

PILOTER LA PREMIERE INTERVENTION SUR ORGE D'HIVER

E. GOURDAIN ⁽¹⁾, G. COULEAUD ⁽¹⁾, F. PIRAUX ⁽¹⁾

⁽¹⁾ ARVALIS-INSTITUT DU VEGETAL 91720 BOIGNEVILLE FRANCE
e.gourdain@arvalisinstitutduvegetal.fr

RÉSUMÉ

Les orges d'hiver subissent en moyenne sur les 10 dernières années 1,7 traitements fongicides en cours de campagne. La stabilité de cet indicateur dans le temps, traduite par un écart type de 0,05, soulève la question de l'adaptation des pratiques aux pressions de l'année avec des variations allant de 7 q/ha de nuisibilité des années comme 2010 et 2011 à presque 25 q/ha en 2007. Par ailleurs, sur cette même période, le profil de sensibilité des variétés a nettement évolué avec l'arrivée de variétés globalement plus résistantes aux maladies. Dans ce contexte, il nous a semblé important de pouvoir proposer une aide au pilotage du premier traitement sur orge à la parcelle. Pour cela, deux modèles de prédictions des risques helminthosporiose et rhynchosporiose sur stades précoces de l'orge, avant sortie des barbes ont été développés. Ces modèles, basés sur une méthode de régression logistique intégrant des co-variables climatiques et agronomiques, ont été paramétrés depuis les données issues des réseaux d'épidémiosurveillance sur l'ensemble du territoire depuis 2010 (Vigicultures® et VGOBS®). Pour gagner en qualité prédictive, nous avons travaillé à l'intégration des données des réseaux d'épidémiosurveillance en cours de campagne dans les modèles. Les résultats de la première année de test ont montré que l'utilisation de tels modèles permettait de réduire considérablement les interventions sur stades précoces avec un taux de faux négatifs de l'ordre de 3%.

Mots-clés : orge, modélisation, pilotage, rhynchosporiose, helminthosporiose.

ABSTRACT

MANAGEMENT OF SPRAYING AT EARLY STAGE OF WINTER BARLEY

Over the last 10 years, the average of fungicide treatments on winter barley during the campaign is 1.7. The stability of this indicator in time, highlighted by a standard deviation of 0.05, raises the question of managing practices to fungus pressures of the year regarding to yield losses variations ranging from 7 q/ha in 2010 and 2011 to almost 25 q/ha in 2007. Moreover, over the same period, the susceptibility profile of varieties has evolved with the arrival of varieties that are more resistant to diseases. In this context, Arvalis felt it was important to be able to provide assistance in piloting the first barley treatment. To do so, two prediction models of net blotch and scald risks on early stages of barley have been developed. These models, based on a logistic regression method incorporating climatic and agronomic co-variables, have been parameterized from data from epidemic monitoring networks throughout the territory since 2010 (Vigicultures® and VGOBS®). To gain in predictive quality, we worked to integrate data from this networks during the campaign into the models. The results showed that the use of such models resulted in a drastic reduction in early stage interventions with a false negative rate of around 3%.

Keywords: barley, modeling, management, scald, net blotch.

INTRODUCTION

Avec 1,4 millions d'hectare et 9 millions de tonnes récoltées en 2017, l'orge d'hiver, brassicole ou fourragère, est la seconde céréales à paille cultivée en France. Les orges d'hiver, comme plus généralement les céréales à paille, sont particulièrement sensibles aux maladies du feuillage qui provoquent des pertes de rendement moyen de 15 q/ha pour cette culture. Cette nuisibilité est compensée aujourd'hui par des interventions fongicides autour de 1,7 interventions par an (Choisir et décider, 2016). La stabilité de cet indicateur depuis ces 10 dernières années, traduite par un écart type de 0,05, soulève la question de l'adaptation des pratiques aux pressions de l'année avec des variations allant de 7 q/ha de nuisibilité des années comme 2010 et 2011 à presque 25 q/ha en 2007. Or, la volonté de réduire le nombre d'interventions phytosanitaires à travers le plan Ecophyto couplé à l'augmentation de la fréquence des épisodes climatiques extrêmes comme ceux du printemps 2016 doit conduire les agriculteurs à raisonner davantage leurs interventions en prenant mieux en compte les contextes pédoclimatiques de leur exploitation et le climat en cours et à venir. Par ailleurs, le profil de sensibilité des variétés d'orge a également évolué depuis 10 ans avec l'arrivée de variétés résistantes. Il convient donc de s'interroger sur l'intérêt des traitements sur cette espèce et de proposer des outils d'aide à la décision aux producteurs s'appuyant sur des modèles prévisionnels de risque d'apparition et de développement des maladies.

Dans chaque processus de modélisation se pose la question de la disponibilité des données nécessaires à la construction d'un modèle. Classiquement, pour modéliser un processus biologique, il est nécessaire de passer par une étape d'acquisition de données jusqu'alors basée sur l'expérimentation et nécessitant souvent plusieurs années. Il est désormais possible d'accéder à de nouvelles sources de données acquises dans une optique généralement très éloignée des questions de modélisation. C'est le cas notamment des données acquises dans les réseaux d'épidémiosurveillance. Ces données permettent de réaliser une analyse de risque afin de connaître l'état sanitaire du territoire en temps réel et permettre la certification phytosanitaire des produits exportés ou circulant dans l'Union Européenne, et de rédiger le Bulletin de Santé du Végétal (BSV) afin de donner aux agriculteurs les moyens de raisonner leurs pratiques et développer des méthodes alternatives. Pour accéder à ces données d'observation, deux outils de saisie et de partage ont été développés dès 2008, Vigicultures® mis en place par les instituts techniques français et qui concernent 11 régions (Siné et al., 2010) et VGOBS® développé par la FNLR pour la Bretagne et les Pays-de-la-Loire. Ces portails collectent chaque semaine les données de plus de 3000 observateurs en France et génèrent ainsi un volume conséquent de données d'observation dont la valorisation se limite jusqu'ici principalement à l'édition du BSV.

Ainsi, ces données spécifiques ont été utilisées pour tenter de répondre à la question de modélisation suivante : peut-on prédire sur une parcelle d'orge d'hiver la nécessité de traiter ? Pour y répondre, deux modèles de prédictions des risques helminthosporiose et rhynchosporiose sur stades précoces de l'orge, avant sortie des barbes, ont été construits. Le choix de ces deux maladies a été dicté par leur nuisibilité importante sur orge d'hiver dans les grands bassins de production français. La rhynchosporiose, *Rhynchosporium secalis*, est une maladie fongique polycyclique dont l'inoculum primaire se trouve sur les résidus des cultures précédentes. De nombreuses générations peuvent se produire sans symptômes apparents. Cette maladie est connue pour s'installer en premier sur les orges cultivées dans des climats humides et frais, puis elle progresse de feuille en feuille vers le haut avec les pluies (Rapilly et al., 1971). L'helminthosporiose, provoquée en France principalement par *Helminthosporium teres*, est également une espèce fongique polycyclique mais pour laquelle l'apparition des symptômes sur feuille est plus tardive car son besoin en température est plus élevé que pour *Rhynchosporium* (Rapilly et al., 1971). La lutte contre ces deux espèces fongiques passe par le choix variétal dont les notes de résistances sont comprises entre 4 et 8 et par la lutte fongicide pour laquelle la première intervention en végétation aurait comme effet la réduction de l'inoculum primaire, quand la seconde intervention autour de l'épiaison permet de protéger les deux dernières feuilles qui contribuent fortement au rendement final et au calibrage de l'orge. Le travail s'est donc focalisé sur la première intervention afin de conseiller une intervention ou non à cette période du cycle de l'orge.

MATERIEL ET MÉTHODE

QUESTION DE MODELISATION

L'enjeu est de proposer un modèle pour prédire sur une parcelle d'orge d'hiver la nécessité de traiter (oui/non) avec un fongicide en phase précoce, soit entre les stades épi 1cm et deux noeuds.

JEUX DE DONNEES

Données d'épidémiosurveillance

L'ajustement des modèles est réalisé à partir des données d'épidémiosurveillance extraites des outils Vigicultures® et VGOBS®, sur les parcelles d'orge d'hiver semées à l'automne des campagnes allant de 2010 à 2018. Les parcelles sont observées de manière hebdomadaire selon un protocole de notation harmonisé en libre accès sur internet qui consiste à observer 20 feuilles des 3 ou 4 derniers étages foliaires sur une placette de la parcelle clairement identifiée et non traitée. Sur chacun de ces étages foliaires, la fréquence de feuilles touchées par la maladie observée est notée. Le tableau I présente le nombre d'observations et de parcelles suivies pour chacune des deux maladies.

Tableau I: Années d'étude, nombre d'observations et nombres de parcelles de données d'épidémiosurveillance extraites de Vigicultures® et VGOBS®

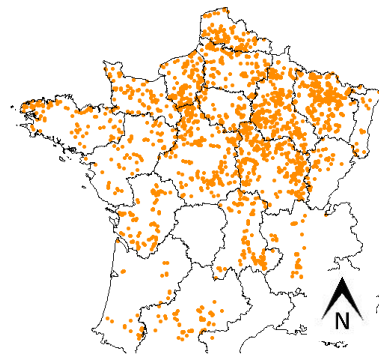
Years, number of observations and number of fields coming from epidemic monitoring networks Vigicultures® and VGOBS®

Maladie	Années d'étude	Nb observations totales	Nb parcelles
Rhynchosporiose	2010-2018	55586	2988
Helminthosporiose	2011-2018	53483	2931

Chaque parcelle est qualifiée par un identifiant, ainsi que par des informations géographiques (localisation, type de sol) et agronomiques (date de semis, nom de la variété, travail du sol, précédent, interventions phytosanitaires le cas échéant). La maladie et la feuille notée sont renseignées. Ainsi une date d'observation génère 3 ou 4 lignes dans la base (une par feuille) identifiée par l'organe, la maladie et la note de fréquence d'attaque. En plus de l'intérêt de ces données de par leur distribution temporelle, la distribution spatiale est aussi intéressante puisque la cartographie des parcelles suivies met en évidence que tous les grands bassins de production d'orge sont représentés (Fig. 1).

Figure 1: Cartographie des parcelles de données d'épidémiosurveillance extraites de Vigicultures® et VGOBS®

Mapping of fields coming from epidemic monitoring networks Vigicultures® and VGOBS®



Jeu de données

La première étape consiste à nettoyer le jeu de données pour supprimer toutes les données erronées. Les suppressions et corrections portent sur les critères agronomiques principalement, ainsi nom de variété et date de semis sont corrigés dans le mesure du possible ou les parcelles supprimées si ces données spécifiques ne sont pas renseignées.

Le jeu de données est ensuite complété afin de répondre à la problématique. Il s'agit de calculer pour chaque parcelle un ensemble de variables agro-climatiques explicatives de la présence de rhynchosporiose et/ou helminthosporiose. L'étude est centrée sur une plage courte entre 2 stades de développement de l'orge, Z30 (stade épi 1cm), et Z32 (stade 2 noeuds) (Zadoks et al., 1974). Les variables sont recherchées dans la littérature ou à dire d'experts si la bibliographie est insuffisante, ce qui est le cas pour ces deux maladies. Dans un premier temps, pour chaque parcelle, les stades clés de la culture sont calculés par un modèle phénologique décrit par Gouache et al. (2012). Ainsi pour chaque parcelle sont calculées les stades levée, épi 1cm (E1C), un nœud (1N) et deux nœuds (2N). Puis

les variables climatiques depuis la date de levée et jusqu'au stade 2N sont calculées sur 4 fenêtres climatiques depuis le semis et jusqu'au stades deux nœuds espacées d'environ 300°J. Le calcul des variables climatiques est basé sur quatre paramètres journaliers, la température, la pluie, l'évotranspiration et le rayonnement. Ces variables sont calculées à partir de données climatiques spatialisées (Deudon *et al.*, 2017) aux points de coordonnées géographiques de chaque parcelle. Ensuite, le nombre de rangs de l'orge (4 ou 6) ainsi que la note de résistance variétale sont récupérés, puis pour cette dernière mise en classe de sensibilité : [4,5] = sensible ; [6-8] = peu sensible. Enfin, chaque parcelle est affectée à une des 5 régions dont les frontières ont été dessinées à dire d'experts afin d'obtenir des zones de production homogènes.

Concernant la variable à expliquer, une variable binaire (0/1) a été créée pour répondre à la question de modélisation. Ainsi pour chaque parcelle et chaque maladie observée, la valeur 1 est affectée si la moyenne de fréquence d'attaque observée entre les stades E1C et 2N est supérieure au seuil d'intervention proposé par le fongiscope orges® dont le principe est décrit dans Henriksen *et al.*, (2000) et les seuils définis en France à dire d'experts. Ce seuil est fonction de la résistance variétale et vaut 10 % de feuilles atteintes pour une variété dite sensible (note < 6) contre 25 % pour une variété dite tolérante (note >= 6). Si la moyenne d'attaque est inférieure à ces seuils, la valeur affectée à la parcelle est 0. Lors des premiers travaux, plusieurs seuils ont été travaillés, 10%, 20% d'attaque mais il s'avère que le seuil proposé par le fongiscope orges® apportait les meilleurs résultats en terme de qualité prédictive des modèles.

Les nombres de parcelles utilisées pour la modélisation sont repris dans le tableau II.

Tableau II : nombre de parcelles pour chaque maladie par année et région.
number of fields per year and region for each disease

année	Helminthosporiose					Rhynchosporiose				
	CENTRE	EST	NORD	OUEST	SUD	CENTRE	EST	NORD	OUEST	SUD
2010	0	1	1	0	4	9	79	40	7	13
2011	33	118	46	16	17	34	116	49	15	23
2012	38	78	57	21	22	33	80	59	21	19
2013	30	89	53	15	27	30	103	52	20	29
2014	28	90	43	13	32	35	109	44	13	32
2015	36	111	55	43	31	40	125	54	44	30
2016	53	114	57	50	49	54	137	61	46	35
2017	46	128	41	49	42	50	156	40	49	39
2018	35	117	35	44	58	42	144	46	46	53

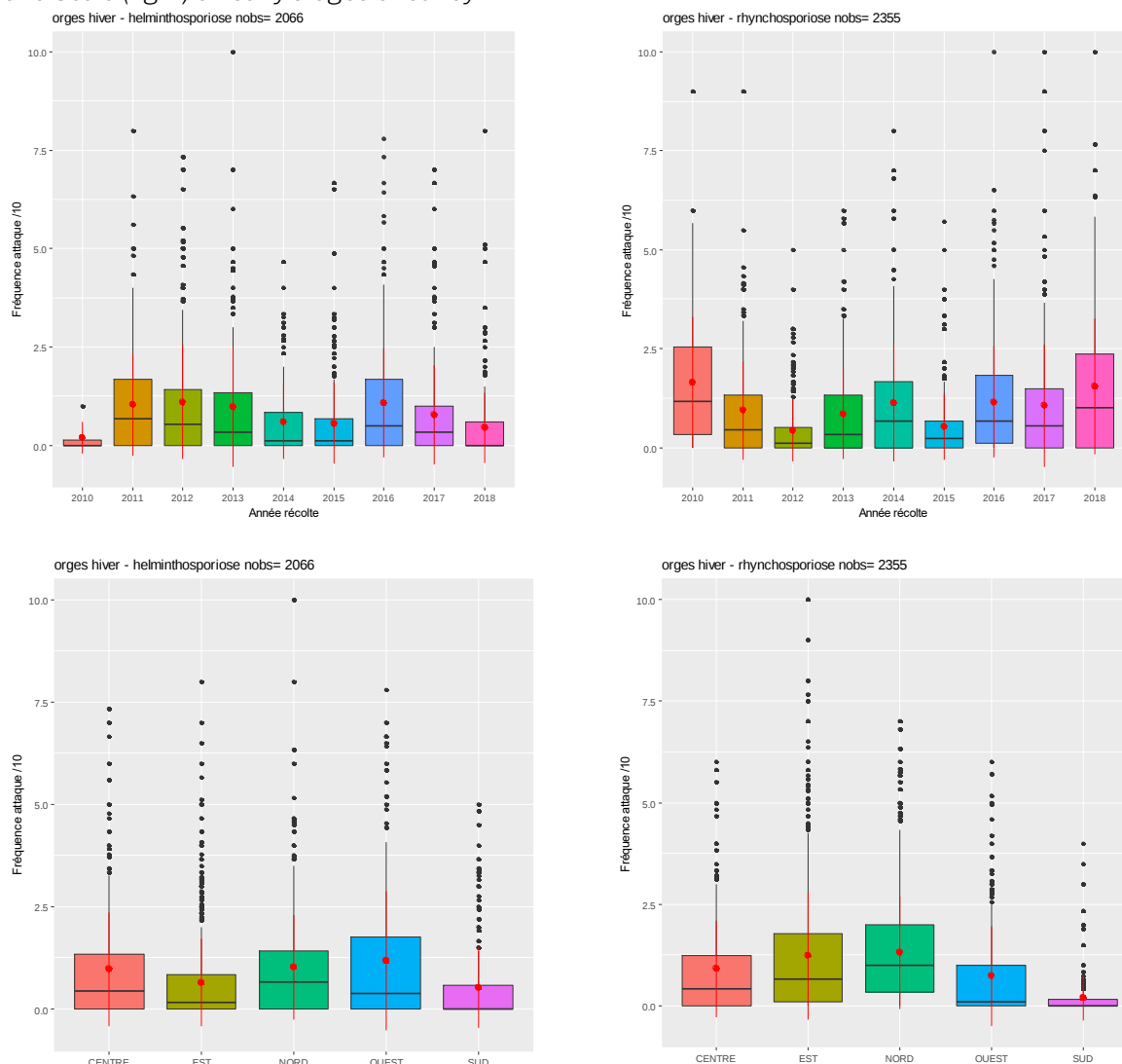
A l'exception de l'année 2010 pour l'helminthosporiose, les régions - années sont assez bien représentées avec un nombre de données supérieur à 10. Le découpage en cinq régions a été réalisé à dire d'experts en fonction des profils variétaux et pressions maladies. La région centre est composée du Centre et du Poitou-Charentes ; la région est de la Bourgogne, de la Champagne-Ardennes et de la Lorraine ; la région nord, de la Haute-Normandie, des Hauts-de-France et l'Île-de-France ; la région ouest de la Bretagne, des Pays-de-la-Loire et de la Basse-Normandie ; enfin, la région sud, de l'Aquitaine, Midi-Pyrénées, de l'Auvergne et de Rhône-Alpes.

Concernant les variétés, 114 variétés sont recensées dans le jeu de données dont les 8 plus fréquentes représentent 60% des données. Il s'agit d'Etincel (27%), Cervoise (9%), KWS Cassia (7%), Esterel (6%), Ketos (5%), Isocel, Arturio et Abondance (3%). 49% des variétés du jeu de données sont classées sensibles en helminthosporiose et en rhynchosporiose.

Enfin, les fréquences de feuilles touchées moyennes entre E1C et 2N (Fig. 2) et le taux de parcelles notée 1 (Tab. III) sont étudiées afin d'évaluer le niveau de variabilité.

Figure 2: variabilité inter-annuelle (haut) et inter-régionale (bas) de fréquence d'attaque pour l'helminthosporiose (gauche) et la rhynchosporiose (droite) sur stades précoces de l'orge.

inter-annual (top) and inter-regional (down) attack frequency variability for net blotch (left) and Scald (right) on early stages of barley



La rhynchosporiose présente plus de variabilité annuelle et régionale avec des années à faible attaque (2012 et 2015) et des régions globalement plus touchées (Est et Nord) alors que pour l'helminthosporiose, le niveau d'attaque semble plus stable d'une année et d'une région sur l'autre.

Tableau III : taux de parcelles qui déclenchent selon la règle Fongiscope orge® pour les deux maladies par année et région

fields rate which need a spraying according to Fongiscope orge® rules for both diseases per year and region

	2010	2011	2012	2013	2014	2015	2016	2017	2018
Helminthosporiose	0%	24%	20%	23%	17%	17%	30%	24%	12%
Rhynchosporiose	43%	24%	6%	23%	34%	18%	37%	36%	44%

	CENTRE	EST	NORD	OUEST	SUD
Helminthosporiose	31%	18%	27%	23%	12%
Rhynchosporiose	33%	35%	40%	16%	4%

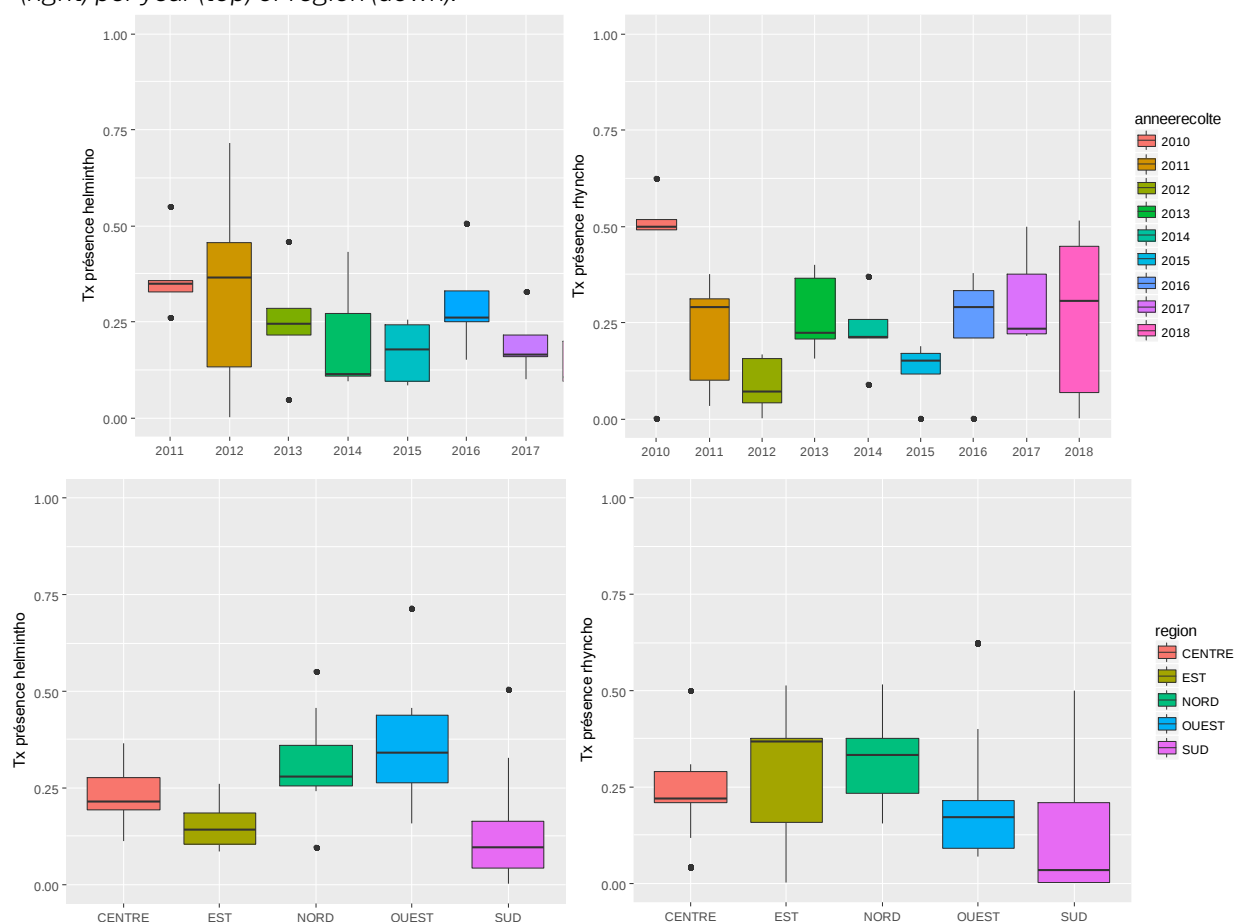
Les deux maladies sont modélisées séparément après vérification par la méthode du Xhi2 de l'indépendance entre la présence des deux maladies.

Plusieurs méthodes de classification ont été testées et présentées dans Gourdain *et al.*, 2018. Parmi les méthodes testées, la régression logistique est apparue comme la méthode présentant les meilleurs résultats pour une mise en œuvre simple. La régression logistique est un cas particulier du modèle linéaire généralisé. Cette méthode de régression permet de modéliser l'effet d'un ensemble de variables sur une variable binaire. Elle modélise la probabilité d'observer l'évènement d'intérêt (ici présence de maladie supérieure au seuil d'intervention) en fonction des variables explicatives. Le modèle utilisé est un modèle de type additif. Pour cette approche, les variables les plus explicatives ont été sélectionnées par la méthode stepwise. Une fois le modèle ajusté aux données, la probabilité prédite par le modèle est utilisée pour décider s'il faut intervenir sur la parcelle ou non. Cela nécessite de fixer un seuil pour la probabilité prédite à partir duquel on considère qu'il est nécessaire de traiter. Afin de prédire au plus juste la présence de maladie, le seuil de probabilité est optimisé par la méthode ROC (bibliothèque ROCR) afin de maximiser la sensibilité (capacité du modèle à prédire la présence de la maladie pour les parcelles malades) et la spécificité (capacité du modèle à ne pas prédire la présence de maladie pour les parcelles saines).

Pour gagner en qualité prédictive, un offset a été ajouté au modèle. Il s'agit en quelque sorte de fournir au modèle un *a priori*, qui dans notre cas est le taux moyen de fréquence de parcelles dépassant le seuil d'intervention l'année donnée dans la région étudiée, ce taux moyen étant corrigé de l'effet variétal. L'utilisation de cet offset est rendu possible car dans un usage du modèle en cours de campagne, les données d'épidémiosurveillance sont remontées au fur et à mesure et les modèles de prévisions peuvent ainsi bénéficier de cette information. Le taux moyen utilisé par année ou par région est fourni en figure 3. Cet offset est calculé au stade 1N, 2N ou en pluriannuel.

Figure 3: boxplot des taux moyens régionaux de fréquence de parcelles dépassant le seuil décisionnel par année (haut) ou des taux moyens annuels de fréquence par région (bas) en Helminthosporiose (gauche) ou rhynchosporiose (droite)

Boxplot of average rate of plots exceeding the decision threshold for net blotch (left) or scald (right) per year (top) or region (down).



Afin de nous assurer de leur qualité prédictive, les modèles sont testés en validation croisée. La validation croisée consiste à retirer une année du jeu de données, d'ajuster le modèle sur les données restantes et de prédire l'année ainsi retirée. Faute de données suffisantes en 2010 en rhynchosporiose, cette année a été retirée de la validation croisée. Le tableau IV représente la matrice de confusion calculée chaque année.

	Prédiction = 0	Prédiction = 1
observation = 0	VN	FP
observation = 1	FN	VP

Tableau IV : matrice de confusion où VN = vrais négatifs, FP = faux positifs (false positives), FN = faux négatifs (false negatives) et VP = vrais positifs (true positives).

Confusion matrix where VN = true negatives, FP = false positives, FN = false negatives and VP = true positives.

A partir de cette matrice, cinq indicateurs d'évaluation sont calculés : la sensibilité ($\text{SENS} = \text{VP}/(\text{VP}+\text{FN})$), la spécificité ($\text{SPEC} = \text{VN}/(\text{VN}+\text{FP})$), le taux de bien classées ($\text{TBC} = (\text{VN}+\text{VP}) / (\text{VN}+\text{FP}+\text{VP}+\text{FN})$), le taux de faux positifs ($\text{TFP} = \text{FP}/(\text{VN}+\text{FP}+\text{VP}+\text{FN})$) ainsi que celui de faux négatifs ($\text{TFN} = \text{FN}/(\text{VN}+\text{FP}+\text{VP}+\text{FN})$).

In fine, les modèles sont ajustés sur l'ensemble des données et peuvent ainsi être utilisés en cours de campagne dans les outils de gestion du risque.

RESULTATS

VALIDATION CROISEE

Les résultats de la validation croisée sur l'ensemble des années pour chacune des deux maladies sont compilés dans les tableaux V et VI. Les indicateurs sont les résultats regroupés pour chaque année en fonction du type d'offset utilisé lors de l'ajustement ainsi que lors de l'évaluation. Dans le tableau V, le même offset (calculé à 2N, 1N, pluriannuel ou égal à zéro) a été utilisé pour l'ajustement et pour l'évaluation sur l'année retirée alors que le tableau VI présente les résultats pour un ajustement avec l'offset de l'année calculé entre E1C et 2N mais lors de l'évaluation sur la nouvelle année, sont testés un offset nul, un offset représentant la pression entre E1C et 1N ainsi qu'un offset correspond à la moyenne pluriannuelle. L'objectif de tester ces différents offset est d'étudier le comportement du modèle en se replaçant en cours de campagne et en imaginant que la pression de l'année n'est pas calculable (offset = nul ou moyenne pluriannuelle) ou seulement sur stades précoces (1N).

Tableau V : résultats de la validation croisée par maladie avec un ajustement et une évaluation des modèles utilisant différents offsets.

Results of cross-validation for each disease with adjustment and evaluation of models using different offsets.

helmintho n=1837				rhyncho n=1953			
offset	SENS	SPEC	TBC	offset	SENS	SPEC	TBC
nul	64%	60%	61%	nul	62%	74%	71%
2N	68%	66%	66%	2N	77%	74%	75%
1N	68%	65%	66%	1N	79%	71%	74%
pluriannuel	60%	61%	61%	pluriannuel	58%	74%	69%

Tableau VI : résultats de la validation croisée par maladie avec un ajustement des modèles utilisant un offset à 2N et un offset pour l'évaluation variable (nul, 1N, pluriannuel)

Results of cross-validation by disease with adjustment of models using 2N offset and variable offset for evaluation (NUL, 1N, multi-year)

helmintho n=1837				rhyncho n=1953			
offset évaluation	SENS	SPEC	TBC	offset évaluation	SENS	SPEC	TBC
nul	87%	32%	44%	nul	89%	56%	66%
1N	69%	61%	63%	1N	75%	75%	75%
pluriannuel	62%	64%	63%	pluriannuel	67%	71%	70%

Les résultats montrent que les meilleures performances en validation croisée sont obtenus pour un offset calculé l'année en cours à 2N avec respectivement un taux de parcelles bien classées de 66% pour l'helminthosporiose à 75% pour la rhynchosporiose. Par ailleurs, ils démontrent bien l'intérêt d'utiliser cette information soit avec une valeur à 2N soit à 1N avec des résultats bien meilleurs que sans.

A noter que ces résultats cachent de fortes disparités annuelles notamment pour l'helminthosporiose où les années 2011 et 2016 sont moins bien prédites.

AJUSTEMENT

Nous décidons après cette première étape de retenir les modèles avec un offset à 2N et réalisons l'ajustement sur l'ensemble de nos données disponibles. Pour chacun des deux modèles la classe de sensibilité variétale est retenue et pèse le plus fortement sur la prévision, viennent ensuite les variables agro-climatiques résumées dans le tableau VII.

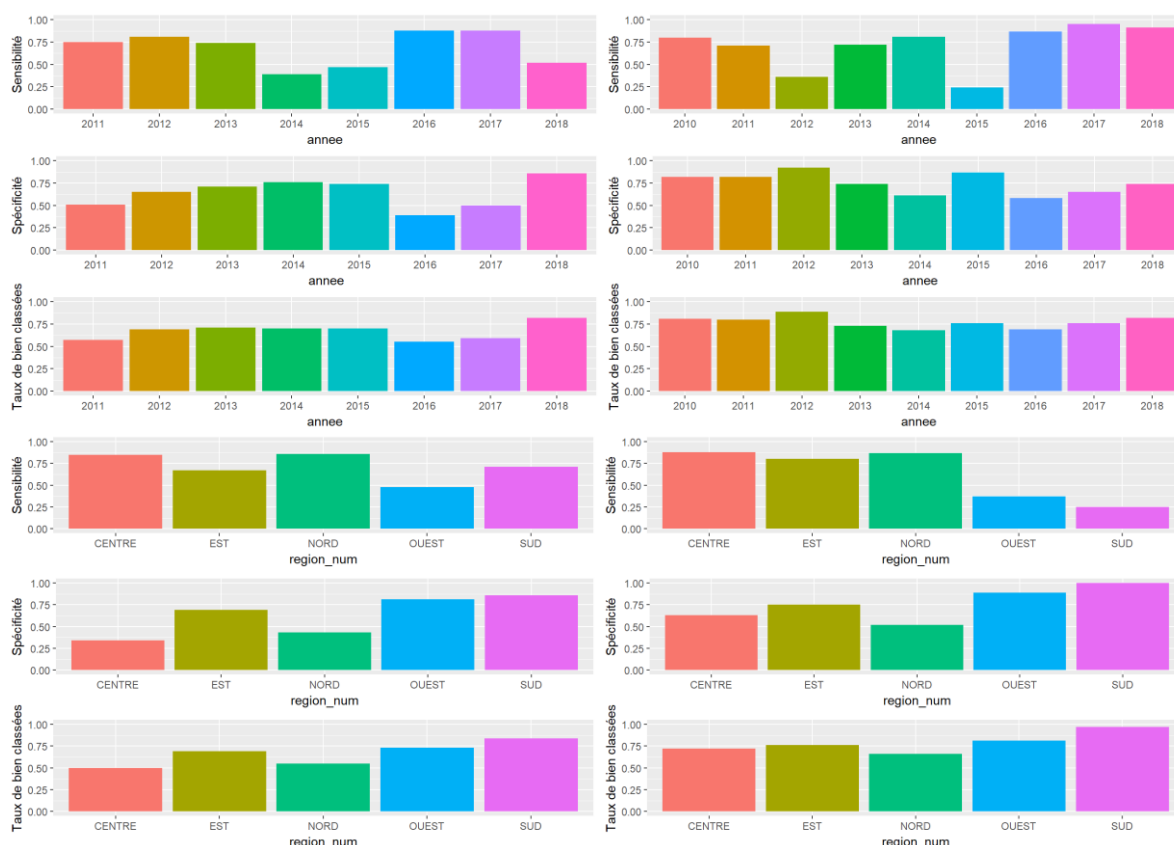
Tableau VII : paramètres climatiques retenus dans les modèles
climatiques parameters selected into both models

Helmintho		
Paramètres	début	fin
Somme Tmax	Semis	levée+150°C
Somme Tmin	levée+150°C	E1C-300°C
Somme Rayonnement	levée+150°C	E1C-300°C
Nb jours pluie et $2 < T_{moy} < 20^{\circ}C$	levée+150°C	E1C-300°C
Somme Tmin	E1C-300°C	E1C
Somme pluie	E1C-300°C	E1C
Nb jours $T_{moy} < 2^{\circ}C$	E1C-300°C	E1C
Rhyncho		
Paramètres	début	fin
Nb de jours	levée	2N
Nb jours $T_{moy} > 20$	Semis	levée+150°C
Somme pluie	levée+150°C	E1C-300°C
Somme Tmax	E1C-300°C	E1C
Nb jours pluie et $2 < T_{moy} < 20^{\circ}C$	E1C-300°C	E1C
Nb jours $T_{moy} < 2^{\circ}C$	E1C	2N

Pour les deux maladies, les températures hivernales sont fortement représentées comme les pluies autour du stade épi 1cm. S'agissant de modèles statistiques nous resterons prudent sur l'interprétation épidémiologique de ces variables.

Le taux de bien classés moyen s'élève à 66% pour l'helminthosporiose (soit un taux d'erreur de 34%) avec 73% de sensibilité contre 76% pour la rhynchosporiose (soit un taux d'erreur de 24%) avec 80% de sensibilité. Les performances annuelles et régionales sont présentées dans la figure 4 à travers l'étude de la sensibilité, spécificité et du taux de bien classées.

Figure 4: performances des modèles (%) ajustés par année (haut) et région (bas) pour l'helminthosporiose (gauche) et la rhynchosporiose (droite)
adjusted model performance (%) by year (top) and region (bottom) for blight (left) and scald (right)



Les modèles ont globalement tendances à surestimer le risque pour chacune des maladies, c'est particulièrement notable en 2016 et 2017, et dans les régions Nord et Centre pour l'helminthosporiose tout comme en 2017 et dans le Nord pour la rhynchosporiose.

DISCUSSION

Les résultats présentés dans cette étude ont mis en évidence qu'après nettoyage et transformation des données acquises dans le cadre des réseaux d'épidémiosurveillance, il était envisageable de développer des modèles prédictifs d'intervention sur helminthosporiose et rhynchosporiose de l'orge. Pour aller plus loin dans l'utilisation de ces deux modèles, nous avons travaillé sur leur combinaison afin de proposer un conseil unique d'intervention ou non dans la parcelle. Deux types de cas ont été testés sur les 1739 parcelles présentant des observations des deux maladies :

- Cas 1 : la parcelle est considérée à traiter si au moins une des deux maladies est présente;
- Cas 2 : la parcelle est considérée à traiter si et seulement si les deux maladies sont présentes.

La difficulté de ce travail complémentaire réside dans l'incapacité à déclarer qu'il fallait effectivement traiter les parcelles. En effet, les données d'épidémiosurveillance ne contiennent aucune information sur la nuisibilité finale. Par ailleurs dans le cas que nous travaillons, seul une expérimentation avec deux modalités T1+T2 et T2 pourrait nous permettre d'évaluer la qualité des deux modèles combinés à prédire un traitement. Ces essais ont été réalisés en 2017 et 2018 et l'analyse reste encore à faire. Néanmoins, nous pouvons à partir de nos données comparer le conseil réalisé sur base des observations au conseil réalisé sur base des prédictions par les modèles (Tab. VIII).

Tableau VIII : indicateurs de qualité prédictive (%) entre conseil lié aux observations et conseil lié aux prédictions selon les deux cas d'étude

predictive quality indicators (%) between observation advice and predictive advice by both study cases

	% Parcelles conseil traitement	SENS	SPEC	TBC	TFP	TFN
cas 1	58%	83%	57%	67%	27%	6%
cas 2	26%	66%	77%	76%	20%	3%

SENS : sensibilité (sensitivity) ; SPEC : spécificité (specificity) ; TBC : taux de bien classés (rate of well graded) ; TFP : taux de faux positifs (rate of false positives) ; TFN : taux de faux négatifs (rate of false negatives).

Les résultats montrent que le cas 1 génère beaucoup de faux positifs (27%) et surestime donc fortement le risque alors que le cas 2 permet de limiter les faux positifs à 20% et les faux négatifs à 3%, élément rassurant pour l'utilisateur. Enfin, pour le cas 2, que nous privilégierons, si on considère qu'actuellement toutes les parcelles sont traitées au stade 2N alors il permettrait de réduire ce taux car le conseil de traitement ne concerne que 26% des parcelles. Cette utilisation du modèle semble donc répondre aux différents enjeux, qui peuvent être parfois contradictoires, à savoir réduire l'usage des produits phytosanitaires tout en sécurisant l'agriculteur dans sa prise de décision.

Néanmoins, malgré les résultats encourageants, il convient de s'interroger sur les explications concernant les taux d'erreur moyens (1 – taux de bien classés) qui restent malgré tout assez marqués avec respectivement 34% pour l'helminthosporiose et 24% pour la rhynchosporiose. La modélisation statistique de phénomènes biologiques se heurte à une difficulté majeure, celles des variables climatiques qui sont souvent corrélées entre elles et qui jouent pour la plupart du temps en interaction. C'est la succession de différents épisodes mêlant pluie, température, rayonnement, etc. qui conduisent au développement d'une maladie sur une culture, le tout en interaction forte avec la plante, sa sensibilité intrinsèque au pathogène mais aussi son stade de développement. Il semble donc que le choix de ces variables doit être regardé plus finement en amont pour en réduire le nombre et ne choisir que les plus pertinentes ainsi que le choix de la méthode de modélisation qui pour certaines peuvent gérer les corrélations et interactions entre elles. Dans notre travail, nous avons testé les variables climatiques en interaction mais les résultats obtenus n'apportaient aucune plus-value par rapport à un modèle additif. Par ailleurs, sur le choix des variables, le faible niveau d'information trouvé dans la littérature nous a imposé un choix un peu arbitraire de variables en lien avec les données climatiques à notre disposition (données de température, pluie journalières). Or, on peut également s'interroger sur ces données climatiques, même si les méthodes de spatialisation utilisées montrent des résultats très bons sur les températures, il n'en est pas de même sur les pluies (Deudon *et al.*, 2017), facteurs très importants pour le développement des pathogènes. Par ailleurs, on peut se poser la question de la pertinence de données journalières plutôt que horaires et de l'absence de l'humidité relative dans nos données souvent documentée comme influente sur les pathosystèmes. Ces limites peuvent expliquer les taux d'erreurs encore élevés sur les modèles les plus performants sélectionnés.

CONCLUSION

En conclusion, les données d'épidémiosurveillance acquises sur le territoire pour l'édition des Bulletins de Santé du Végétal sont une source intéressante de données pour la construction de

modèles prédictifs sous réserve d'avoir au préalable bien posé la question à laquelle le modèle doit répondre afin d'adapter les données en ce sens. Les résultats obtenus en validation croisée montre que ces modèles pourraient être utilisés en cours de campagne sous réserve de disposer en flux continu des données acquises l'année en cours. Un dernier travail doit être réalisé à l'automne sur des expérimentations en microparcelles croisant sensibilité variétale et traitements fongicides afin d'évaluer l'intérêt économique au-delà de l'intérêt environnemental d'utiliser ces modèles.

REMERCIEMENTS

Ce travail n'aurait pu être entrepris sans les animateurs des réseaux Vigicultures® et VGOBS® ainsi que tous les observateurs qui chaque semaine font remonter leurs observations. Par ailleurs, nous remercions la Direction Générale de l'Environnement et la Recherche pour leur contribution financière à ces travaux dans le cadre du projet de recherche SMARTPIC.

BIBLIOGRAPHIE

Choisir et décider 2. Edition National 2016. ARVALIS-Institut du végétal.

Deudon O., Le Bris X., Piraux F., 2017. Interest and implementation of a spatialization method of meteorological data used for in Agricultural Decision Support Tools. Proc. 2017 EFITA CONGRESS – Montpellier, France.

Gouache D., Le Bris X., Bogard M., Deudon O., Pagé C., Gate P., 2012. Evaluating agronomic adaptation options to increasing heat stress under climate change during wheat grain filling in France. European Journal of Agronomy, 39, 62–70.

Henriksen K.E., Jorgensen L.N., Nielsen G.C., 2000. Plant Protection, a tool to reduce fungicide input in winter wheat, winter barley and spring barley in Denmark. Brighton Crop Protection Conference. Pest and diseases, 835-840.

Gourdain E., Piraux F., Couleaud G., Grignon G., Launay M., Deudon D., Gaucher D., Moreau F., Le Bris X., 2017. Datasciences for fungal diseases modeling of cereals. EFITA CONGRESS – Montpellier, France.

Rapilly F., Lemaire H.J.M., Cassini R., 1971. Les principales maladies cryptogamiques des céréales. INRA-ITCF, 186p.

Siné M., Morin E., Simonneau D., Brochard M., De Cosnac G., Escriou H., 2010. VIGICULTURES - An early warning system for crop pest management. Proc Scientific and Technical Information and Rural Development, IAALD XIIIth World Congress, Montpellier, 26-29 avril 2010.

Zadoks J.C., Chang T.T., Konzak C.F., 1974. A decimal code for the growth stages of cereals. Weed research, 14, 415–421.