **144** (717),
489 2381–2391.
- 490 Bouttier, F., and L. Raynaud, 2018b: Clustering and selection of boundary conditions for limited
491 area ensemble prediction. *Quarterly Journal of the Royal Meteorological Society*, **144**, 2381–
492 2391.
- 493 Calanca, P., D. Bolius, A. Weigel, and M. Liniger, 2011: Application of long-range weather
494 forecasts to agricultural decision problems in europe. *The Journal of Agricultural Science*,
495 **149** (1), 15–22.
- 496 Candille, G., C. Côté, P. Houtekamer, and G. Pellerin, 2007: Verification of an ensemble prediction
497 system against observations. *Mon. Wea. Rev.*, **135** (7), 2688–2699.
- 498 Candille, G., and O. Talagrand, 2005: Evaluation of probabilistic prediction systems for a scalar
499 variable. *Quarterly Journal of the Royal Meteorological Society: A J. Atmos. Sci., applied
500 meteorology and physical oceanography*, **131** (609), 2131–2150.
- 501 Chavent, F., 1983: Etude de la dynamique des populations d’Eudémis de la vigne (Lobesia botrana
502 Den. et Schiff.). Adaptation d’un modèle d’évolution aux conditions climatiques de Provence,
503 mémoire de fin d’études, In French.
- 504 Cheung, J., A. Hally, J. Heijstek, A. Marsman, and J.-L. Brenguier, 2015: Recommendations on
505 trajectory selection in flight planning based on weather uncertainty. *Proc. 5th SESAR Innovation
506 Days (SID2015)*, Bologna, Italy, 1–8.

- 507 Christ, E. H., P. J. Webster, G. D. Collins, V. E. Toma, and S. A. Byrd, 2015: Using precipitation
508 forecasts to irrigate cotton. *The Journal of Cotton Science*, 351–358.
- 509 Clark, M., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby, 2004: The schaake shuffle:
510 A method for reconstructing space–time variability in forecasted precipitation and temperature
511 fields. *J. Hydrometeor.*, **5** (1), 243–262.
- 512 Cooke, B. M., D. G. Jones, B. Kaye, and Coauthors, 2006: *The epidemiology of plant diseases*,
513 Vol. 2. Springer.
- 514 Descamps, L., C. Labadie, A. Joly, E. Bazile, P. Arbogast, and P. Cébron, 2015: Pearp, the météo-
515 france short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological*
516 *Society*, **141** (690), 1671–1685.
- 517 Fundel, V. J., N. Fleischhut, S. M. Herzog, M. Göber, and R. Hagedorn, 2019: Promoting the use of
518 probabilistic weather forecasts through a dialogue between scientists, developers and end-users.
519 *Quarterly Journal of the Royal Meteorological Society*, **145**, 210–231.
- 520 Ghaffary, M. T., 2011: *Efficacy and mapping of resistance to Mycosphaerella graminicola in*
521 *wheat*.
- 522 Giorgino, T., and Coauthors, 2009: Computing and visualizing dynamic time warping alignments
523 in R: the DTW package. *Journal of statistical Software*, **31** (7), 1–24.
- 524 Gneiting, T., 2014: Calibration of medium-range weather forecasts. *ECMWF Technical Memorandum*,
525 **719**.
- 526 Gneiting, T., and A. E. Raftery, 2005: Weather forecasting with ensemble methods. *Science*,
527 **310** (5746), 248–249.

- 528 Gneiting, T., A. E. Raftery, A. H. Westveld III, and T. Goldman, 2005: Calibrated probabilistic
529 forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea.*
530 *Rev.*, **133** (5), 1098–1118.
- 531 Hagelin, S., J. Son, R. Swinbank, A. McCabe, N. Roberts, and W. Tennant, 2017: The Met Office
532 convective-scale ensemble MOGREPS-UK. *Quart. J. Roy. Meteor. Soc.*, **143**, 2846–2861.
- 533 Hanley, J. A., and B. J. McNeil, 1982: The meaning and use of the area under a receiver operating
534 characteristic (roc) curve. *Radiology*, **143** (1), 29–36.
- 535 Hemri, S., M. Scheuerer, F. Pappenberger, K. Bogner, and T. Haiden, 2014: Trends in the predictive
536 performance of raw ensemble weather forecasts. *Geophysical Research Letters*, **41** (24), 9197–
537 9205.
- 538 Kollers, S., and Coauthors, 2013: Genetic architecture of resistance to septoria tritici blotch
539 (*mycosphaerella graminicola*) in european winter wheat. *Molecular breeding*, **32** (2), 411–423.
- 540 Kuhn, H. W., 1955: The hungarian method for the assignment problem. *Naval research logistics*
541 *quarterly*, **2** (1-2), 83–97.
- 542 Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics*,
543 **227** (7), 3515–3539.
- 544 Miedaner, T., Y. Zhao, M. Gowda, C. F. H. Longin, V. Korzun, E. Ebmeyer, E. Kazman, and J. C.
545 Reif, 2013: Genetic architecture of resistance to septoria tritici blotch in european wheat. *BMC*
546 *genomics*, **14** (1), 858.
- 547 Molteni, F., R. Buizza, T. N. Palmer, and T. Petroligias, 1996: The ecmwf ensemble prediction
548 system: Methodology and validation. *Quarterly journal of the royal meteorological society*,
549 **122** (529), 73–119.

- 550 Moyer, M. M., D. M. Gadoury, W. F. Wilcox, and R. C. Seem, 2016: Weather during critical
551 epidemiological periods and subsequent severity of powdery mildew on grape berries. *Plant*
552 *disease*, **100** (1), 116–124.
- 553 Olatinwo, R., T. Prabha, J. Paz, D. Riley, and G. Hoogenboom, 2011: The Weather Research and
554 Forecasting (WRF) model: application in prediction of TSWV-vectors populations. *Journal of*
555 *Applied Entomology*, **135** (1-2), 81–90.
- 556 Olatinwo, R. O., T. V. Prabha, J. O. Paz, and G. Hoogenboom, 2012: Predicting favorable conditions
557 for early leaf spot of peanut using output from the Weather Research and Forecasting (WRF)
558 model. *International journal of biometeorology*, **56** (2), 259–268.
- 559 Palmer, T., 2019: The ecmwf ensemble prediction system: Looking back (more than) 25 years and
560 projecting forward 25 years. *Quarterly Journal of the Royal Meteorological Society*, **145**, 12–24.
- 561 Pappenberger, F., F. Wetterhall, E. Dutra, F. Di Giuseppe, K. Bogner, L. Alfieri, and H. L. Cloke,
562 2013: Seamless forecasting of extreme events on a global scale. *Climate and Land Surface*
563 *Changes in Hydrology*, edited by: Boegh, E., Blyth, E., Hannah, DM, Hisdal, H., Kunstmann,
564 H., Su, B., and Yilmaz, KK, IAHS Publication, Gothenburg, Sweden, 3–10.
- 565 Pertot, I., and Coauthors, 2017: A critical review of plant protection tools for reducing pesticide use
566 on grapevine and new perspectives for the implementation of ipm in viticulture. *Crop Protection*,
567 **97**, 70–84.
- 568 Pinson, P., H. A. Nielsen, H. Madsen, and G. Kariniotakis, 2009: Skill forecasting from ensemble
569 predictions of wind power. *Applied Energy*, **86** (7-8), 1326–1334.
- 570 Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using bayesian model
571 averaging to calibrate forecast ensembles. *Mon. Wea. Rev.*, **133** (5), 1155–1174.

- 572 Raynaud, L., and F. Bouttier, 2016: Comparison of initial perturbation methods for ensemble
573 prediction at convective scale. *Quarterly Journal of the Royal Meteorological Society*, **142 (695)**,
574 854–866.
- 575 Sakoe, H., S. Chiba, A. Waibel, and K. Lee, 1990: Dynamic programming algorithm optimization
576 for spoken word recognition. *Readings in speech recognition*, 159–165.
- 577 Schefzik, R., 2011: Ensemble copula coupling. *Master's thesis, Faculty of Mathematics and*
578 *Informatics, University of Heidelberg, Germany.*
- 579 Schefzik, R., T. L. Thorarinsdottir, T. Gneiting, and Coauthors, 2013: Uncertainty quantification
580 in complex simulation models using ensemble copula coupling. *Statistical Science*, **28 (4)**,
581 616–640.
- 582 Scheuerer, M., 2014: Probabilistic quantitative precipitation forecasting using ensemble model
583 output statistics. *Quarterly Journal of the Royal Meteorological Society*, **140 (680)**, 1086–1096.
- 584 Shipp, J. L., and N. D. Clarke, 1999: Decision tools for integrated pest management. *Integrated*
585 *Pest and Disease Management in Greenhouse Crops*, Springer, 168–182.
- 586 Suffert, F., I. Sache, and C. Lannou, 2011: Early stages of septoria tritici blotch epidemics of
587 winter wheat: build-up, overseasoning, and release of primary inoculum. *Plant Pathology*,
588 **60 (2)**, 166–177.
- 589 Taillardat, M., A.-L. Fougères, P. Naveau, and O. Mestre, 2019: Forest-based and semiparametric
590 methods for the postprocessing of rainfall ensemble forecasting. *Wea. Forecasting*, **34**, 617–634.
- 591 Thiéry, D., P. Louâpre, L. Muneret, A. Rusch, G. Sentenac, F. Vogelweith, C. Iltis, and J. Moreau,
592 2018: Biological protection against grape berry moths. a review. *Agronomy for sustainable*
593 *development*, **38 (2)**, 15.

- 594 Thiéry, D., and Coauthors, 2013: Histoire de l'installation de quelques ravageurs, In French.
- 595 Wetterhall, F., and F. Di Giuseppe, 2018: The benefit of seamless forecasts for hydrological
596 predictions over europe. *Hydrology & Earth System Sciences*, **22**, 3409–3420.
- 597 White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (s2s) predic-
598 tions. *Meteorological Applications*, **24** (3), 315–325, doi:10.1002/met.1654.
- 599 Wilcoxon, F., S. Katti, and R. A. Wilcox, 1970: Critical values and probability levels for the
600 wilcoxon rank sum test and the wilcoxon signed rank test. *Selected tables in mathematical*
601 *statistics*, **1**, 171–259.
- 602 Worsnop, R. P., M. Scheuerer, and T. M. Hamill, 2019: Extended-range probabilistic fire-weather
603 forecasting based on ensemble model output statistics and ensemble copula coupling. *Mon. Wea.*
604 *Rev.*, **148**, 499–521.
- 605 Zamo, M., 2016: Statistical post-processing of deterministic and ensemble wind speed forecasts
606 on a grid. Ph.D. thesis, University of Paris-Saclay.

607 **LIST OF TABLES**

608 **Table 1.** Characteristics of the three EPSs used (corresponding to their operational con-
609 figurations over the period 1 June 2018 - 31 May 2019). 31

610 **Table 2.** Percentage of selected ARPEGE-EPS and IFS-EPS members with the different
611 assignment strategies for the first matching step, over different time periods.
612 Theoretical values for the random selection are 40.7% for ARPEGE-EPS and
613 59.3% for IFS-EPS, according to their respective ensemble size. 32

Model Characteristics	AROME-EPS	ARPEGE-EPS	IFS-EPS
Size	12 perturbed members	34 perturbed members + 1 control forecast	50 perturbed members + 1 control forecast
Horizontal resolution	2.5km	stretched grid, 10km over France	18km
Vertical resolution (levels)	90	90	91
Initial time of forecast	21 UTC	18 UTC	00 UTC
Lead times	51h	102h	360h
Output frequency	1h	3h	3h up to 144h, 6h beyond
Dynamical core	non-hydrostatic	hydrostatic	hydrostatic
Archive grid resolution (regular lat/lon)	0.025°×0.025°	0.1°×0.1°	0.5°×0.5°

614 TABLE 1. Characteristics of the three EPSs used (corresponding to their operational configurations over the
615 period 1 June 2018 - 31 May 2019).

Time Period	ARPEGE-EPS			IFS-EPS		
	RN	NN	HU	RN	NN	HU
2018-06-01-2018-08-31	40.7 %	45.65 %	45.41 %	59.3 %	54.35 %	54.59 %
2018-09-01-2018-11-30	40.7 %	49.60 %	49.10 %	59.3 %	50.40 %	50.90 %
2018-12-01-2019-02-28	40.7 %	53.00 %	52.50 %	59.3 %	47.00 %	47.50 %
2019-03-01-2019-05-31	40.7 %	46.27 %	45.87 %	59.3 %	53.73 %	54.13 %

616 TABLE 2. Percentage of selected ARPEGE-EPS and IFS-EPS members with the different assignment strategies
617 for the first matching step, over different time periods. Theoretical values for the random selection are 40.7% for
618 ARPEGE-EPS and 59.3% for IFS-EPS, according to their respective ensemble size.

619 **LIST OF FIGURES**

620 **Fig. 1.** Schematic representation of the analysis and prediction steps. During the analysis step the
621 agronomic model is run with observation data from an initialization date to the forecast
622 starting date. The model is then run using weather forecasts from the forecast starting date
623 to the maximum lead time. 35

624 **Fig. 2.** Locations used for the weather and agronomic simulations. The red points correspond to
625 vineyards and the green points correspond to wheat fields. The latitude-longitude coordinates
626 of each point are also provided. 36

627 **Fig. 3.** Bias (left) and RMSE (right) of ensemble mean forecast for 2-meter temperature, as a function
628 of forecast lead time. Scores are computed for the 11 sites over the period 01/06/2018-
629 31/05/2019 with raw (solid line) and calibrated (dashed line) ensembles. Top: AROME-
630 EPS, middle: ARPEGE-EPS and bottom: IFS-EPS. Top + (resp. -) signs indicate that the
631 performance of the calibrated ensemble is statistically better (resp. worse) than the raw
632 ensemble performance, according to the Wilcoxon test. 37

633 **Fig. 4.** CRPS of 2-meter temperature forecasts as a function of forecast lead time. Scores are
634 computed for the 11 sites over the period 01/06/2018-31/05/2019 for the three calibrated
635 EPS. Top + (resp. -) signs indicate that the performance of the AROME-EPS and the
636 ARPEGE-EPS is statistically better (resp. worse) than the IFS-EPS performance, according
637 to the Wilcoxon test. 38

638 **Fig. 5.** Frequency histograms of the number of independent members in the NN ensembles, com-
639 puted for the 11 sites over the period 01/06/2018-31/05/2019, (a) after the 48h-merging step
640 and (b) after the 96h-merging step. 39

641 **Fig. 6.** Boxplots of absolute temperature forecast differences, averaged over ensemble members,
642 between lead times 48 and 49h (in black) and lead times 96 and 97h (in gray), computed for
643 the 11 sites over the period 01/06/2018-31/05/2019. These forecast differences are presented,
644 from left to right, for IFS-EPS members and for the seamless RN, NN and HU ensemble
645 members. 40

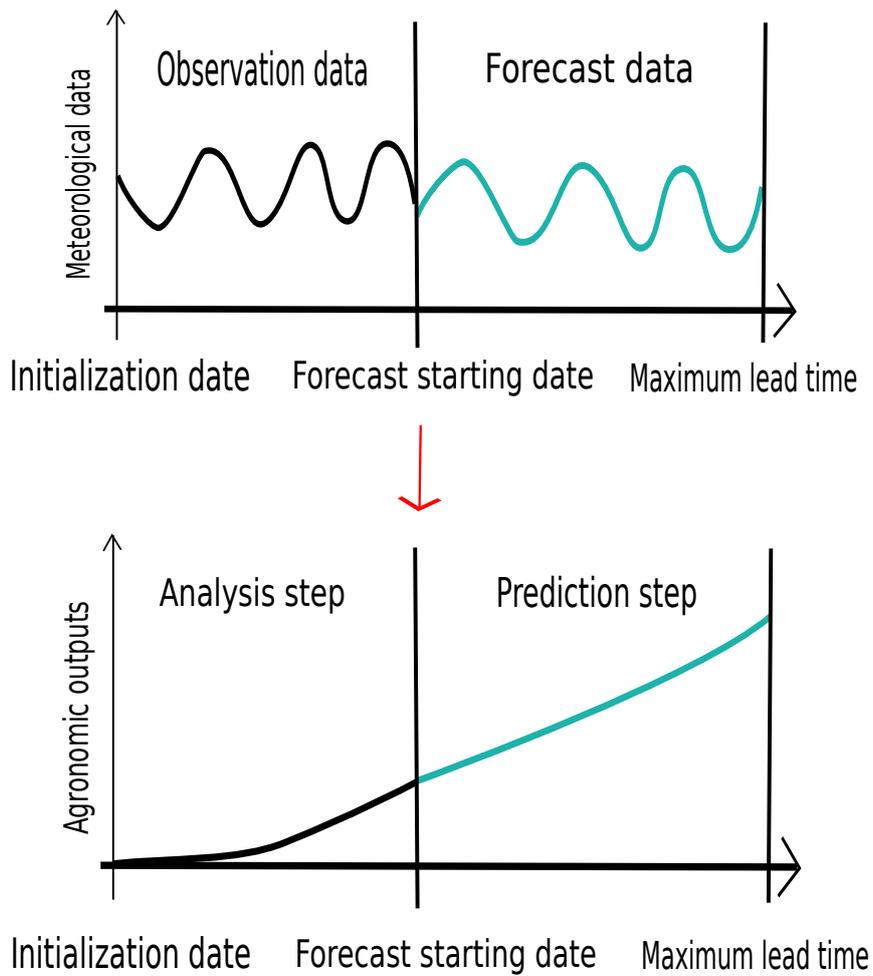
646 **Fig. 7.** CRPS of 2-meter temperature forecasts computed for the 11 sites over the period 01/06/2018-
647 31/05/2019, as a function of forecast day. Results for the random 12-member IFS-EPS and
648 the three seamless ensembles are overlaid. Top + (resp. -) signs indicate that the performance
649 of a seamless ensemble is statistically better (resp. worse) than the IFS-EPS performance,
650 according to the Wilcoxon test. 41

651 **Fig. 8.** AUC of 2-meter temperature forecasts above 15°C, computed for the 11 sites over the period
652 01/06/2018-31/05/2019, as a function of forecast day. Results for the random 12-member
653 IFS-EPS and the three seamless ensembles are overlaid. Top + (resp. -) signs indicate that
654 the performance of a seamless ensemble is statistically better (resp. worse) than the IFS-EPS
655 performance, according to the Wilcoxon test. 42

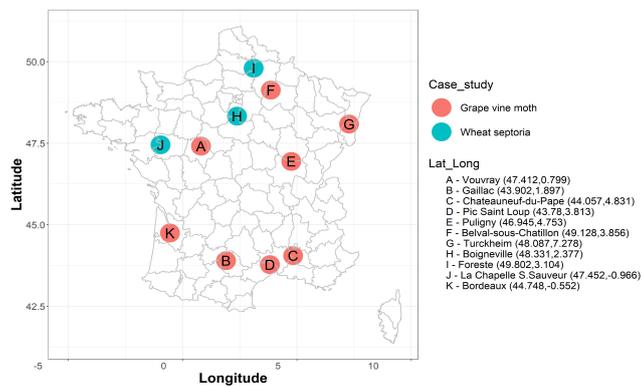
656 **Fig. 9.** Reliability diagram for 2-meter temperature forecasts above 15°C, computed for the 11 sites
657 over the period 01/06/2018-31/05/2019. (a) lead times 0-48h, (b) lead times 0-360h. Results
658 for the random 12-member IFS-EPS and the three seamless ensembles are overlaid. 43

659 **Fig. 10.** CRPS of 2-meter temperature forecasts computed for the 11 sites over the period 01/06/2018-
660 31/05/2019, as a function of forecast day. Results are given for seamless HU ensembles with
661 a first merging at 24h (dashed-dotted line) and 48h (solid line). Top + (resp. -) signs

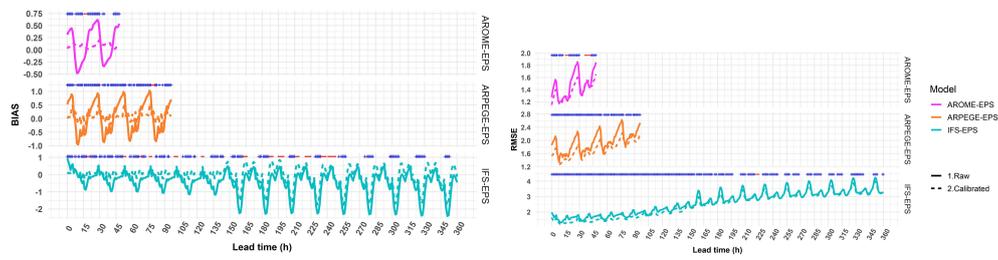
662	indicate that the performance of the 24h-merging is statistically better (resp. worse) than the	
663	performance of the 48h-merging, according to the Wilcoxon test.	44
664	Fig. 11. Bias of the ensemble mean treatment date forecast, as a function of the time interval (in	
665	day) between the observation-based date and the forecast starting date. Scores are averaged	
666	over the 8 vineyard sites and the period of interest 01/04-19/09 for the years 2018 and	
667	2019. Results from using raw IFS-EPS, calibrated IFS-EPS and the three seamless ensemble	
668	forecasts are overlaid.	45
669	Fig. 12. CRPS for the forecast treatment date, as a function of the time interval (in day) between the	
670	observation-based date and the forecast starting date. Scores are averaged over the 8 vineyard	
671	sites and the period of interest 01/04-19/09 for the years 2018 and 2019. Results from using	
672	raw IFS-EPS, calibrated IFS-EPS and the three seamless forecasts are overlaid.	46
673	Fig. 13. Observation coverage, as a function of the time interval (in day) between the observation-	
674	based treatment date and the forecast starting date. The percentage values of observation	
675	coverage are averaged over the 8 vineyard sites and the period of interest 01/04-19/09 for the	
676	years 2018 and 2019. Results from using raw IFS-EPS, calibrated IFS-EPS and the three	
677	seamless forecasts are overlaid.	47



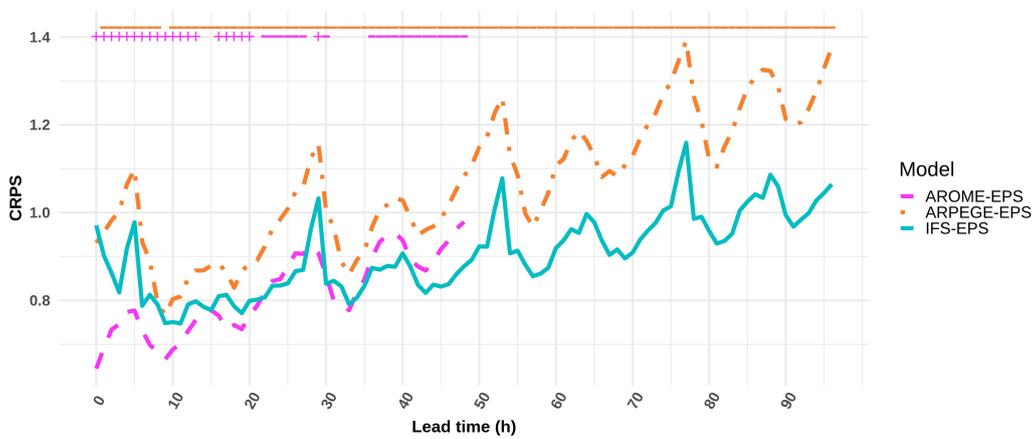
678 FIG. 1. Schematic representation of the analysis and prediction steps. During the analysis step the agronomic
 679 model is run with observation data from an initialization date to the forecast starting date. The model is then run
 680 using weather forecasts from the forecast starting date to the maximum lead time.



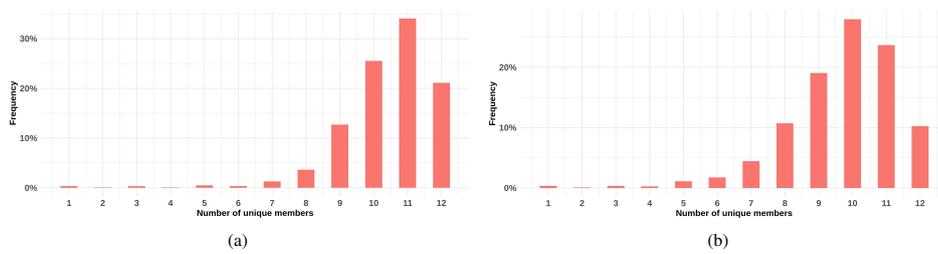
681 FIG. 2. Locations used for the weather and agronomic simulations. The red points correspond to vineyards and
 682 the green points correspond to wheat fields. The latitude-longitude coordinates of each point are also provided.



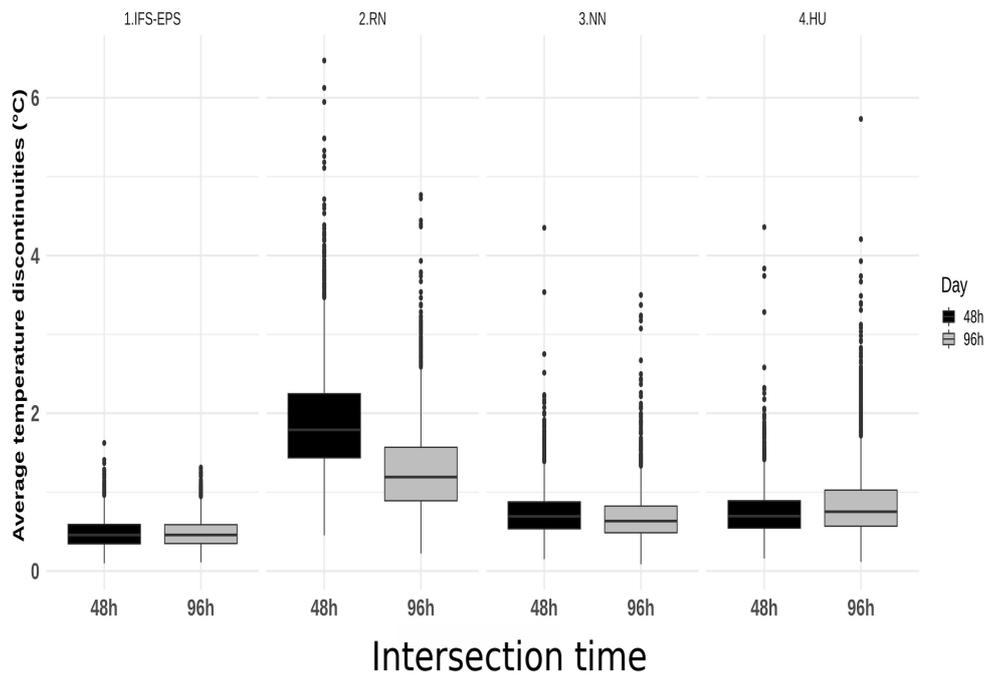
683 FIG. 3. Bias (left) and RMSE (right) of ensemble mean forecast for 2-meter temperature, as a function of
 684 forecast lead time. Scores are computed for the 11 sites over the period 01/06/2018-31/05/2019 with raw (solid
 685 line) and calibrated (dashed line) ensembles. Top: AROME-EPS, middle: ARPEGE-EPS and bottom: IFS-EPS.
 686 Top + (resp. -) signs indicate that the performance of the calibrated ensemble is statistically better (resp. worse)
 687 than the raw ensemble performance, according to the Wilcoxon test.



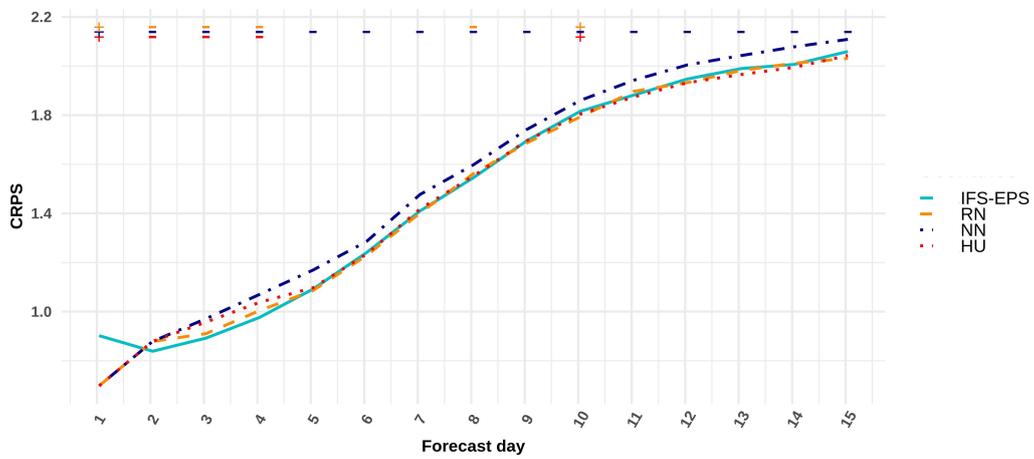
688 FIG. 4. CRPS of 2-meter temperature forecasts as a function of forecast lead time. Scores are computed for
 689 the 11 sites over the period 01/06/2018-31/05/2019 for the three calibrated EPS. Top + (resp. -) signs indicate
 690 that the performance of the AROME-EPS and the ARPEGE-EPS is statistically better (resp. worse) than the
 691 IFS-EPS performance, according to the Wilcoxon test.



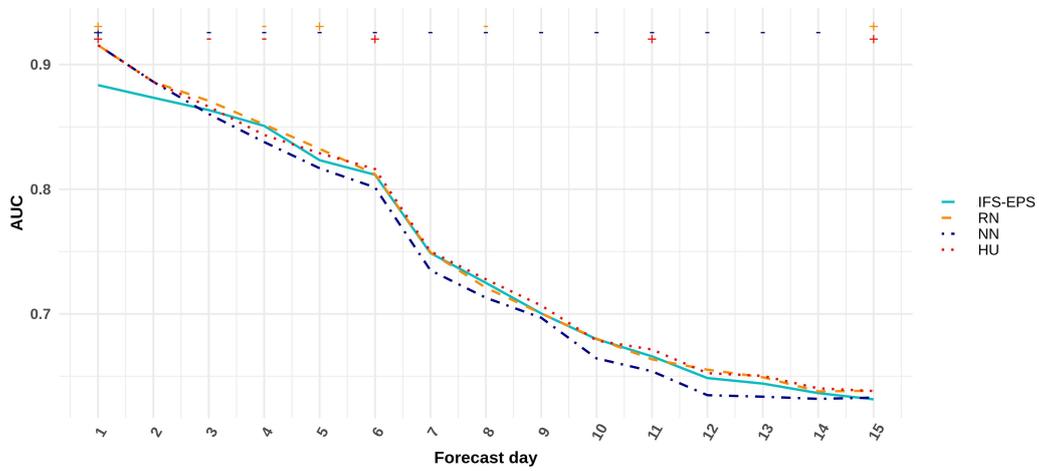
692 FIG. 5. Frequency histograms of the number of independent members in the NN ensembles, computed for the
 693 11 sites over the period 01/06/2018-31/05/2019, (a) after the 48h-merging step and (b) after the 96h-merging
 694 step.



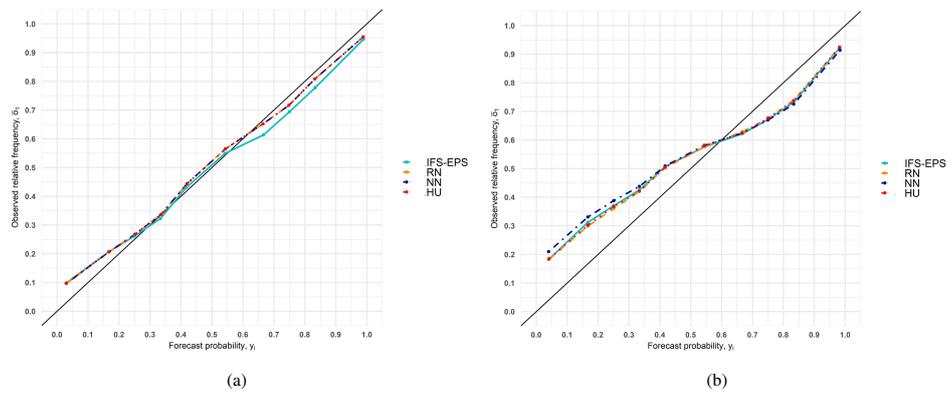
695 FIG. 6. Boxplots of absolute temperature forecast differences, averaged over ensemble members, between lead
 696 times 48 and 49h (in black) and lead times 96 and 97h (in gray), computed for the 11 sites over the period
 697 01/06/2018-31/05/2019. These forecast differences are presented, from left to right, for IFS-EPS members and
 698 for the seamless RN, NN and HU ensemble members.



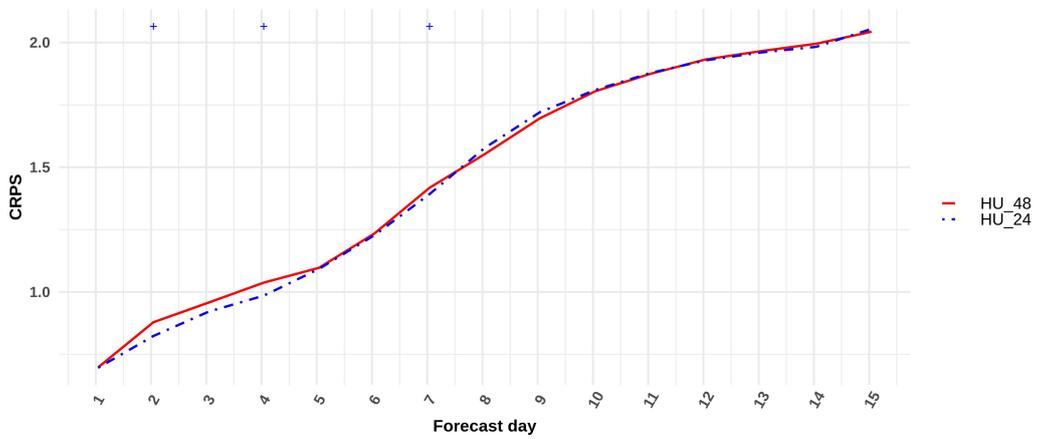
699 FIG. 7. CRPS of 2-meter temperature forecasts computed for the 11 sites over the period 01/06/2018-
700 31/05/2019, as a function of forecast day. Results for the random 12-member IFS-EPS and the three seamless
701 ensembles are overlaid. Top + (resp. -) signs indicate that the performance of a seamless ensemble is statistically
702 better (resp. worse) than the IFS-EPS performance, according to the Wilcoxon test.



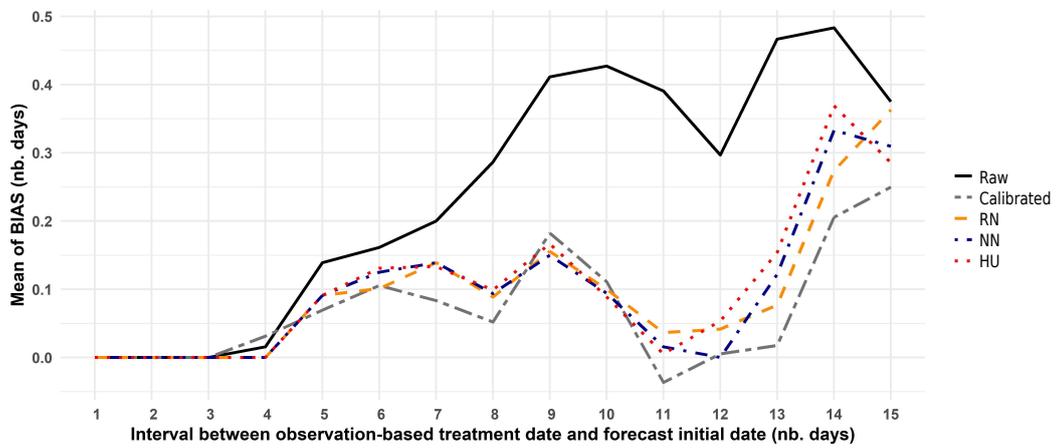
703 FIG. 8. AUC of 2-meter temperature forecasts above 15°C, computed for the 11 sites over the period 01/06/2018-
 704 31/05/2019, as a function of forecast day. Results for the random 12-member IFS-EPS and the three seamless
 705 ensembles are overlaid. Top + (resp. -) signs indicate that the performance of a seamless ensemble is statistically
 706 better (resp. worse) than the IFS-EPS performance, according to the Wilcoxon test.



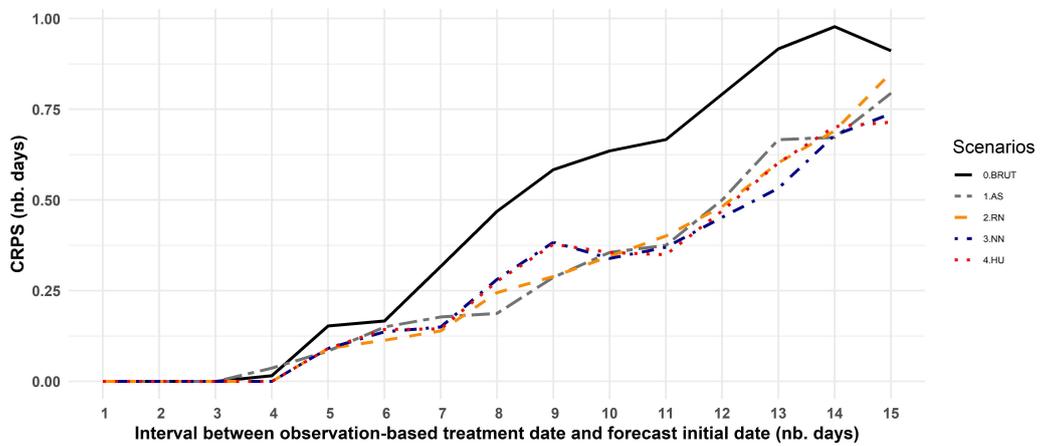
707 FIG. 9. Reliability diagram for 2-meter temperature forecasts above 15°C, computed for the 11 sites over the
 708 period 01/06/2018-31/05/2019. (a) lead times 0-48h, (b) lead times 0-360h. Results for the random 12-member
 709 IFS-EPS and the three seamless ensembles are overlaid.



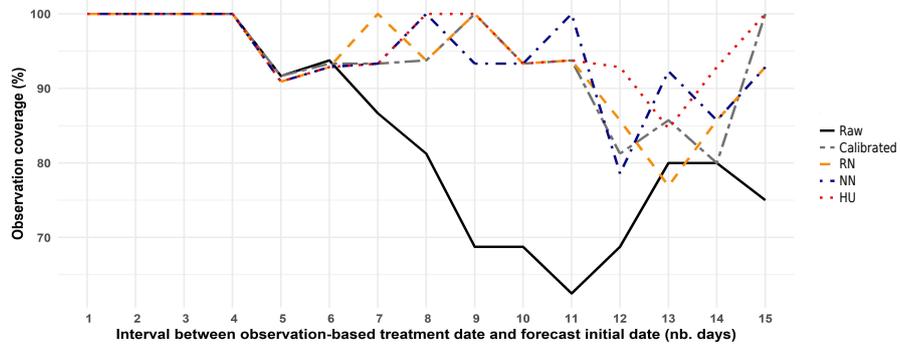
710 FIG. 10. CRPS of 2-meter temperature forecasts computed for the 11 sites over the period 01/06/2018-
 711 31/05/2019, as a function of forecast day. Results are given for seamless HU ensembles with a first merging
 712 at 24h (dashed-dotted line) and 48h (solid line). Top + (resp. -) signs indicate that the performance of the
 713 24h-merging is statistically better (resp. worse) than the performance of the 48h-merging, according to the
 714 Wilcoxon test.



715 FIG. 11. Bias of the ensemble mean treatment date forecast, as a function of the time interval (in day) between
 716 the observation-based date and the forecast starting date. Scores are averaged over the 8 vineyard sites and the
 717 period of interest 01/04-19/09 for the years 2018 and 2019. Results from using raw IFS-EPS, calibrated IFS-EPS
 718 and the three seamless ensemble forecasts are overlaid.



719 FIG. 12. CRPS for the forecast treatment date, as a function of the time interval (in day) between the
 720 observation-based date and the forecast starting date. Scores are averaged over the 8 vineyard sites and the period
 721 of interest 01/04-19/09 for the years 2018 and 2019. Results from using raw IFS-EPS, calibrated IFS-EPS and
 722 the three seamless forecasts are overlaid.



723 FIG. 13. Observation coverage, as a function of the time interval (in day) between the observation-based
 724 treatment date and the forecast starting date. The percentage values of observation coverage are averaged over
 725 the 8 vineyard sites and the period of interest 01/04-19/09 for the years 2018 and 2019. Results from using raw
 726 IFS-EPS, calibrated IFS-EPS and the three seamless forecasts are overlaid.

Travaux complémentaires

E.1 Analyse de sensibilité du modèle SeptoLIS aux variables d'entrée agronomiques et météorologiques

Le projet METEOPREC a pour objectif une amélioration de l'utilisation des données météorologiques pour obtenir des outils d'aide à la décision plus performants dans le domaine agricole. La sensibilité des modèles agronomiques tels que SeptoLIS à cette donnée climatique ou prévisionnelle est encore méconnue. Une amélioration de tels outils d'aide à la décision repose également sur la compréhension de cette sensibilité. Cette étude propose un plan d'analyse de sensibilité applicable aux modèles agronomiques pour quantifier la sensibilité du modèle SeptoLIS à la météorologie dite de prévision. Ce plan d'analyse se déroule en plusieurs étapes : un criblage via une analyse de Morris pour sélectionner les paramètres propres au modèle les plus influents. Un Map Labelling pour vérifier l'importance de l'influence de la prévision météorologique sur le modèle. Une analyse par la méthode des filtres pour quantifier l'impact d'une application des données de prévision de manière générale et des analyses de sensibilité situationnelles pour quantifier des impacts de manière plus localisée. Les résultats montrent que la donnée météorologique est importante pour le modèle SeptoLIS. Une transformation de données climatologiques en données de prévision l'est également mais de manière situationnelle. Le plan d'analyse de sensibilité a été testé et pourra être appliqué à d'autres modèles agronomiques mais également pour étudier d'autres variables potentiellement influentes comme la précision des capteurs.

Relative impact of uncertainties in weather forecast and model parameters for decision in crop protection

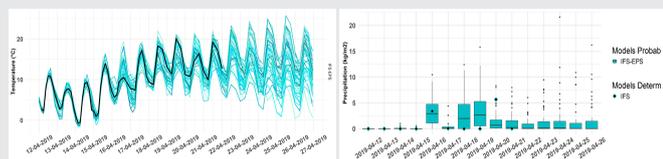
Ivana ALEKSOVSKA^{1,3,6,7}, Roland ALLART⁵, Laure RAYNAUD¹, Robert FAIVRE², François BRUN³, Olivier DEUDON⁵, Marc RAYNAL⁴, Sébastien ROUX², Samuel BUIS²

Introduction

- **Crop diseases** are highly dependent on weather conditions.
- Different **Decision Support Tools (DST)** take them into account to maintain good pesticide consumption practice.
- Weather forecasts remain uncertain which can impact the reliability of estimated **treatment date**.
- **Probabilistic forecasting** that cover different time and length scales to estimate the uncertainty. Their advantages and an assessment of the risks associated with this decision for crop protection has been demonstrated through DST with applications on vines and wheat [1].
- The purpose of this study is to present the **sensitivity analysis** results that quantify the relative impact of the two main sources of uncertainties: agronomic model parameters and weather forecasts.

Ensemble weather forecasting

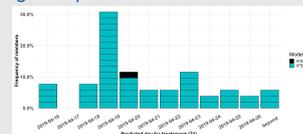
- Weather forecasts information is affected by uncertainty.
- **Ensemble forecast** methods to estimate forecast uncertainty [2].
- Ensemble Prediction Systems (EPS) : **set of perturbed forecasts to sample the probability distribution**.
- Available EPS models for France: AROME-EPS, ARPEGE-EPS (Météo-France), **IFS-EPS** (ECMWF-European Centre for Medium-Range Weather Forecasts).
- Focus on IFS-EPS (Zone: Globe, m=51 members, 18km horizontal resolution, Lead-time Day+15).



Initial time of the forecast 12/04/2019 at Boigneville (France). IFS deterministic (black) and IFS-EPS ensemble forecasts (color) of temperature (left) and precipitation (right).

Septo-LIS[®] for wheat septoria disease

- Model that predicts : the growth stage-appearance of leaves and the evolution of disease [3].
- Input variables : variety of wheat and seeding date, functional inputs (**precipitation and temperature**) and 48 model parameters.
- Optimal treatment rule: leaf contamination threshold (3 last leaves) → Date of treatment (T1).
- Ensemble forecasting + Septo-LIS[®] → **distribution of dates T1**.

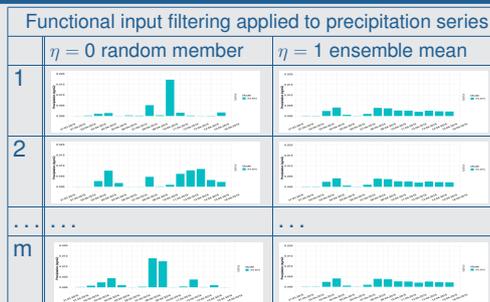


Distribution of dates of treatment T1. Initial time of the forecast 12/04/2019 at Boigneville.

- Sensibility analysis (Morris) → the 5 most influencing model parameters: HauteurMaxFeuille3 (maximum height leaf 3), HauteurMaxFeuille2 (maximum height leaf 2), LorinTmoyCoef2 (model temperature coefficient), SeuilTLorin (model temperature threshold) and CoefFracLatence1 (latency infection - necrosis).

A filter-based approach for sensitivity analysis with weather forecast series as functional inputs

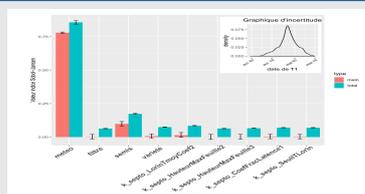
- **Filter method** [4] : analyse and evaluate the sensitivity of the model to the transformation of dynamic input.
- Steps :
 - 1) Transform the meteorological data variables in order to modify one of its element ;
 - 2) Quantify the effect of applying or not this transformation to the model on its output by performing a global sensitivity analysis including a boolean factor that activates or not the transformation.
- More details → presentation : "New indices for Multivariate Sensitivity Analysis using fuzzy clustering", Sebastien Roux, Matieyendou Lamboni and Samuel Buis (SAMO 2019).



Filter scenarios for Auch initial time of the forecast 01/04/2018

Map labeling + filter method in Septo-LIS[®]

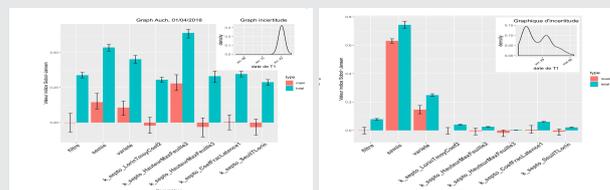
- 30 situations : 15 places, 2 years ;
- Experimental design : Latin Hypercube Sampling (LHS), sample size $N = 15000$;
- Method to compute first order and total variance-based sensitivity indices : **Sobol Jansen**.



Results of the sensitivity analysis applying the filter method to the precipitation and temperature.

The applied transformation have a very low impact compared to the meteo in average on all situations.

Location-time-based analyses



Significant filter effect in Auch initial time of the forecast 01/04/2018 (left) and negligible filter effect in Beauvais initial time of the forecast 15/04/2019 (right).

Weather forecast uncertainties are important in some specific cases depending on different places and various situations.

References

1. Aleksovsk, I., Brun, F., Raynaud, L., Faivre, R., Raynal, M. and Deudon, O., Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures, Végéphyt - 12e conférence internationale sur les maladies des plantes - Tours (France) - 10, 11 et 12 décembre 2018.
2. Bouittier, F. and Buizza, R., 2018. La prévision d'ensemble. La Météorologie(100), pp.29-36.
3. Gouache, D., Bensadoun, A., Brun, F., Pagé, C., Makowski, D. and Wallach, D., 2013. Modelling climate change impact on Septoria tritici blotch (STB) in France: accounting for climate model and disease model uncertainty. Agricultural and forest meteorology 170, pp.242-252.
4. Roux S., Loisel P., Buis S., A filter-based approach for global sensitivity analysis of models with functional inputs, Reliability Engineering & System Safety Elsevier , 187, PP:119-128.

Conclusions

- Ensemble forecast in Septo-LIS[®] leads to distribution of dates of treatment.
- Filter-based approach is used to quantify the relative impact of weather forecast uncertainty.
- Only in particular situation the uncertainty of the forecasts is more important than the uncertainty of the model's parameters or other inputs.

Contact: ivana.aleksovsk@meteo.fr

E.2 Prise en compte des incertitudes des prévisions météorologiques à court terme pour les outils d'aide à la décision pour la vigne et le maïs

Les prévisions d'ensemble météorologiques sont utilisées dans cette étude pour prendre en compte leurs incertitudes dans les bilans hydriques de vigne et maïs. Quatre sites au sud de la France sont étudiés pour l'année 2018 et 2019. Deux bilans hydriques sont utilisés pour les deux cultures, en prenant compte des facteurs environnementaux. Comme prévu, l'incertitude s'accroît quand on s'éloigne du premier jour de prévision. Pourtant, cet incertitude varie beaucoup entre les sites et les dates qui permet ainsi de prendre en compte des conditions météorologiques spécifiques. Localisation, profondeur du sol et enherbement apportent de grands impacts aux bilans hydriques et leurs incertitudes. Afin d'améliorer la prévision, une méthode de débiaisage basée sur une simple régression linéaire est directement appliquée sur les sorties de bilan hydrique. Parmi les quatres sites étudiés, seulement une localisation montre une réduction significative de biais après le débiaisage. Les conséquences de l'incertitude sur les dates d'irrigation simulées par le modèle pour le maïs sont également étudiées.

Bibliographie

- Aldrich, John *et al.* (1997). « R.A Fisher and the making of maximum likelihood 1912-1922 ». In : *Statistical science* 12.3, p. 162-176 (cf. p. 30).
- Aleksovska, I., F. Brun, L. Raynaud, R. Faivre, M. Raynal et O. Deudon (2018). « Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures ». In : *Végéphyt - 12e conférence internationale sur les maladies des plantes - Tours (France) - 10, 11 et 12 décembre 2018* (cf. p. 60, 63).
- Baudin, Paul (2015). « Prévision séquentielle par agrégation d'ensemble : application à des prévisions météorologiques assorties d'incertitudes ». Thèse de doct. Paris Saclay (cf. p. 107).
- Bauer, Peter, Alan Thorpe et Gilbert Brunet (2015). « The quiet revolution of numerical weather prediction ». In : *Nature* 525.7567, p. 47 (cf. p. 8).
- Bellier, Joseph, Guillaume Bontron et Isabella Zin (2017). « Using meteorological analogues for reordering postprocessed precipitation ensembles in hydrological forecasting ». In : *Water Resources Research* 53.12, p. 10085-10107 (cf. p. 2).
- Bellier, Joseph, Isabella Zin et Guillaume Bontron (2018). « Generating Coherent Ensemble Forecasts After Hydrological Postprocessing : Adaptations of ECC-Based Methods ». In : *Water Resources Research* 54.8, p. 5741-5762 (cf. p. 2).
- Berndt, Donald J et James Clifford (1994). « Using dynamic time warping to find patterns in time series. » In : *KDD workshop*. T. 10. 16. Seattle, WA, p. 359-370 (cf. p. 36).
- Bouttier, F. et R. Buizza (2018). « La prévision d'ensemble ». In : *La Météorologie* 100, pp. 29-36 (cf. p. 8).
- Bouttier, François et Laure Raynaud (2018). « Clustering and selection of boundary conditions for limited area ensemble prediction ». In : *Quarterly Journal of the Royal Meteorological Society*, p. 2381-2391 (cf. p. 12).
- Bouttier, François, Laure Raynaud, Olivier Nuissier et Benjamin Ménétrier (2016). « Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX ». In : *Quarterly Journal of the Royal Meteorological Society* 142.S1, p. 390-403 (cf. p. 10, 12).
- Bouttier, François, Benoît Vié, Olivier Nuissier et Laure Raynaud (2012). « Impact of stochastic physics in a convection-permitting ensemble ». In : *Monthly Weather Review* 140.11, p. 3706-3721 (cf. p. 10).
- Breiman, Leo (2001). « Random forests ». In : *Machine learning* 45.1, p. 5-32 (cf. p. 31).
- Brier, Glenn W (1950). « Verification of forecasts expressed in terms of probability ». In : *Monthly weather review* 78.1, p. 1-3 (cf. p. 43).
- Bröcker, Jochen et Leonard A Smith (2007). « Increasing the reliability of reliability diagrams ». In : *Weather and forecasting* 22.3, p. 651-661 (cf. p. 45).
- Buizza, Roberto, J Barkmeijer, TN Palmer et DS Richardson (2000). « Current status and future developments of the ECMWF Ensemble Prediction System ». In : *Meteorological Applications* 7.2, p. 163-175 (cf. p. 10).
- Buizza, Roberto et TN Palmer (1995). « The singular-vector structure of the atmospheric global circulation ». In : *Journal of the Atmospheric Sciences* 52.9, p. 1434-1456 (cf. p. 9).
- Calanca, P, D Bolius, AP Weigel et MA Liniger (2011). « Application of long-range weather forecasts to agricultural decision problems in Europe ». In : *The Journal of Agricultural Science* 149.1, p. 15-22 (cf. p. 33).
- Canal, Nicolas (2014). « Application à l'agriculture de la prévision saisonnière : évaluation à l'échelle de la France ». Thèse de doct. (cf. p. 106).
- Candille, Guillem et Olivier Talagrand (2005). « Evaluation of probabilistic prediction systems for a scalar variable ». In : *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.609, p. 2131-2150 (cf. p. 43).
- Chaulagain, Bhim, Ian Small, James M Shine, Clyde Fraisse, Richard N Raid et Philippe Charles Emile Rott (2019). « Weather-based predictive modeling of orange rust of sugarcane in Florida ». In : *Phytopathology* ja (cf. p. 1).
- Chavent, F. (1983). *Etude de la dynamique des populations d'Eudémis de la vigne (Lobesia botrana Den. et S hi.). Adaptation d'un modèle d'évolution aux conditions climatiques de Provence, Mémoire de fin n d'étude* (cf. p. 14-16).
- Cheung, Jacob, Alan Hally, Jaap Heijstek, Adri Marsman et Jean-Louis Brenguier (2015). « Recommendations on trajectory selection in flight planning based on weather uncertainty ». In : *Proc. 5th SESAR Innovation Days (SID2015), Bologna, Italy*, p. 1-8 (cf. p. 2).
- Christ, Emily H, Peter J Webster, Guy D Collins, Violeta E Toma et Seth A Byrd (2015). « Using Precipitation Forecasts to Irrigate Cotton ». In : *The Journal of Cotton Science*, p. 351-358 (cf. p. 2).
- Clark, Martyn, Subhrendu Gangopadhyay, Lauren Hay, Balaji Rajagopalan et Robert Wilby (2004). « The Schaake shuffle : A method for reconstructing space-time variability in forecasted precipitation and temperature fields ». In : *Journal of Hydrometeorology* 5.1, p. 243-262 (cf. p. 32).
- Courtier, P., C. Freyrier, J.-F Geleyn, F. Rabier et M. Rochas (1991). « The Arpege project at Météo-France ». In : *Proceedings of the ECMWF Workshop on numerical methods in atmospheric models. 9-13 September*, p. 193-231 (cf. p. 10).

- Courtier, Philippe, J-N Thépaut et Anthony Hollingsworth (1994). « A strategy for operational implementation of 4D-Var, using an incremental approach ». In : *Quarterly Journal of the Royal Meteorological Society* 120.519, p. 1367-1387 (cf. p. 7).
- Descamps, L, C Labadie, A Joly, E Bazile, Ph Arbogast et P Cébron (2015). « PEARP, the Météo-France short-range ensemble prediction system ». In : *Quarterly Journal of the Royal Meteorological Society* 141.690, p. 1671-1685 (cf. p. 10, 12).
- Desmazieres, JBHJ (1842). « Neuvieme notice sur quelques plantes cryptogames ». In : *Ann des Sci Nat, Bot Ser 2*, p. 91-118 (cf. p. 16).
- Eyal, Zahir (1999). « The Septoria tritici and Stagonospora nodorum blotch diseases of wheat ». In : *European Journal of Plant Pathology* 105.7, p. 629-641 (cf. p. 16).
- Fundel, Vanessa J, Nadine Fleischhut, Stefan M Herzog, Martin Göber et Renate Hagedorn (2019). « Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users ». In : *Quarterly Journal of the Royal Meteorological Society* (cf. p. 2).
- Gebetsberger, Manuel, Jakob W Messner, Georg J Mayr et Achim Zeileis (2018). « Estimation Methods for Nonhomogeneous Regression Models : Minimum Continuous Ranked Probability Score versus Maximum Likelihood ». In : *Monthly Weather Review* 146.12, p. 4323-4338 (cf. p. 30, 31).
- Ghaffary, MS Tabib (2011). *Efficacy and mapping of resistance to Mycosphaerella graminicola in wheat* (cf. p. 16).
- Ghaffary, Seyed Mahmud Tabib, Olivier Robert, Valerie Laurent, Philippe Lonnet, Eric Margalé, Theo AJ van der Lee, Richard GF Visser et Gert HJ Kema (2011). « Genetic analysis of resistance to Septoria tritici blotch in the French winter wheat cultivars Balance and Apache ». In : *Theoretical and Applied Genetics* 123.5, p. 741-754 (cf. p. 16).
- Giorgino, Toni *et al.* (2009). « Computing and visualizing dynamic time warping alignments in R : the DTW package ». In : *Journal of statistical Software* 31.7, p. 1-24 (cf. p. 37).
- Glahn, Harry R et Dale A Lowry (1972). « The use of model output statistics (MOS) in objective weather forecasting ». In : *Journal of applied meteorology* 11.8, p. 1203-1211 (cf. p. 29).
- Gneiting, Tilmann (2014). *Calibration of medium-range weather forecasts*. European Centre for Medium-Range Weather Forecasts (cf. p. 9, 29).
- Gneiting, Tilmann et Adrian E Raftery (2005). « Weather forecasting with ensemble methods ». In : *Science* 310.5746, p. 248-249 (cf. p. 9, 29).
- Gneiting, Tilmann, Adrian E Raftery, Anton H Westveld III et Tom Goldman (2005). « Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation ». In : *Monthly Weather Review* 133.5, p. 1098-1118 (cf. p. 30).
- Gouache, D., A. Bensadoun, F. Brun, C. Pagé, D. Makowski et D. Wallach (2013). « Modelling climate change impact on Septoria tritici blotch (STB) in France : accounting for climate model and disease model uncertainty ». In : *Agricultural and forest meteorology* 170, p. 242-252 (cf. p. 14, 16).
- Gouache, D et G Couleaud (2009). « Le positionnement des traitements fongicides : enjeu pour la septoriose et intérêt du modèle "SeptoLIS" ». In : *AFPP. 9ème CIMA-Tours, France* (cf. p. 16).
- Gouache, D, G Couleaud *et al.* (2009). « Timing of fungicide treatments against Septoria : stakes and potential of the "SeptoLIS" model. » In : *Association Française de Protection des Plantes, 9ème conférence internationale sur les maladies des plantes, Tours, France, 8 et 9 décembre 2009*. Association Française de Protection des Plantes (AFPP), p. 455-463 (cf. p. 16).
- Hanley, James A (1982). « Characteristic (ROC) curvel ». In : *Radiology* 743.2 (cf. p. 43).
- Hanley, James A et Barbara J McNeil (1982). « The meaning and use of the area under a receiver operating characteristic (ROC) curve. » In : *Radiology* 143.1, p. 29-36 (cf. p. 43).
- Hastie, Trevor, Robert Tibshirani et Jerome Friedman (2009). « Unsupervised learning ». In : *The elements of statistical learning*. Springer, p. 485-585 (cf. p. 29).
- Hersbach, H. (2000). « Decomposition of the continuous ranked probability score for ensemble prediction systems ». In : *Weather and Forecasting* 15.5, p. 559-570 (cf. p. 44).
- Hervé, T. (2011). *Modèle SeptoLIS - Appropriation, comparaisons et analyses*. Arvalis - Institut du végétal, Acta, INRA (cf. p. 17).
- Isaac, Henri et Marine Pouyat (2015). « Les défis de l'agriculture connectée dans une société numérique ». In : *Renais-sance numérique, nov* (cf. p. 1).
- James, Gareth, Daniela Witten, Trevor Hastie et Robert Tibshirani (2013). *An introduction to statistical learning*. T. 112. Springer (cf. p. 29).
- Jolliffe, Ian T et David B Stephenson (2012). *Forecast verification : a practitioner's guide in atmospheric science*. John Wiley Sons (cf. p. 55).
- Jørgensen, Lise Nistrup, GC Nielsen, Jens Erik Ørum, JE Jensen et HO Pinnschmidt (2008). « Integrating disease control in winter wheat—optimizing fungicide input ». In : *Outlooks on Pest management* 19.5, p. 206-213 (cf. p. 16).
- Kollers, Sonja, Bernd Rodemann, Jie Ling, Viktor Korzun, Erhard Ebmeyer, Odile Argillier, Maike Hinze, Jörg Plieske, Dagmar Kulosa, Martin W Ganal *et al.* (2013). « Genetic architecture of resistance to Septoria tritici blotch (*Mycosphaerella graminicola*) in European winter wheat ». In : *Molecular breeding* 32.2, p. 411-423 (cf. p. 16).
- Kuhn, Harold W (1955). « The Hungarian method for the assignment problem ». In : *Naval research logistics quarterly* 2.1-2, p. 83-97 (cf. p. 35).

- Lawless, Conor et Mikhail A Semenov (2005). « Assessing lead-time for predicting wheat growth using a crop simulation model ». In : *Agricultural and forest meteorology* 135.1-4, p. 302-313 (cf. p. 2).
- Leutbecher, M., S-J. Lock, P. Ollinaho, S. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, S. English, M. Fisher, R. M. Forbes, J. Goddard, T. Haiden, R. J. Hogan, S. Juricke, H. Lawrence, D. MacLeod, L. Magnusson, S. Malardel, S. Massart, I. Sandu, P. K. Smolarkiewicz, A. Subramanian, F. Vitart, N. Wedi et A. Weisheimer (2017). « Stochastic representations of model uncertainties at ECMWF : state of the art and future vision ». In : *Quart. J. Roy. Meteor. Soc.* 143, p. 2315-2339 (cf. p. 9).
- Leutbecher, Martin et Tim N Palmer (2008). « Ensemble forecasting ». In : *Journal of Computational Physics* 227.7, p. 3515-3539 (cf. p. 8, 12).
- Linville, DE, CE Drye *et al.* (1995). « Assessment of peanut leaf spot disease control guidelines using climatological data. » In : *Plant disease* 79.9, p. 876-879 (cf. p. 2).
- Lorenz, Edward N (1963). « Deterministic nonperiodic flow ». In : *Journal of the atmospheric sciences* 20.2, p. 130-141 (cf. p. 8).
- Maher, Nevile, Denis Thiery et Erich Städler (2006). « Oviposition by *Lobesia botrana* is stimulated by sugars detected by contact chemoreceptors ». In : *Physiological Entomology* 31.1, p. 14-22 (cf. p. 15).
- Matheson, James E et Robert L Winkler (1976). « Scoring rules for continuous probability distributions ». In : *Management science* 22.10, p. 1087-1096 (cf. p. 43).
- Meinshausen, Nicolai (2006). « Quantile regression forests ». In : *Journal of Machine Learning Research* 7.Jun, p. 983-999 (cf. p. 31).
- Miedaner, Thomas, Yusheng Zhao, Manje Gowda, C Friedrich H Longin, Viktor Korzun, Erhard Ebmeyer, Ebrahim Kazman et Jochen C Reif (2013). « Genetic architecture of resistance to *Septoria tritici* blotch in European wheat ». In : *BMC genomics* 14.1, p. 858 (cf. p. 16).
- Molteni, Franco, Roberto Buizza, Tim N Palmer et Thomas Petroliaigis (1996). « The ECMWF ensemble prediction system : Methodology and validation ». In : *Quarterly journal of the royal meteorological society* 122.529, p. 73-119 (cf. p. 10).
- Moyer, Michelle M, David M Gadoury, Wayne F Wilcox et Robert C Seem (2016). « Weather during critical epidemiological periods and subsequent severity of powdery mildew on grape berries ». In : *Plant disease* 100.1, p. 116-124 (cf. p. 1).
- Murphy, Allan H (1993). « What is a good forecast ? An essay on the nature of goodness in weather forecasting ». In : *Weather and forecasting* 8.2, p. 281-293 (cf. p. 40, 42).
- Murphy, Allan H et Robert L Winkler (1987). « A general framework for forecast verification ». In : *Monthly weather review* 115.7, p. 1330-1338 (cf. p. 40).
- Najmeddine, Hala, Frédéric Suard, Arnaud Jay, Philippe Marechal et Marié Sylvain (2012). « Mesures de similarité pour l'aide à l'analyse des données énergétiques de bâtiments ». In : *RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)*, p. 978-2 (cf. p. 36).
- Norton, Geoff A et John D Mumford (1993). *Decision tools for pest management* (cf. p. 13).
- Ogutu, Geoffrey EO, Wietse HP Franssen, Iwan Supit, P Omondi et Ronald WA Hutjes (2018). « Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts ». In : *Agricultural and forest meteorology* 250, p. 243-261 (cf. p. 2).
- Palmer, Tim (2019). « The ECMWF ensemble prediction system : Looking back (more than) 25 years and projecting forward 25 years ». In : *Quarterly Journal of the Royal Meteorological Society* 145, p. 12-24 (cf. p. 10).
- Pertot, I, Tito Caffi, Vittorio Rossi, L Mugnai, C Hoffmann, MS Grandi, C Gary, D Lafond, C Duso, D Thierry *et al.* (2017). « A critical review of plant protection tools for reducing pesticide use on grapevine and new perspectives for the implementation of IPM in viticulture ». In : *Crop Protection* 97, p. 70-84 (cf. p. 1).
- Picart, Delphine (2009). « Modélisation et estimation des paramètres liés au succès reproducteur d'un ravageur de la vigne (*Lobesia botrana* DEN. & SCHIFF.) » Thèse de doct. Université Sciences et Technologies-Bordeaux I (cf. p. 15).
- Pietravalle, S, MW Shaw, SR Parker et F Van Den Bosch (2003). « Modeling of relationships between weather and *Septoria tritici* epidemics on winter wheat : a critical approach ». In : *Phytopathology* 93.10, p. 1329-1339 (cf. p. 1).
- Pinson, Pierre, H Aa Nielsen, Henrik Madsen et Georges Kariniotakis (2009). « Skill forecasting from ensemble predictions of wind power ». In : *Applied Energy* 86.7-8, p. 1326-1334 (cf. p. 2).
- Raftery, Adrian E, Tilmann Gneiting, Fadoua Balabdaoui et Michael Polakowski (2005). « Using Bayesian model averaging to calibrate forecast ensembles ». In : *Monthly weather review* 133.5, p. 1155-1174 (cf. p. 9, 29).
- Raynaud, Laure et F Bouttier (2016). « Comparison of initial perturbation methods for ensemble prediction at convective scale ». In : *Quarterly Journal of the Royal Meteorological Society* 142.695, p. 854-866 (cf. p. 9).
- Rossi, Vittorio, Tito Caffi et Francesca Salinari (2012). « Helping farmers face the increasing complexity of decision-making for crop protection ». In : *Phytopathologia Mediterranea*, p. 457-479 (cf. p. 1, 14).
- Rossing, WAH, RA Daamen et MJW Jansen (1994). « Uncertainty analysis applied to supervised control of aphids and brown rust in winter wheat. Part 2. Relative importance of different components of uncertainty ». In : *Agricultural systems* 44.4, p. 449-460 (cf. p. 2).
- Roulin, E (2006). « Skill and relative economic value of medium-range hydrological ensemble predictions ». In : (cf. p. 2).
- Roulin, Emmanuel et Stéphane Vannitsem (2005). « Skill of medium-range hydrological ensemble predictions ». In : *Journal of Hydrometeorology* 6.5, p. 729-744 (cf. p. 2).

- Roux, S., F. Brun et D. Wallach (2014). « Combining input uncertainty and residual error in crop model predictions : A case study on vineyards ». In : *European journal of agronomy* 52, p. 191-197 (cf. p. 106).
- Ruth, David P, Bob Glahn, Valery Dagostaro et Kathryn Gilbert (2009). « The performance of MOS in the digital age ». In : *Weather and Forecasting* 24.2, p. 504-519 (cf. p. 29).
- Sakoe, Hiroaki, Seibi Chiba, A Waibel et KF Lee (1990). « Dynamic programming algorithm optimization for spoken word recognition ». In : *Readings in speech recognition* 159, p. 224 (cf. p. 37).
- Schefzik, Roman, Thordis L Thorarinsdottir, Tilmann Gneiting *et al.* (2013). « Uncertainty quantification in complex simulation models using ensemble copula coupling ». In : *Statistical Science* 28.4, p. 616-640 (cf. p. 32).
- Seity, Y, P Brousseau, S Malardel, G Hello, P Bénard, F Bouttier, C Lac et V Masson (2011). « The AROME-France convective-scale operational model ». In : *Monthly Weather Review* 139.3, p. 976-991 (cf. p. 10).
- Semenov, Mikhail A (2004). « Using weather generators in crop modelling ». In : *VII International Symposium on Modelling in Fruit Research and Orchard Management 707*, p. 93-100 (cf. p. 2).
- Shipp, J Leslie et Norman D Clarke (1999). « Decision tools for integrated pest management ». In : *Integrated Pest and Disease Management in Greenhouse Crops*. Springer, p. 168-182 (cf. p. 14).
- Sprague, Roderick (1938). « The status of Septoria graminum ». In : *Mycologia* 30.6, p. 672-678 (cf. p. 16).
- Stanski, Henry R, Laurence J Wilson et William R Burrows (1989). « Survey of common verification methods in meteorology ». In : (cf. p. 42).
- Suffert, F, I Sache et C Lannou (2011). « Early stages of septoria tritici blotch epidemics of winter wheat : build-up, overseasoning, and release of primary inoculum ». In : *Plant Pathology* 60.2, p. 166-177 (cf. p. 16).
- Taillardat, Maxime (2017). « Méthodes Non-Paramétriques de Post-Traitement des Prévisions d'Ensemble ». Thèse de doct. Paris Saclay (cf. p. 31).
- Taillardat, Maxime, Anne-Laure Fougères, Philippe Naveau et Olivier Mestre (2019). « Forest-based and semiparametric methods for the postprocessing of rainfall ensemble forecasting ». In : *Weather and Forecasting* 34, p. 617-634 (cf. p. 9).
- Te Beest, DE, ND Paveley, MW Shaw et F Van Den Bosch (2008). « Disease–weather relationships for powdery mildew and yellow rust on winter wheat ». In : *Phytopathology* 98.5, p. 609-617 (cf. p. 1).
- Thépaut, Jean-Noél et Philippe Courtier (1991). « Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model ». In : *Quarterly Journal of the Royal Meteorological Society* 117.502, p. 1225-1254 (cf. p. 7).
- Thiéry, Denis (2008). « Les tordeuses nuisibles à la vigne ». In : *Féret Publication, Bordeaux, France* (cf. p. 14).
- Thiéry, Denis, Yves Carton, Claire Vidal, Nathalie Gauthier, Sylvie Derridj, Bernard Vercambre, François-Régis Goebel, Jean-Claude Grégoire et François Lieutier (2013). « Histoire de l'installation de quelques ravageurs ». In : (cf. p. 14).
- Thiéry, Denis, Philippe Louâpre, Lucile Muneret, Adrien Rusch, Gilles Sentenac, Fanny Vogelweith, Corentin Iltis et Jérôme Moreau (2018). « Biological protection against grape berry moths. A review ». In : *Agronomy for sustainable development* 38.2, p. 15 (cf. p. 15).
- Thorey, J. (2017). « Ensemble forecasting using sequential aggregation for photovoltaic power applications ». Thèse de doct. Université Pierre & Marie Curie-Paris 6 (cf. p. 107).
- Wetterhall, Fredrik et Francesca Di Giuseppe (2018). « The benefit of seamless forecasts for hydrological predictions over Europe ». In : *Hydrology & Earth System Sciences* 22, p. 3409-3420 (cf. p. 34).
- White, Christopher J., Henrik Carlsen, Andrew W. Robertson, Richard J.T. Klein, Jeffrey K. Lazo, Arun Kumar, Frederic Vitart, Erin Coughlan de Perez, Andrea J. Ray, Virginia Murray, Sukaina Bharwani, Dave MacLeod, Rachel James, Lora Fleming, Andrew P. Morse, Bernd Eggen, Richard Graham, Erik Kjellström, Emily Becker, Kathleen V. Pegion, Neil J. Holbrook, Darryn McEvoy, Michael Depledge, Sarah Perkins-Kirkpatrick, Timothy J. Brown, Roger Street, Lindsey Jones, Tomas A. Remenyi, Indi Hodgson-Johnston, Carlo Buontempo, Rob Lamb, Holger Meinke, Berit Arheimer et Stephen E. Zebiak (2017). « Potential applications of subseasonal-to-seasonal (S2S) predictions ». In : *Meteorological Applications* 24.3, p. 315-325. DOI : 10.1002/met.1654 (cf. p. 106).
- Wilcoxon, Frank, SK Katti et Roberta A Wilcox (1970). « Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test ». In : *Selected tables in mathematical statistics* 1, p. 171-259 (cf. p. 46).
- Wilks, Daniel S. (2015). « Multivariate ensemble Model Output Statistics using empirical copulas ». In : *Quarterly Journal of the Royal Meteorological Society* 141.688, p. 945-952 (cf. p. 31).
- Wilks, Daniel S (2018). « Univariate ensemble postprocessing ». In : *Statistical postprocessing of ensemble forecasts*. Elsevier, p. 49-89 (cf. p. 31).
- Wold, Svante, Kim Esbensen et Paul Geladi (1987). « Principal component analysis ». In : *Chemometrics and intelligent laboratory systems* 2.1-3, p. 37-52 (cf. p. 87).
- Worsnop, Rochelle P, Michael Scheuerer et Thomas M Hamill (2019). « Extended-range probabilistic fire-weather forecasting based on Ensemble Model Output Statistics and Ensemble Copula Coupling ». In : *Monthly Weather Review* 2019 (cf. p. 2).
- Zamo, M. (2016). « Statistical Post-processing of Deterministic and Ensemble Wind Speed Forecasts on a Grid ». Thèse de doct. Université Paris-Saclay (cf. p. 107).
- Zhao, Tongtiengang, Quan J Wang, Andrew Schepen et Morwenna Griffiths (2019). « Ensemble forecasting of monthly and seasonal reference crop evapotranspiration based on global climate model outputs ». In : *Agricultural and forest meteorology* 264, p. 114-124 (cf. p. 2).
- Écophyto (2020). *Ministère de l'agriculture et de l'alimentation, PLAN ÉCOPHYTO II+* (cf. p. 1).

Résumé — Nous proposons de montrer l'intérêt des prévisions d'ensemble météorologiques dans l'anticipation par les agriculteurs d'application de produits phytosanitaires. En nous appuyant sur les modèles EVA relatif au ver de la grappe en vigne et Septo-LIS à la septoriose du blé, nous illustrons le potentiel des prévisions d'ensemble en les comparant aux données fréquentielles. Nous construisons des ensembles de prévision cohérents, dits "sans couture", bénéficiant de trois systèmes de prévision d'ensemble couvrant différentes échelles spatio-temporelles que nous évaluons d'un point de vue météorologique et agronomique. Après calibrage des prévisions par une approche paramétrique, un raccordement sans couture est effectué au moyen d'une mesure de distance entre prévisions et d'un algorithme d'affectation. Nous montrons que l'affectation dite hongroise permet d'obtenir des prévisions cohérentes de température satisfaisant les critères de continuité temporelle et d'unicité des membres avec une meilleure performance météorologique aux courtes échéances. Les prévisions du modèle EVA sont significativement améliorées par le calibrage des prévisions de température. Le gain des prévisions sans couture n'est pas significatif sur cette application.

Mots clés : prévision d'ensemble météorologique, propagation des incertitudes, outils d'aide à décision, scénarios "sans couture", protection des cultures, vigne, blé.

Abstract — This PhD-thesis demonstrates the potential of ensemble weather forecasts in the decision support tools developed to assist farmers in anticipating the application of phytosanitary treatments. We consider the models EVA that simulates the dynamics of grape berry moth and Septo-LIS that forecasts the development of wheat septoria. We illustrate the potential of using ensemble weather forecasts in agronomic models compared to frequency data. We then propose strategies to design seamless ensemble weather forecasts that combine information from three ensembles with different spatio-temporal scales. Finally these seamless forecasts are evaluated from a meteorological and agronomic point of view. The design of seamless ensemble predictions is considered as a concatenation problem. Ensemble predictions are first calibrated using a parametric approach, then the concatenation of forecasts is handled with a distance measure and an assignment algorithm. We show that the so-called Hungarian method is able to provide ensembles of independent and temporally consistent forecasts. It is shown that the EVA model is significantly improved by the calibration of temperature forecasts, while the benefit of seamless forecasts is not significant.

Keywords : ensemble weather prediction, propagation of uncertainty, decision support tools, seamless forecasts, crop protection, vine, wheat.
