Statistical modelling of pests and pathogens

By

Ntwampe katlego Levy

Location : Thierval Grignon
Supervised by : Corentin barbu
08/07/2020

Contents

Table of figures

List of tables

Abstract

It is well documented that the pest and pathogens affect the status of field crops, resulting in lower yields. These factors provide an incentive for farmers to use synthetic chemicals such as pesticides and fungicides. However, such control mechanism have detrimental repercussions on the environment and the health of the farmers being the first to be exposed. The aim of this study is to predict statistically the presence and abundance for multiple pests and diseases in cultivated field crops based on climate parameters. The statistical models could provide aid to farmers in order to construct better control mechanisms and assist in the process of decision-making. Our methodology depended on the examination of the French epidemiological data, consisting of 9 years of observations. Thirty-(30), major pathogens and pests of field crops (winter wheat, rapeseed, potato, maize and barely) were jointly analyzed with the climate variables. For most of them, no predictive models existed. Explanatory variables from the weather data that favor the occurrence of the diseases were selected by the Lasso regression. The regression selects temperature and rainfall as the major determinants of the occurrence. The geographical points (longitude and latitude) of the observations are also selected in some models including *Septoria tritici, Sclerotinia sclerotiorum, Phytophthora infestans etc*. On average, we were able to explain 36.8% of diseases (pathogens) presence variation and 35.5 % of pest presence variation. The occurrence of diseases consists of the interaction between pests/pathogens, host crops and environmental conditions. This complex interaction calls for an integration between crop modellers, agronomist and biologist to enhance knowledge and awareness to farmers.

Key words: Regression, Temperature, Rainfall, Pathogens, Pests

Résumé

Il est établi que les ravageurs et les agents pathogènes affectent l'état des cultures, entraînant une baisse des rendements. Ces facteurs incitent les agriculteurs à utiliser des produits chimiques de synthèse tels que les pesticides et les fongicides. Cependant, ces produits entraînent des répercussions néfastes sur l'environnement ainsi que sur la santé des agriculteurs, ces dernier étant les premiers à y être exposés. L'objectif de cette étude est de prédire statistiquement la présence et l'abondance de divers ravageurs et maladies dans les cultures en se basant sur des paramètres climatiques. Les modèles statistiques fournissent un soutien aux agriculteurs afin de construire de meilleurs mécanismes de contrôle et d'aider au processus de prise de décision. Notre méthodologie s'est appuyée sur l'examen des données épidémiologiques françaises constituées de 9 années d'observations. Trente (30) pathogènes et ravageurs majeurs de cultures (blé d'hiver, colza, pomme de terre, maïs et orge) ont été analysés conjointement avec les composantes climatiques. Les variables explicatives des données météorologiques favorisant l'apparition des maladies ont été sélectionnées par la régression Lasso. La régression sélectionne la température et les précipitations comme principaux déterminants de l'apparition des maladies. Les points géographiques (longitude et latitude) des observations sont également sélectionnés dans certains modèles dont *Septoria tritici*, *Sclerotinia sclerotiorum*, *Phytophthora infestans,* etc. En moyenne, nous avons pu expliquer 36,8 % de la variation de présence de maladies (pathogènes) et 35,5 % de la variation de présence de ravageurs. L'apparition des maladies est due à l'interaction entre ravageurs/pathogènes, les cultures hôtes et les conditions

environnementales. Cette interaction complexe nécessite une intégration entre les modélisateurs de cultures, les agronomes et les biologistes afin d'améliorer les connaissances des agriculteurs et les sensibiliser.

Mots clés : Régression, Température, Précipitations, Pathogènes, Ravageurs

Acknowledgement
I would like to offer my sincere gratitude to my supervisor Corentin Barbu and co-supervisor Blanche Collard for their teaching, and continued support throughout the work. I would also like to thank the team working under Corentin Barbu for exchanging ideas throughout the project.

1. Introduction

Bio-aggressors such as pests, pathogens and diseases are among factors that affect the status of field crops (wheat, rapeseed, maize, etc) (Donatelli et al., 2017). These factors are a burden to farmers, and have detrimental repercussions on several components such as yield, with an economical origin. The bio-aggressors provide an incentive for farmers to apply phytosanitary products on field crops. The application of such products has raised a global concern over the human health and environmental status (Rizzati et al., 2016). France's farming system is ranked third in the world for the use of phytosanitary products (Jacquet et al., 2011). Therefore, several studies have proposed the possibility of reducing the use of synthetic chemicals on field crops, which will have a positive impact on the environment, the health of the farmers and local residents, as well as the biodiversity (Pelosi et al., 2013, Jacquet et al., 2011, Delaune et al., 2019). Reducing the use of pesticides and fungicides may be achieved through various farming practices that will favor the natural control of pests and diseases (Delaune et al., 2019) and resilience of crops to the pests and diseases (Perez-Hedo et al., 2017).

Statistical tools including the epidemiological models are regarded as a way to estimate and predict the presence of pests and diseases. The models provide aid to farmers in order to construct better control mechanisms and assist in the process of decision-making (Dalal & Singh, 2017). The study by Tonnang et al. (2017) depicts that the diseases and pest forecasting brings awareness of the actual timing of incidence. This approach helps in achieving quality results in terms of control strategies, and help avoid applying pesticides when the risks are low, with positive impact on economic aspects and environmental benefits. This process is critical as it leads to a sustainable pest control management (Tonnang et al., 2017).

Jacquet et al., (2011) shows that the models of pests and diseases often take into account the landscape. It is evident in the literature that the composition of the landscape plays a vital role in the abundance of pests(Bianchi et al., 2006, Delaune et al., 2019). However, efficient models need to take into account other parameters such as weather and a spatial epidemiology dimension. The landscape composition and dispersal mechanisms have an essential influence in the dynamics of bio-aggressors, which are considered in the inferential process (Werf et al., 1989 ,Blangiardo et al., 2013). This approach makes the combination of time and space important (Delaune et al., 2019). However, studies that include the entire range of interactions that have an influence on the presence of pests and disease are few (Tonnang et al., 2017) and they usually focus on a single pest infestation in relation to a particular crop, or have minimal considerations of time and space data (Tonnang et al., 2017, Delaune et al., 2019)

However, it is hard to routinely develop mechanistic models accounting even only for space and time for multiple pests and diseases, statistical models could then be preferable for large-scale routine approaches.

The autoregressive models are well suited for modelling the presence and abundance of pests and diseases whose distributions are influenced by external and biological parameters(van Maanen & Xu, 2003). Several studies have used the autoregressive models to evaluate the relation between the landscape and pests infestations (Alkindi et al., 2017). These models are able to integrate the relationship between pest infestations and farmer practices such pesticide, fungicide use, methods of spraying (Alkindi et al., 2017) which aligns well with our aim. These autoregressive models are more efficient in providing the prediction of pests and diseases than simple regression (Vinatier et al., 2011). The spatial auto regressive process reflects the spatial and temporal correlations of the disease or pest presence but also in both predictor and response variables, or in the error term(van Maanen & Xu, 2003). These correlations can be observed for each variable at different scales or temporal lags (Dormann et al., 2007). The Integrated Nested Laplace Approximation methodology (INLA method) model is among the auto-regressive models [a fast approximation of the random field based auto-regressive models] (Bakka et al., 2018) and is consequently a good candidate for our routine production of pests and disease models.

However, the methodology of INLA could not be applied fully, due to delays in communication with other partners of the project due to the unfortunate pandemic of Coronavirus. This situation required us to adjust our methodological procedure of analysis. An alternative was chosen to use generalised linear models and the least absolute shrinkage and selection operator commonly known as LASSO (Robert Tibshirani, 1996) The application of lasso is essential as it selects the explanatory variables that have an influence on the response variables (pests and pathogens). A recent review on statistical modeling techniques by (Kim et al., 2014) reflected that Bayesian Lasso technique is suitable for the evaluation of crop pests and pathogens. Additionally, this technique served as prerequisite for Audsley et al.(2005) by developing a simulation model for foliar disease of wheat linked with yield, to contribute to the disease management system. Therefore, this paper has followed recent trends in diseases modelling.

To adjust the statistical models, national scale datasets are available. The agricultural epidemiological services of the French government have organized the mandatory recordings and centralization of the observations of pests and diseases since 2008. Experts in the field on the arable crops carry out the observations. For this study, we analyse the 30 identified common pests and pathogens of field crops.

The aim of this study is to predict statistically the presence and abundance for multiple pests and diseases in cultivated field crops. Our regressions have to account for the spatial attributes (space and time) and the weather as the bases of the analysis. This study serves as a prerequisite in order to develop a complex model that can be robust and efficient. Altogether, several observations different pests and diseases of arable crops were jointly studied with the climate variables.

2. Materials and methods

2.1 Pests and diseases data

The agricultural epidemiological services of the French government organize and centralize the mandatory observations of pests and diseases since 2008. The French agricultural institute ARVALIS-Institut du végétal developed a system called Vigicultures® to centralize a large share of those observations. Together with an other system centralizing data on beetroot, this represents data for beetroot, wheat, barely, maize, potatoes and oilseeds from 2009 to 2018. We attempted unsuccessfully to complement these data for missing regions with an other system (epyphyt) To feed all these databases, different plots of crops are chosen every year, georeferenced and visited each week, to assess the condition of pests and diseases epidemics. Therefore, different observations are made on each crop, in relation to the condition of pest infestations. Several institutes contributes to the observations following a set of national protocols. When several protocols had been used for the same pest, we quantify here the presence and abundance of each pest and diseases using the metric that had most number of observations, *see table 1 and table 2.* In total, 17 diseases affecting winter wheat, rapeseed, potatoes and beetroot were analysed as well as 13 pests that affects maize, winter wheat and rapeseed. For common English names of pest and Diseases See, the supplementary table five.

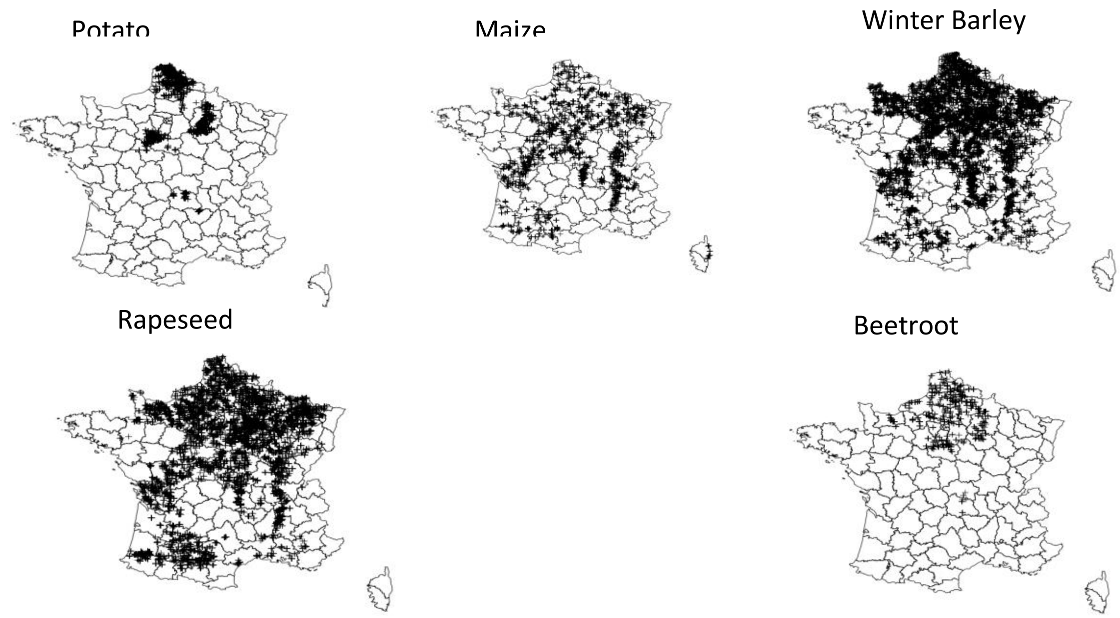


Figure 1.Distribution of agricultural plots over French departments (year: 2010-2017)

Table 1. Pathogens/ diseases data

Crop name	Disease [1]	Observation period[2]	Observation metric[3]
Winter Wheat	*Septoria tritici*	March - June	% of leaves affected
	Puccinia triticina	May - June	% of leaves affected
	Puccinia striiformis	March - June	% of affected leafs
	Fusarium graminearum	March - June	% of affected plants
	Helminthosporium wheat	March - June	% of affected leafs
	Blumeria graminis	February - June	% of organs affected
	Gaeumannomyces graminis	February - June	% of affected roots
	Oculimacula spp	March – June	% of affected plants
Rapeseed	*Leptosphaeria maculan*	February - June	% of plants with leaf macules
	Sclerotinia sclerotiorum	March – May	% of plants affected/stem affected
Beetroot	*Erysiphe betae*	June - August	% of leaves affected
	Uromyces betae	June – August	% of leaves affected-non treated
	Cercospora beticola	June - August	% of leaves affected-non treated
	Ramularia betae	February - June	% of leaves affected-non treated
Winter Barley	*Rhynchosporium secalis*	March – June	% of leaves affected % of leaves leafs
Potatoes	*Phytophthora infestans*	May - October	% of leaves affected

[1] Species of interest
[2] Observations perfomed by experts between 2010 – 2017
[3] Observational metric

Table 2. Pest data

Crop name	pests[1]	Observation period[2]	Obervation metric
Winter wheat	*Sitodiplosis mosellan*	March - June	#[3] captured in yellow trap
	Deroceras/arion/limax	October - December	% plants affected
	Rhopalosiphum padi	October – December	% of affected leafs
	Sitobion avenae	April – June	# captured in yellow trap
Rapeseed			
	Psylliodes chrysocephala	September - November	% plants affected
	Phyllotreta nemorum	September - November	# captured in yellow trap
	Ceutorhynchus napi	September - October	# captured in traps
	Ceutorhynchus picitarsis	September - November	# captured in traps
	Ceutorhynchus assimilis	February - April	# captured in traps
	Meligethes aeneus	February - June	# Average no. of individuals per plant

	Brevicoryne brassicae	February - April	
	myzus persicae	february – June	% plants with presence
		October - November	% plants with presence
Maize			
	*ostrinia nubilali*s	May - October	# captured in traps

[1] species of interest
[2] Observations perfomed by experts between 2010 -2017
[3] Measured metric

2.2 Climate data

The prediction of pests and diseases were analysed based on the climate data in France for the period of 2010 - 2017 at the spatial scale of French departments, based on Safran meteorological model from Météo France. These data initially give all meteorological variables each day on a grid of 8 km distant points. We used a version of these data aggregated by month and French department.

The weather inputs include the following parameters: temperature which is decomposed into average and extremum temperatures (Tmax and Tmin °C), solar radiation, precipitation, evapotranspiration, the number of days with average temperature between 0 and 10° C , the number of days with precipitation and total precipitation. In addition to the monthly variable, we also accounted for groups of months: autumn namely October, November, December (OND) and spring months, April, May, June, July (AMJJ) were averaged. The climate parameters serve as our explanatory variables in predicting the presence and abundance of pests and diseases. These parameters were abbreviated for better visualization on the lasso model; table 3 provides description of the variables.

Table 3. Description of weather variables

Abbreviated variables in the lasso model	Description of the weather variables
pr_per	Average number of rainy days per month
tx34 and tn17	Number of (non-consecutive) days above 34°C (for tx) and number of (non-consecutive) days below -17.2°C (for tn). (lethal temperatures)
tx_0_10	Number of days between 0 and 10°C for Tx (non-consecutive) (vernalization).
tn	Minimum temperature
Tx	Maximum temperature
Rv	radiation
Etp	evapotranspiration
Lat	latitude
Long	longitude

2.3 Pest or disease abundance quantification

The pest or disease, commonly referred to as the bioagresor, abundance is the response variable that is calculated from the vigicultures observational metric. For each variable, we computed the median on all observations across years and points which we applied as the threshold. For each year of observation, we calculated the number of times the observations exceeded the threshold as well as the observations below the threshold creating a binomial variable. The number of positive observations (above threshold) among the total quantity of observations in a given year.

$$Nobs.BelowThreshold = \sum value > threshold$$

&

$$Nobs.AboveThrehold = \sum value \leq threshold$$

2.4 Statistical analysis

The variables were analysed by fitting a generalised linear model (GLM) by means of penalized maximum likelihood (LASSO) utilizing a binomial model of distribution of the observations (Delaune et al., 2019). The LASSO technique allow to jointly select variables and fit models. We used the implementation in the *glmnet* R package; we use standard cross-validation procedures to automatize the choice of the most relevant selection level of the variables. The weather variables were standardized prior to usage in the lasso regression to allow comparison of the respective impacts of the variables on the presence of the pests and diseases.

We estimate the quality of the estimates provided by the models by fitting a linear regression between the observed and the predicted ratios of positive observations over the total number of observations. The number of observations realized at each point weights this regression. The datasets of the pests and pathogens was split into train and hold-out data (20%) in order to evaluate the predictive ability of the model. As this introduces a share of randomness in the process, we evaluate the consistency of the predictive ability of the model by running the whole process 600 times on each pests and pathogen. Therefore we obtained for the adjusted R^2 (adj R^2) the median of the repetitions from the model and its variation interval given by quantile 2.5% and 97.5 % of the predicted and observed values. he train data consisted of random observations amounting to 80% of the data while the test data was 20% of the random observations.

2.5 Lasso Validation

In order to optimize our model for accuracy and better predictive ability we select the lambda parameter by cross-validation *(figure 2).*

Cross deviance validation plot for *Septoria tritici*

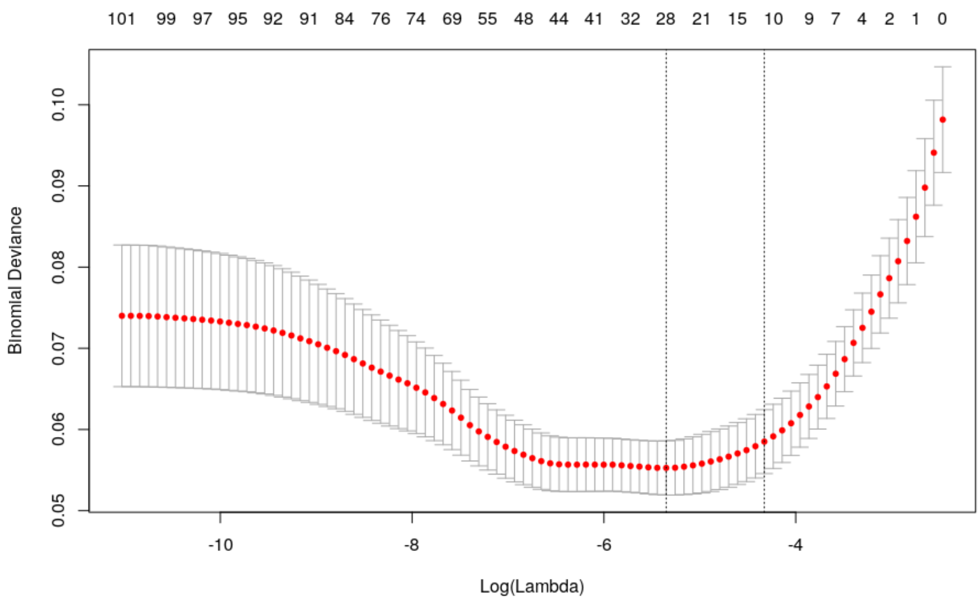


Figure 2. Cross validation plot for lambda selection of Septoria tritici.

Lambda parameter selection for the lasso model of *septoriosis*, used 10 folds cross validation. The two dashed vertical lines shows the two possible choices of the λ (lambda). The left vertical line is the one that minimizes (minimum criteria) the predictive error (denoted as: λmin) and the one on the right show one standard error of the minimum criteria (denoted as λ1se). The top values reflects the numbers of the variables (predictors) that have non-zero coefficients.

We made a choice to use the λ1se for our validation in order to reduce the overfitting factor that may occur during the prediction process. This will give a better accuracy on other datasets.

3. Results

3.1 Pests and diseases model performance

We firstly analysed the prediction model on *Septoria* of winter wheat. Figure 1 depicts that the predicted ratios of observations above the threshold (the median of all observations) for *Septoria trtici (*leaf blotch) correlates with the corresponding observed ratios, which serves as an indication that the predictor model achieves a reasonable goodness of fit.

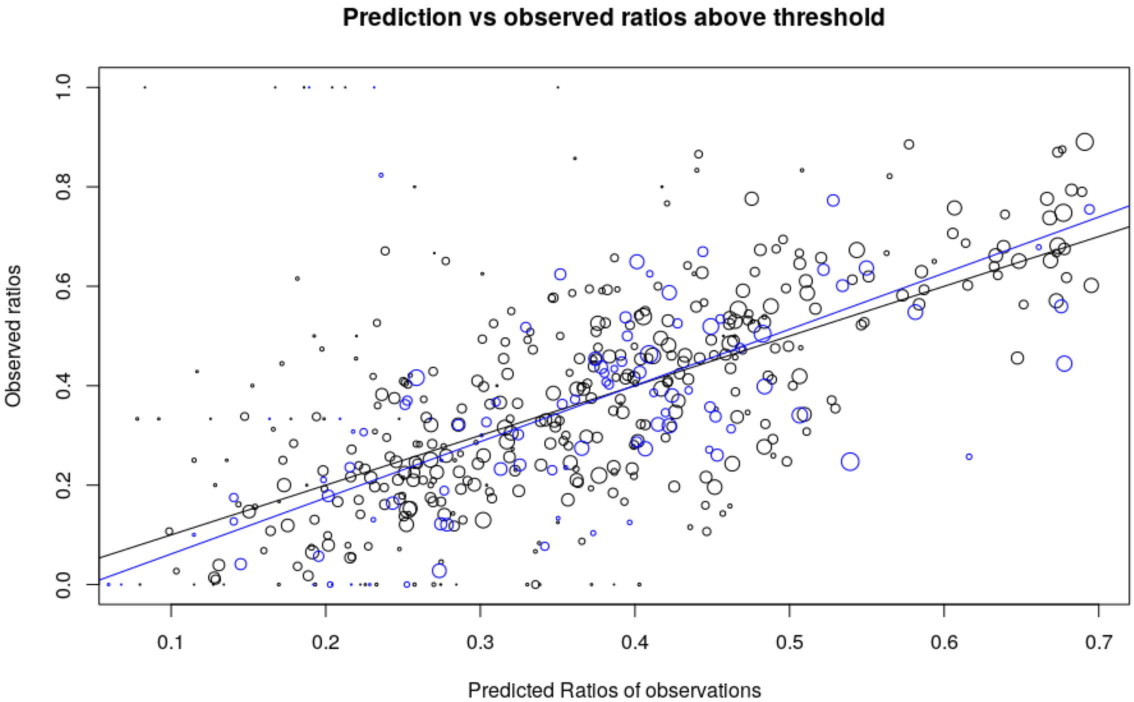


Figure 3. Relationship between observed ratios and predicted ratios of observations for Septoria tritic.

The blue circles are the observed ratios; the black circles are the predicted ratios, with a line of best fit (1:1 relationship). The radius of the points is proportional to the square root of the number of observations *(figure 3).*
The Adjusted R-squared (adj R^2) was utilised in order to quantify the presence/prediction as well as the variability in the response variable shown by the model. The adj R^2 from the model of *septoria* was 0.58 (58%) on data used to fit the model, but went down to 0.50 (50%) for the 20% (n= 96) of holdout data. Therefore, our model was able to account for 58% of the variation explained by our regression line out of the total variation. The R^2 of 0.50 is a moderate value that further accounts that there are other variables responsible for the presence of *Septoria,* although we have a significant trend with climate parameters.

3.2 Lasso coefficients

The lasso model extracted the climatic parameters that have an influence on the presence of *septoria (figure3).*
The lower temperature in January, February, March andNovember, favors the occurrence of S*eptoria* in winter wheat. Increased precipitation in February, March, and April provides suitable conditions for the occurrence of S*eptoria*.
Precipitation in July, October, November, reduces the risk of *Septoria* in wheat. The minimum temperature in September also decreases the risk of S*eptoria.* The number of

days where the Minimum temperature was -17°C decreased the presence of S*eptoria as well as the number* of days where the temperature was between 0 and 10°C for October, November, and December

Lasso plot for *Septoria tritici*

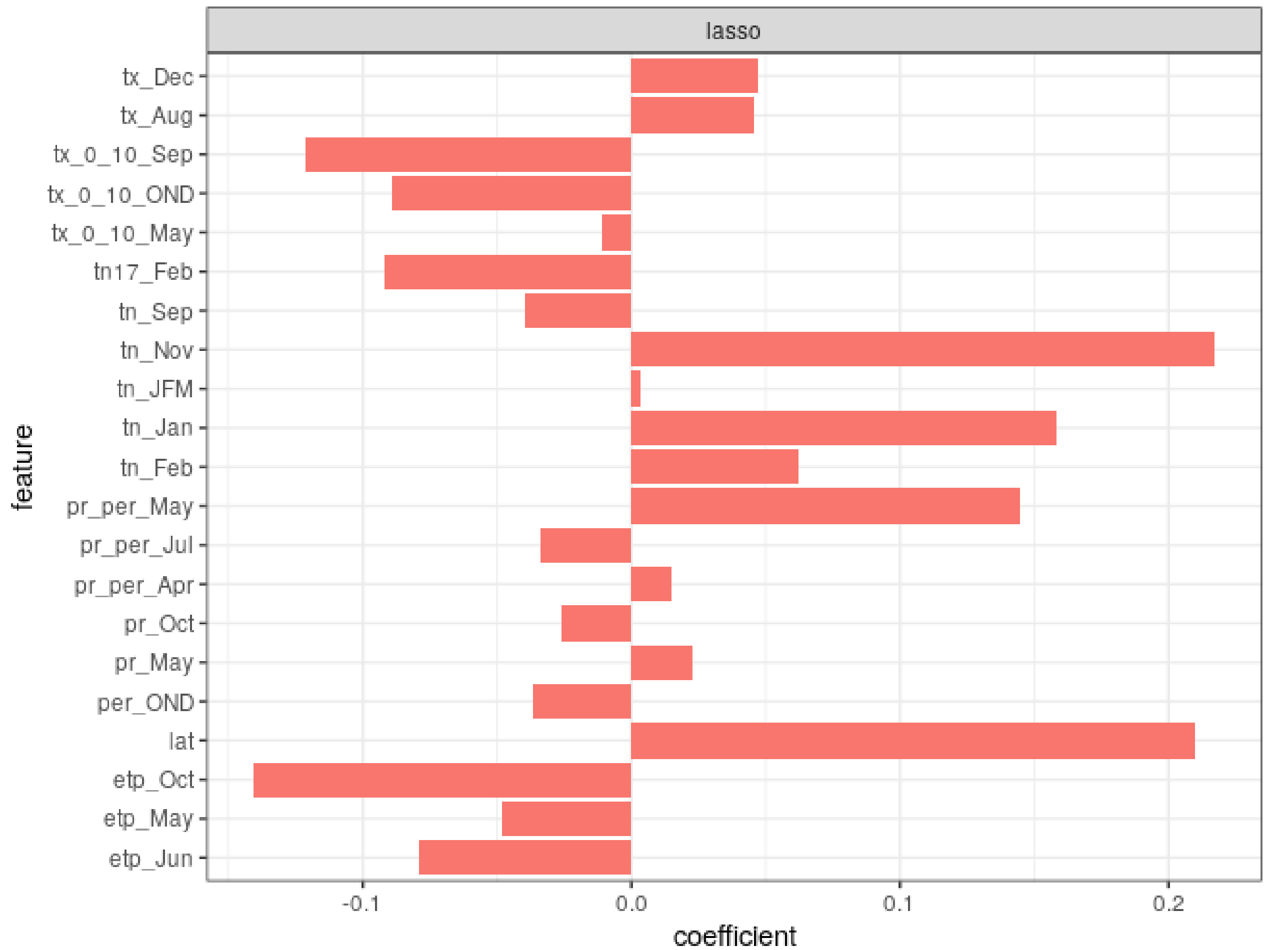


Figure 4. Lasso model with climatic variables for Septoria tritici.

The lasso model of *Septoria tritici* with climatic variables show the coefficients that have an influence on the presence of the pathogen. The positive coefficients imply increased risk of *Septoria* tritici, while the negative coefficients imply lower risk of *Septoria tritici.*

The latitude is a major determinant of the presence of *Septoria* in our data. This is understandable as in metropolitan France; increased latitudes increase the likelihood of lower temperatures and higher precipitations. It is noticeable that the meteorological variables have an impact further and beyond the impact of the latitude. The latitude, the average number of rainy days in May and the minimum temperature in November are the main determinants in the presence of *Septoria*.

3.4 Summary of the results for the rest of the diseases models.

Such models have been run independently for the thirty pests and diseases studied. We summarise hereafter the quality of the fit reached for each of these models (table4 and 5).

Table 4. The proportion of the variation explained by the models for different pathogens

Crop name	Name of the disease	R^2 (Train)[a]	R^2 (Test)[b]
Winter Wheat	*Septoria tritici*	58.6 [51.6, 64.4]	50.3 [31.2, 64.6]
	Puccinia triticina	62.6 [50.8, 69]	52.7[25.4,71.3]
	Puccinia striiformis	59.4 [50.5, 65.6]	49.4 [23.2, 69.2]
	Fusarium graminearum	24.8 [11.5,41]	14.9[1.4,32.9]
	Helminthosporium wheat	23.2 [17.6,30.2]	16 [5.1, 30]
	Blumeria graminis	0 [0, 34.8]	0 [0, 2.6]
	Gaeumannomyces graminis	0 [0, 27.6]	0 [-1.5, 12]
	Oculimacula spp	25.1 [14.6, 33.2]	17.7 [5, 34]
Rapeseed	*Leptosphaeria maculan*	0	0
	Sclerotinia sclerotiorum	46 [40.4, 50.5]	39.4 [22.5, 53.4]
Beetroot	*Erysiphe betae*	0	0
	Uromyces betae	87.9 [0, 95.3]	64.1 [-45.8 ,98]
	Cercospora beticola	73.3 [0, 95.6]	0 [-49.1, 91.3]
	Ramularia betae	0	0
Winter Barley	*Rhynchosporium secalis*	55.1 [49.9, 59.7]	49 [31.5 , 64.1]
	Helminthosporium (oat)	24.9 [13.4, 34.5]	12.6 [1.1, 28.9]
Potatoes	*Phytophthora infestans*	86.3[75.8,95	62.5[5.6,90.7]

[a] The adjusted R squared for train data given as the median based on repetitions of the model, with quantiles of 2.5% to 97.5 %

[b] The adjusted R squared for test data given as the median based on repetitions of the model, with quantiles of 2.5% to 97.5 %

Table . The proportion of variation explained by the models for different pests.

Crop name	Name of the disease	R^2	R^2

		(Train)[a]	(Test)[b]
Winter Wheat	*Sitodiplosis mosellan*	0	0
	Deroceras/arion/limax	45.2 [39.7, 51.1]	35.8 [19.2, 50.1]
	Rhopalosiphum padi	32 [18.6, 44.7]	19 [3.8, 38.8]
	Sitobion avenae	17.7 [0, 24.7]	11.8 [0, 28]
Rapeseed	*Psylliodes chrysocephala*	66 [61.6, 69.7]	61.3 [47.2, 72.4]
	Phyllotreta nemorum	20.7 [16.8, 26.5]	17.4 [5.2, 32]
	Ceutorhynchus napi	45.7 [40.2, 50.4]	39.6 [23.7, 56.3]
	Ceutorhynchus picitarsis	62 [54.6, 70.9]	50.7 [35.3, 64.1]
	Ceutorhynchus assimilis	23.2 [14.8, 31.9]	14.1 [2.5, 31.3]
	Meligethes aeneus	51.6 [47.4, 56]	46 [31.2, 60.6]
	Brevicoryne brassicae	19.9 [13.7, 33.7]	12.7 [1.6, 33.4]
	myzus persicae	56.9 [49.4, 64.4]	45.9 [27.3, 63.2]
Maize	*Ostrinia nubilali*s	21.4 [0, 49.4]	11.2 [-0.8, 30.8]

[a]The adjusted R squared for train data given as the median based on repetitions of the model, with quantiles of 2.5% to 97.5 %

[b]The adjusted R squared for test data given as the median based on repetitions of the model, with quantiles of 2.5% to 97.5 %

The majority of the diseases and pest models achieved a reasonable goodness of fit; however, for a small quantity of diseases the evaluation of goodness of fit was not achieved. The Adjusted R^2 varied from 0% in the *Leptosphaeria maculan, Erysiphe betae, Ramularia betae* to 87% in *Cercospora beticola* of diseases models (*table 4).* For the pest models the values range from 0% in *Sitodiplosis mosellan to 66* % in *Psylliodes chrysocephala* model *(table 5).* In average, we were able to explain 36.8% of diseases (pathogens) presence variation and 35.5 % of pest presence variation. Most of the adjusted R^2 decreased when considering the hold data. This indicates that, there are other factors apart from the climate parameters that have an influence on the abundance and presence of pest and pathogens of field crops.

3.5 Climatic Parameters

The lasso model was able to extract the explanatory variables from the weather climate data. These parameters are provided in the supplementary table 6 for each pest or disease. The parameters that are found to be influential for predicting the presence of dieases and pests vary among the models. For the pathogens models we were able to extract influential variables for 14 of the 16 pathogens (81%) and 12 pests out of 13, (92%).

We summarised the essential parameters by the magnitude of their impact, on the presence and abundance of particular pest and pathogen. For the pathogen models the following parameters were often:

- Maximum temperature in October (Tx= Tmax °C)
- Minimum temperature in January, March, July
- Number of days between 0 and 10 °C for April
- Precipitation in June, December and January

For pest models, the following parameters have a major influence on the presence and abundance:

- Evapotranspiration in January and March
- Temperatures between 0 and 10 in February and July
- Radiation in March, December
- Minimum temperature in July
- Maximum temperature in October, February, August, January

4. Discussion

The study highlighted that the climatic parameters have significant effects on the presence and abundance of most of the 30 selected pest and pathogens on field crops as observedover several years of field observations. These parameters are essential as they affect both host crops and pests/pathogens, making the interaction more complex *figure 3* (Olatinwo & Hoogenboom, 2014). An example of such interaction is highlighted by Fones & Gurr, (2015), who showed the presence of *septoria tritici* is well influenced by the leaf moisture and consecutive number of rainy days in order to spread throughout the hosts crops. Additionally the pathogens/ diseases that affect the aerial parts of the crops are a subject of interaction with climatic conditions such as *Erysiphe betae* (powdery mildew) (Kumar, 2016). Temperature is also shown to play a vital role for the stripe rust of wheat which is caused by *Puccinia striiformis*, however canopy temperature has a role in stripe rust which indicates the complex interactions of the host crops and its pathogens (Cheng et al., 2015). These findings are also in agreement with a study by Chaloner et al., (2019) which highlights the role of temperature and rainfall on the initial occurrence and development of the pathogens including *septoria tritici*. This condition shows the dependence of the pathogen on complex factors that are not limited to temperature but include moisture and light which is linked to the evapotranspiration mechanism (indicated by etp in the figure 2). Therefore, our study is part of the several attempts that are being applied in order to model and understands the prediction and effects of the weather on the presence of pests and pathogens.

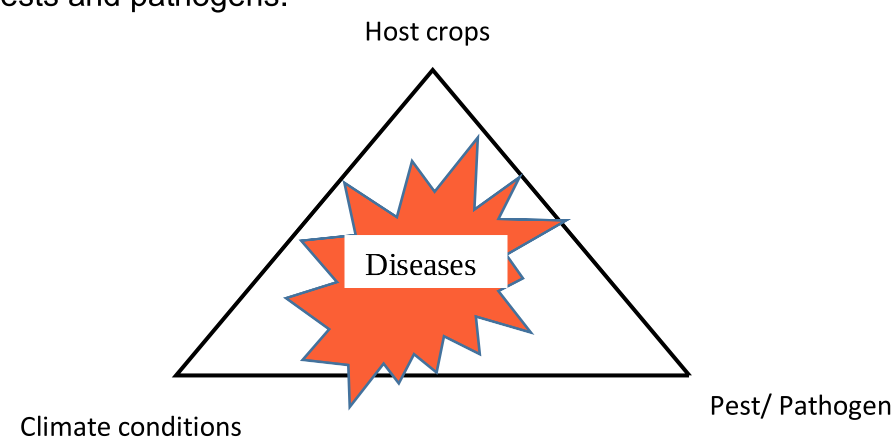


Figure 5. Schematic representation of the relationship between host, climate and pets/pathogens (N.Vharddwick, 1998)

Beest et al. (2009) Reported a positive predictive proportion of 0.61 based on the relationship between weather and occurrence of S*eptoria tritici* which is in line with our predictive proportion of 0.58 – 0.64. This indicates that our model was able to perform fairly. However, most of the published models took into account the experimental observations under controlled conditions and resistant cultivars e.g. greenhouses, which is not the case for our study. This also accounts for a lot of variation within our proportions and published proportions. The typical differences are also highlighted by (Velásquez et al., 2018), that the rate of inoculation for rust caused by *Puccinia striiformis* differed in laboratory from in the fields, the controlled temperature of 21 °C at laboratory did not cause any infections whereas in the fields the infection occurred even when temperature is not constant (18°C - 30°C).

It is shown, in the literature that favorable warm climate or gentle winter conditions that supports the occurrence of pests and pathogens are known to expand the utilization of synthetic chemicals (Olatinwo & Hoogenboom, 2014). Therefore, the ability to estimate future presence based on the trend of climate for particular regions is a crucial activity that will give timing and awareness to farmers.

Our study highlights that temperature and rainfall forms part of the most important variables that affects crop pests and pathogens interactions. Most of the presence is linked with favourable weather conditions/patterns such as a rise in humidity, early or late rains (Olatinwo & Hoogenboom, 2014,Skellern et al., 2017, Pandey et al., 2017). The influence of temperature and rainfall on the occurrence of pests as well as their population with the ability of spreading throughout the crops is summarised, see table 6 (Kumar, 2016).

Table 5. The influence of several climatic parameters on pests.

Climatic parameter	*Influence on pests/ diseases*
Temperature (Tmax & Tmin)	Formative rate of pests
Precipitation (rainfall)	Oviposition, adult appearance/emergence,
Water vapour	Egg hatch
Microclimatic variables such as leaf wetness, humidity etc.	Spread and developmental rate of pests and pathogens

4.1 Temperatures

Our regression model was able to select the temperature with its cardinals (minimum and maximum) for the majority pests and diseases. It is determined that for *Septoria*, *eyespot, Sitobion avenae (grain aphid), Bird cherry-oat aphid, limace (slug), Silver scurf,* are some of the diseases that are hugely influenced by temperature, this include the minimum temperatures in January, November, September and maximum temperatures in October and February. High temperatures are essential for the rate of inoculum development in barely powdery mildew (Kumar, 2016). However, low temperatures can support the longevity of certain diseases and spore formation on crops. For example: *Phytophthora infestans (Potato late blight)* may produce spores at low temperatures, as the optimum temperature for germination of potato leaf blight is about 13°C (Kumar, 2016). Winter warmth is additionally known to move the phenology of bugs and pathogens resulting in colonization of crops and earlier spreading of diseases (Ben-Ari et al., 2018). It is essential to mention that a lot of variables in the fall season are important, as they influence the occurrence of diseases.

4.1 Precipitation

Precipitation/rainfall (denoted by *pr in our lasso model*) or determined moisture encourages the development and dispersal of contagious diseases among crops in spring (Ben-Ari et al., 2018). Our models lead us into a convincible speculation that mild autumn/winter favors a development of pathogens and constancy of inoculum (Ben-Ari et al., 2018). It is evident that moisture plays a vital role in order for the infection to occur, especially for foliar pathogens such as *Erysiphe betae (*powdery mildew) (Bebber, 2015). The wet periods, which are linked to the leaf area by creating a wet surface influences the fungal infections, an example, the rust fungus caused by *Puccinia striiformis* only, need the minimum of 5 hours of leaf wetness for infection to happen (Velásquez et al., 2018). The duration of leaf wetness that resembles the formation of pathogens is affected by the dew accumulation in leaves, that it is longer as a results of dew formation (Rowlandson et al., 2015). The pathogen infection rises at maximum temperatures due to the air's ability to hold more water vapor at extreme temperatures (Velásquez et al., 2018). Furthermore, the severity of S*clerotina sclerotiorum* increases as the air humidity rises (Velásquez et al., 2018). However, the mechanism of such infections depend on the characteristics of pests and pathogens as some can form spores at low relative humidity. We therefore, highlight that knowledge about the frequent diseases of the common crops mentioned in this paper is essential.

Our models also picked up the effect of radiation and evapotranspiration in some of the diseases (Rape flea beetle, Rape steem weevil, etc). The evapotranspiration process is influenced by radiation, air temperature and humidity (Todorovic, 2005). However, this process is associated with abiotic stress factors that have negative impacts on the studied crops. Excessive demand of evapotranspiration due to higher temperatures places an immense pressure on the crop cells that pave the way for the invasion of pathogens and pests hence some of models highlighted this effect. These parameters are more frequent and evident although in less magnitude in pests' models (*Brevicoryne brassicae, Meligethes aeneus, Ceutorhynchus napi, Psylliodes chrysocephala).* It shows that this factors have an influence on the growth and reproduction ability of the pests (their population as whole as well as the distribution among field crops), which in turn could have negative impacts on the crops.

4.2 The effect of the geographical points (latitude and longitude)

Among the models that we fitted, few reflected the effect of the latitude and longitude of the geographical points. Latitude is the most important in *Septoria tritici* (Septoria leaf blotch), *Sclerotinia sclerotiorum* (white mold), *Rhynchosporium secalis* (Barely scald), *Phytophthora infestans* (Potato late blight). As mentioned in the results, the increased latitude in the Metropolitan of France favors the lower temperatures and higher precipitations providing favourable conditions for the mentioned pathogens. This includes the distribution of pests among filed crops, it all amounts back to the importance of temperature and humidity but might describe permanent patterns over the years leading to different average levels of inoculums at the beginning of the crop year in different regions.

4.3 Methodological considerations

Our models did not allow predictions for all of the 30 diseases. None of the variables were selected to predict the abundances of *Leptosphaeria maculans, Erysiphe betae, Ramularia betae and Sitodiplosis mosellan*. This does not necessarily imply that the weather parameters are not influential as the number of observations of the variability in the observations might not have been enough to fit these models. In particular, It is well known that *Leptosphaeria maculans* has not been a major issue in the studied years due to excellent varietal resistances to this disease. There is evidence in the literature that this pests are influenced by the interaction of precipitation and temperature in terms of their population and survival rate (Miao et al., 2019). Therefore, our modeling can be improved by adding more data in order to capture all of the pests and pathogens, adding other places of Europe or more years might solve this issue

Further adaptations are possible by increasing the number of explanatory variables (adding for example the wind or the composition of the environnement (Delaune et al., 2019) or taking into account that some variable might have a non-linear relationship with the diseases and pests. For example, the rainfall has a non-linear relationship with the occurrence of some pests and pathogens. This means that when precipitation occurs, several pathogens may occur, but when precipitation cease then the occurrence of such pests may disappear. This phenomenon is highlighted in the study Tydesley et al. (1980) that the occurance of Sptoria tritici is hugely influenced by rainfall, which may also increase the risk of rust development. It is stated that the less rainfall or minimal irrigation may reduce the risk of Septoria tritici. Though that will add the weather variables in our models, it is essential to be considered.

The modelling applied in this study is rather simple than complex. The results and methods of this study could be used as a basis to develop the autoregressive models that are more complex than the applied general linear modelling for this paper. Importantly, the auto-regressive models would allow to partly reflecting epidemiological dynamics between points. The other notion is the prediction based on the phenological stages of crops. It is essential as the pests and pathogens attack the crops at different stages therefore, knowing which stage is susceptible will be of essence.

5. Conclusion and perspectives

Our study highlighted that temperature with its cadinals (Tmax and Tmin °C) and rainfall are the major determinants of the presence and occurrence of pests and pathogens. The interaction of pathogens and hosts crops is complex and essential in understanding the abundance of the diseases. The models gave various variations of positive proportions, which also highlight that climatic conditions are part of the contributions to crop diseases. Therefore, an interaction study of various components such as landscape, climatic parameters that play a vital role in the occurrence of diseases will be a step forward. Interaction between crop modeling experts, agronomist, farmers and biologist is important.

The study had an underlying aim to contribute to the management of diseases in a more sustainable way that is to limit the application of pesticides. It is known that at times the application of such chemicals are not essential when the diseases are not severe or detrimental to crops. Therefore, awareness about such parameters that have high magnitude on crops is important. This type of predictions or early warnings of the invasion of diseases could serve as an advantage to avoid yield loss.

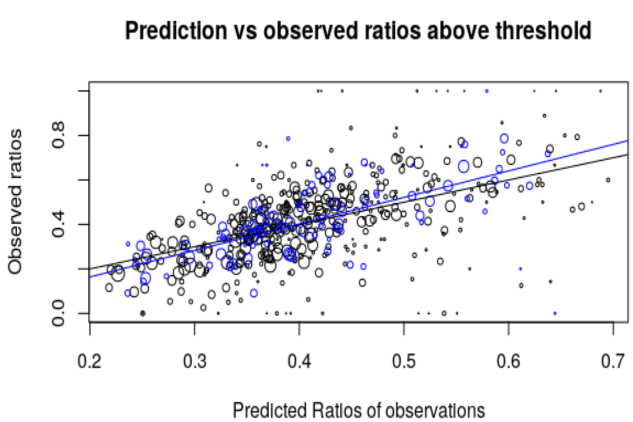
6. Supplementary material

Table 6. Common names of pest and diseases

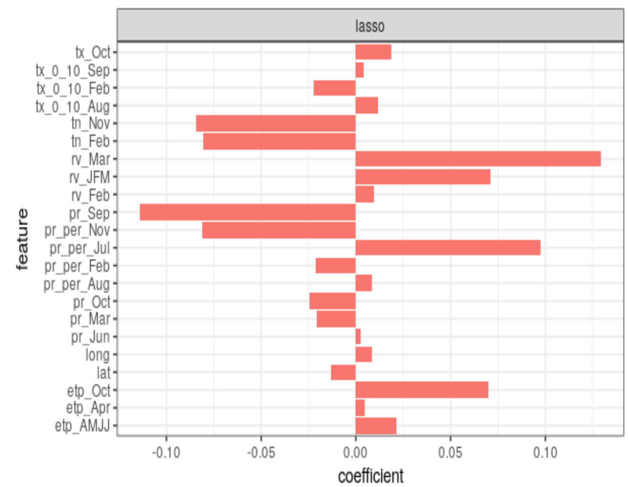
Latin	Common english name
Septoria tritici	Septoria leaf blotch
Puccinia triticina	Brown rust
Puccinia striiformis	Yellow rust
Fusarium graminearum	Fusarium wilt
Helminthosporium wheat	Silver scurf
Blumeria graminis	Brely powdery mildew
Gaeumannomyces graminis	Take-all
Oculimacula spp	Eyespot
Leptosphaeria maculans	Blackleg disease
Sclerotinia sclerotiorum	White mold
Erysiphe betae	Beet powdery mildew
Uromyces betae	Beet rust
Cercospora beticola	Cercospora leaf spot
Ramularia betae	Ramularia leaf spot

Rhynchosporium secalis	Barely scald
Helminthosporium (oat)	Silver scurf
Phytophthora infestans	Potato late blight
Sitodiplosis mosellan	Gall midges
Deroceras/arion/limax	Slug
Rhopalosiphum padi	Bird cherry-oat aphid
Sitobion avenae	Aphid
Psylliodes chrysocephala	Rape flea beetle
Phyllotreta nemorum	Turnip flea beetle
Ceutorhynchus napi	Rape steem weevil
Ceutorhynchus assimilis	Cabbage seed weevil
Ceutorhynchus picitarsis	Rape weevil
Meligethes aeneus	Pollen beetle
Brevicoryne brassicae	Cabbage aphid
myzus persicae	Green peach aphid
*ostrinia nubilali*s	Corn borer

scleretonia

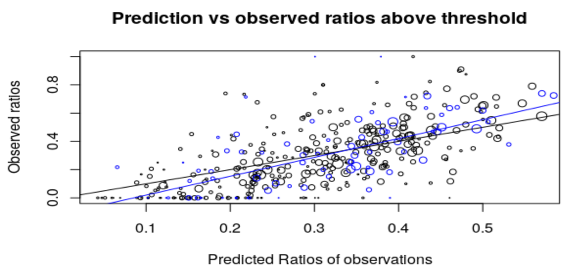


Observed ratios of observations
and predicted ratios

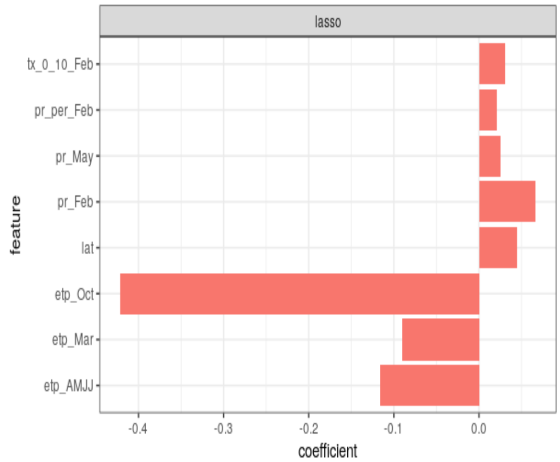


Radiation in march in is major
determinant

Rhynchosporium secalis

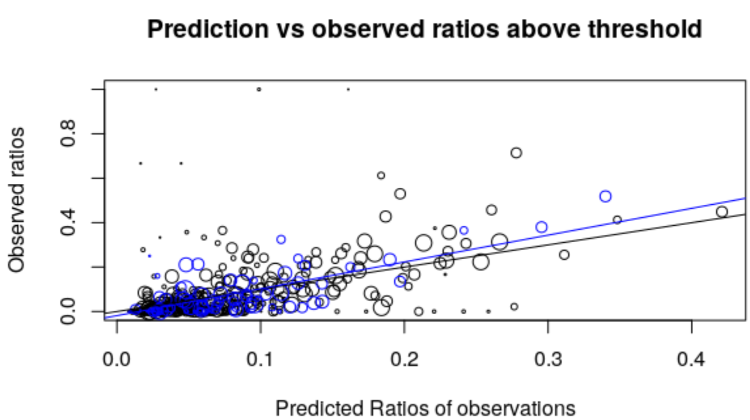


Observed ratios of observations and predicted ratios

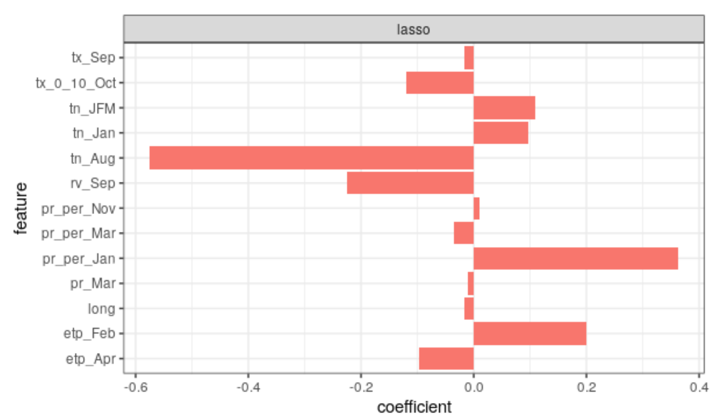


Precipitation in February is major determinant

Puccinia striiformis

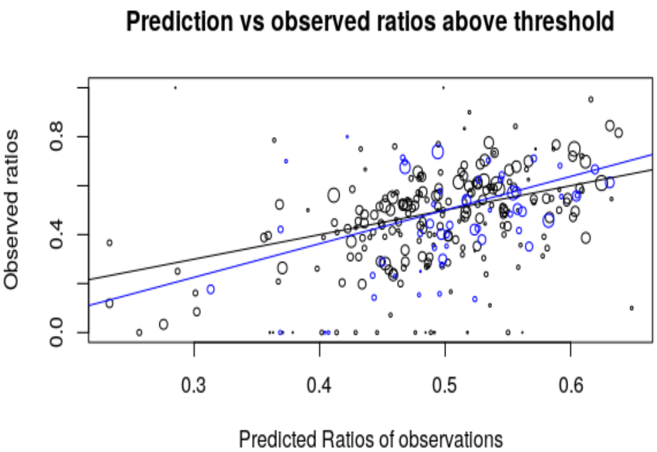


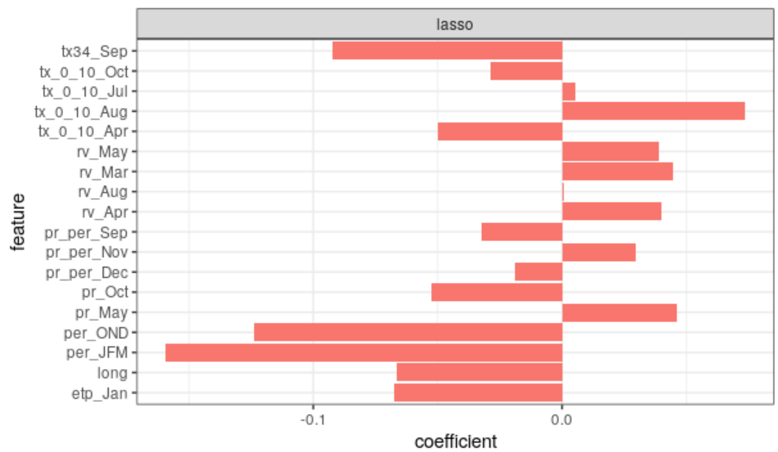
Observed ratios of observations and predicted ratios



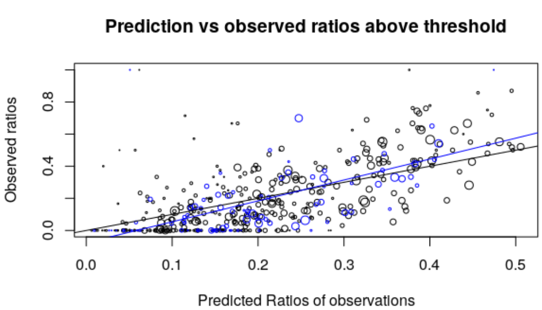
Precipitation in February is major determinant

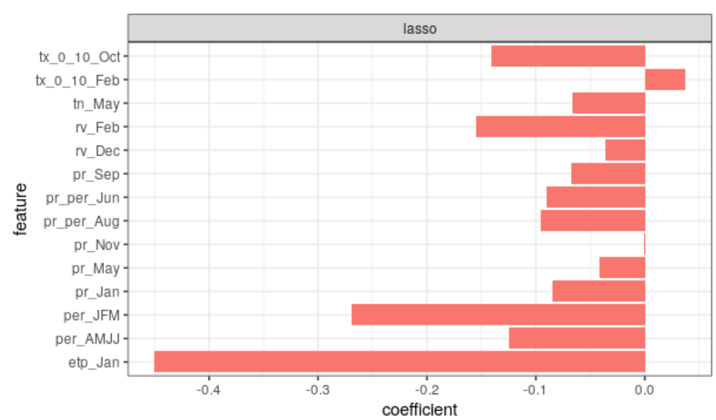
Ostrinia nubilalis





persicae

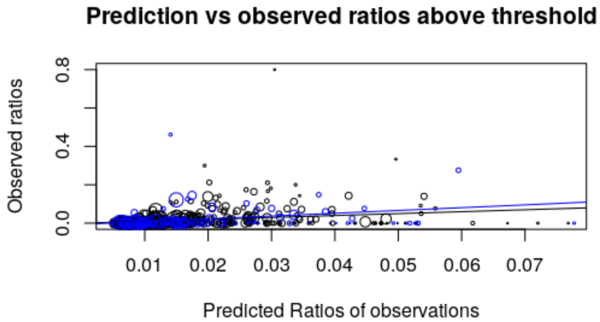


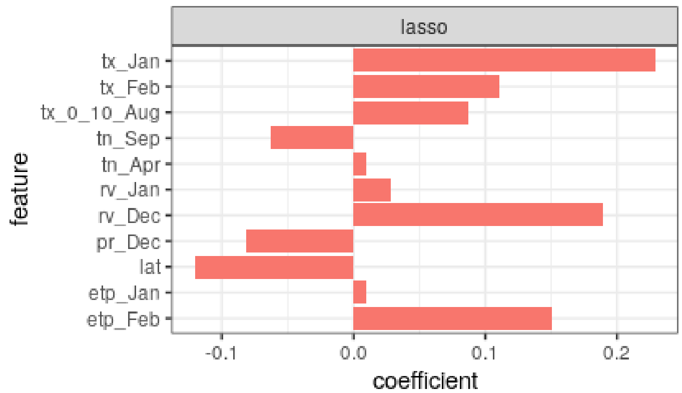


Observed ratios of observations and predicted ratios

Min days in February is major determinant

brevicoyne





Observed ratios of observations and predicted ratios

Minimum temperature in january is major determinant

7. Reference

Alkindi, K., Kwan, P., Andrew, N., & Welch, M. (2017). Remote sensing and spatial statistical techniques for modelling Ommatissus lybicus (Hemiptera: Tropiduchidae) habitat and population densities. *PeerJ*, *5*, e3752. https://doi.org/10.7717/peerj.3752

Audsley, E., Milne, A., & Paveley, N. (2005). A foliar disease model for use in wheat disease management decision support systems. *Annals of Applied Biology*, *147*, 161–172. https://doi.org/10.1111/j.1744-7348.2005.00023.x

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., & Lindgren, F. (2018). Spatial modeling with R-INLA: A review. *WIREs Computational Statistics*, *10*(6), e1443. https://doi.org/10.1002/wics.1443

Bebber, D. (2015). Range-Expanding Pests and Pathogens in a Warming World. *Annual Review of Phytopathology*, *53*. https://doi.org/10.1146/annurev-phyto-080614-120207

Beest, D., Shaw, M., Pietravalle, S., & Bosch, F. (2009). A predictive model for early-warning of Septoria leaf Blotch on winter wheat. *European Journal of Plant Pathology*, *124*, 413–425. https://doi.org/10.1007/s10658-009-9428-0

Ben-Ari, T., Boé, J., Ciais, P., Lecerf, R., Van der Velde, M., & Makowski, D. (2018). Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. *Nature Communications*, *9*(1), 1627. https://doi.org/10.1038/s41467-018-04087-x

Bianchi, F. j. j. a, Booij, C. j. h, & Tscharntke, T. (2006). Sustainable pest regulation in agricultural landscapes: A review on landscape composition, biodiversity and natural pest control. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1595), 1715–1727. https://doi.org/10.1098/rspb.2006.3530

Blangiardo, M., Cameletti, M., Baio, G., & Rue, H. (2013). Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-Temporal Epidemiology*, *4*, 33–49. https://doi.org/10.1016/j.sste.2012.12.001

Chaloner, T. M., Fones, H. N., Varma, V., Bebber, D. P., & Gurr, S. J. (2019). A new mechanistic model of weather-dependent Septoria tritici blotch disease risk. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1775), 20180266. https://doi.org/10.1098/rstb.2018.0266

Cheng, J.-J., Li, H., Ren, B., Zhou, C.-J., Kang, Z.-S., & Huang, L.-L. (2015). Effect of canopy temperature on the stripe rust resistance of wheat. *New Zealand Journal of Crop and Horticultural Science*, *43*, 1–10. https://doi.org/10.1080/01140671.2015.1098708

Dalal, P. K., & Singh, J. (2017). *Role of modeling in insect pest and disease management*. https://www.semanticscholar.org/paper/Role-of-modeling-in-insect-pest-and-disease-Dalal-Singh/a56d3cb3013ad1cac580f390946515fdd8be1064

Delaune, T., Ballot, R., Gouwie, C., Maupas, F., Sausse, C., Félix, I., Brun, F., & Barbu, C. (2019). *Spatiotemporal drivers of crop pests and pathogens abundance at the landscape scale*. https://doi.org/10.1101/641555

Donatelli, M., Magarey, R. D., Bregaglio, S., Willocquet, L., Whish, J. P. M., & Savary, S. (2017). Modelling the impacts of pests and diseases on agricultural systems. *Agricultural Systems*, *155*, 213–224. https://doi.org/10.1016/j.agsy.2017.01.019

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, *30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Fones, H., & Gurr, S. (2015). The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology*, *79*, 3–7. https://doi.org/10.1016/j.fgb.2015.04.004

Jacquet, F., Butault, J.-P., & Guichard, L. (2011). An economic analysis of the possibility of reducing pesticides in French field crops. *Ecological Economics*, *70*(9), 1638–1648. https://doi.org/10.1016/j.ecolecon.2011.04.003

Kim, Y. H., Yoo, S. J., Gu, Y. H., Lim, J. H., Han, D., & Baik, S. W. (2014). Crop Pests Prediction Method Using Regression and Machine Learning Technology: Survey. *IERI Procedia*, *6*, 52–56. https://doi.org/10.1016/j.ieri.2014.03.009

Kumar, A. (2016). *Weather Induced Insects/Pests/Diseases Occurrence & Their Management* (pp. 65–72).

Miao, J., Jianrong, H., Wu, Y., Gong, Z., Li, H., Zhang, G., Duan, Y., Li, T., & Jiang, Y. (2019). Climate factors associated with the population dynamics of Sitodiplosis mosellana (Diptera: Cecidomyiidae) in central China. *Scientific Reports*, *9*. https://doi.org/10.1038/s41598-019-48868-w

Olatinwo, R., & Hoogenboom, G. (2014). Chapter 4—Weather-based Pest Forecasting for Efficient Crop Protection. In D. P. Abrol (Ed.), *Integrated Pest Management* (pp. 59–78). Academic Press. https://doi.org/10.1016/B978-0-12-398529-3.00005-1

Pandey, P., Irulappan, V., Bagavathiannan, M. V., & Senthil-Kumar, M. (2017). Impact of Combined Abiotic and Biotic Stresses on Plant Growth and Avenues for Crop Improvement by Exploiting Physio-morphological Traits. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2017.00537

Pelosi, C., Toutous, L., Chiron, F., Dubs, F., Hedde, M., Muratet, A., Ponge, J.-F., Salmon, S., & Makowski, D. (2013). Reduction of pesticide use can increase earthworm populations in wheat crops in a European temperate region. *Agriculture, Ecosystems & Environment*, *181*, 223–230. https://doi.org/10.1016/j.agee.2013.10.003

Perez-Hedo, M., Suay Cortes, R., Alonso, M., Ruocco, M., Giorgini, M., Poncet, C., & Urbaneja, A. (2017). Resilience and robustness of IPM in protected horticulture in the face of potential invasive pests. *Crop Protection*, *97*, 119–127. https://doi.org/10.1016/j.cropro.2016.11.001

Rizzati, V., Briand, O., Guillou, H., & Gamet-Payrastre, L. (2016). Effects of pesticide mixtures in human and animal models: An update of the recent literature. *Chemico-Biological Interactions*, *254*, 231–246. https://doi.org/10.1016/j.cbi.2016.06.003

Robert Tibshirani, « Regression shrinkage and selection via the lasso », Journal of the Royal Statistical Society. Series B, vol. 58, no 1, 1996, p. 267-288.

Rowlandson, T. L., Gleason, M. L., Sentelhas, P. C., Gillespie, T. J., Thomas, C., & Hornbuckle, B. K. (2015). Reconsidering Leaf Wetness Duration Determination for Plant Disease Management. *Plant Disease*. https://doi.org/10.1094/PDIS-05-14-0529-FE

Skellern, M. P., Welham, S. J., Watts, N. P., & Cook, S. M. (2017). Meteorological and landscape influences on pollen beetle immigration into oilseed rape crops. *Agriculture, Ecosystems & Environment*, *241*, 150–159. https://doi.org/10.1016/j.agee.2017.03.008

Todorovic, M. (2005). Crop Water Requirements. In *Water Encyclopedia* (pp. 557–558). American Cancer Society. https://doi.org/10.1002/047147844X.aw59

Tonnang, H. E. Z., Hervé, B. D. B., Biber-Freudenberger, L., Salifu, D., Subramanian, S., Ngowi, V. B., Guimapi, R. Y. A., Anani, B., Kakmeni, F. M. M., Affognon, H., Niassy, S., Landmann, T., Ndjomatchoua, F. T., Pedro, S. A., Johansson, T., Tanga, C. M., Nana, P., Fiaboe, K. M., Mohamed, S. F., … Borgemeister, C. (2017). Advances in crop insect modelling methods—Towards a whole system approach. *Ecological Modelling*, *354*, 88–103. https://doi.org/10.1016/j.ecolmodel.2017.03.015

Tydesley , J.B., Thomposon, (1980). Forecasting Septoria nodorum on winter wheat in England and wales.

van Maanen, A., & Xu, X. (2003). Modelling Plant Disease Epidemics. *European Journal of Plant Pathology*, *109*, 669–682. https://doi.org/10.1023/A:1026018005613

Velásquez, A. C., Castroverde, C. D. M., & He, S. Y. (2018). Plant–Pathogen Warfare under Changing Climate Conditions. *Current Biology*, *28*(10), R619–R634. https://doi.org/10.1016/j.cub.2018.03.054

Vinatier, F., Tixier, P., Duyck, P.-F., & Lescourret, F. (2011). Factors and mechanisms explaining spatial heterogeneity: A review of methods for insect populations. *Methods in Ecology and Evolution*, *2*(1), 11–22. https://doi.org/10.1111/j.2041-210X.2010.00059.x

Werf, W. van der, Rossing, W. a. H., Rabbinge, R., Jong, M. D. de, & Mols, P. J. M. (1989). Approaches to modelling the spatial dynamics of pests and diseases. *Proc. Congr. Parasitis ’88, Barcelona, 1988, R. Cavalloro, V. Delucchi (Eds.)*, 89–119.