



OPEN Predicting crop disease severity using real time weather variability through machine learning algorithms

Amit Bijlwan^{1,6}, Rajeev Ranjan^{1,7}✉, Manendra Singh^{2,5}, Rahul Purohit³, Arun Jyoti Nath⁴ & Sumit Chakravarty⁵✉

Integrating disease severity with real-time meteorological variables and advanced machine learning techniques has provided valuable predictive insights for assessing disease severity in wheat. This study emphasizes the potential of machine learning models, particularly artificial neural networks (ANN), in predicting wheat disease severity with high accuracy. The field experiment was conducted over two consecutive rabi growing seasons (2023 And 2024) using a randomized block design with four sowing dates to investigate critical weather-disease relationships for two key wheat pathogens: *Puccinia striiformis* f. sp. tritici (yellow rust) and *Blumeria graminis* f. sp. tritici (powdery mildew). Weekly assessments of disease severity were combined with meteorological data and analyzed using ANN and regularized regression models. The ANN model demonstrated superior predictive accuracy for yellow rust and powdery mildew, achieving R-squared values (R^2 of 0.96 And 0.98 for calibration And 0.93 And 0.95 for validation, respectively. Random Forest (RF) models also exhibited robust performance with R^2 values of 0.97 And 0.98 for calibration And 0.93 And 0.90 for validation for yellow rust and powdery mildew, respectively. In contrast, Elastic Net, Lasso, and Ridge regression models showed comparatively moderate predictive capabilities. Principal component analysis (PCA) explained the key meteorological variables influencing disease incidence, with evapotranspiration, temperature, wind speed and humidity emerging as critical factors. Disease prediction is an important aspect of developing a decision support system, and it makes farmers make informed decisions to optimize production.

Keywords Yellow rust, Powdery mildew, Crop weather-disease modelling, Machine learning, Crop management

Wheat is an important global food crop, ranking as the third most-produced cereal after rice and maize and the second most-produced primarily for human consumption¹. Wheat consumption is projected to increase by 11% by 2031, while the production area will expand by only 3%². The Food and Agriculture Organization (FAO) estimates that 20–40% of crops are lost globally due to pests and disease incidence every year³. However, the actual loss in wheat yield is about 13%, mainly caused by pests and diseases⁴. Wheat yield loss is a global concern as new races of rust and powdery mildew pathogens are more virulent and evolve radially^{5–8}. The yield loss caused by powdery mildew in Warsaw And Holland is 1–50% And 1–54%, respectively, in 50 cultivars⁹. In China, wheat yield loss from pests And pathogens ranges from 4.84 to 16.29%¹⁰. Spring wheat was reported to be susceptible to powdery mildew, which reduced the yield by 2.06–26.68% of 23 Egyptian cultivars¹¹. However, under optimal climatic conditions, stripe rust And stem rust can lead to complete yield loss, while leaf rust may result in crop losses ranging from 45 to 50%. Overcoming yield loss with disease and pest resistant varieties takes

¹Department of Agrometeorology, G.B Pant University of Agriculture and Technology, Udham Singh Nagar, 263145 Pantnagar, Uttarakhand, India. ²WOTR Centre for Resilience Studies (W-CReS), Watershed Organization Trust, Pune 411009, Maharashtra, India. ³Department of Plant Pathology, G.B Pant University of Agriculture and Technology, Udham Singh Nagar, 263145 Pantnagar, Uttarakhand, India. ⁴Department of Ecology and Environmental Science, Assam University, Silchar, Assam, India. ⁵Department of Forestry, Uttar Banga Krishi Vishwavidyalaya, 736165 Pundibari, West Bengal, India. ⁶Faculty of Forestry, Birsā Agricultural University, Kanke, 248007 Naugaon, Jharkhand, India. ⁷ICAR-Central Marine Fisheries Research Institute, Kerala, 682018 Kochi, India. ✉email: rajeevranjanagri@gmail.com; c_drsumit@yahoo.com

many years, typically depending on the crop, disease causing pathogen and available technology¹². Estimating yield reductions in wheat crops infected with rust and powdery mildew requires an updated understanding of specific disease progression curve elements for these complex pathosystems. Diseases such as wheat yellow rust (stripe rust), brown rust (leaf rust), and powdery mildew have been recognized as major threats to wheat production worldwide. Among these, the area under the disease progress curve (AUDPC), rate of disease increase (r-value), and final disease severity have been critically linked to wheat yield losses. Higher AUDPC values and faster disease progression rates have been strongly associated with greater reductions in grain yield¹³. Agriculture faces many challenges because of growing environmental concerns related to water use and temperature fluctuations^{14,15}. It is projected that 22 metric tonnes of food grains will be deficient in South Asia by 2030. A recent freeze, followed by hot and dry weather, has adversely affected the winter wheat crop in the region^{16–18}. Historical temperature trends indicate that wheat yield declined by 5.5% from 1980 to 2010, correlating with a 0.13 °C increase in decadal temperatures¹⁹. Meteorological data shows that mean Annual temperatures in regions where wheat, rice, maize, and soybean are cultivated have risen by approximately 1 °C over the last century and are projected to continue rising in the future^{20,21}. The temperature increase is associated with an estimated global wheat yield loss of $6.0 \pm 2.9\%$ per degree Celsius. However, the spatial impacts of temperature increases are highly heterogeneous. Estimated wheat yield losses per degree Celsius are similar to the global average in the United States ($5.5 \pm 4.4\%$) and France ($6.0 \pm 4.2\%$), whereas in India ($9.1 \pm 5.4\%$) and Russia ($7.8 \pm 6.3\%$) are more susceptible to temperature increases^{22–24}.

Powdery mildew is a critical foliar disease of wheat, prevalent in most wheat-growing regions. Infection with powdery mildew can lead to yield reductions of up to 40%, especially under humid rainfed and irrigated conditions^{25,26}. Moreover, wheat blast, caused by the fungal pathogen, poses a significant risk to wheat production in warm and humid regions²⁷. Recent studies have provided a more critical understanding of the factors influencing yield loss due to powdery mildew. Early disease onset has been shown to intensify severity by up to 82%, significantly contributing to yield reductions. In contrast, improved yields are associated with later disease onset, faster progression rates, earlier cultivar maturation, and optimal planting dates. Furthermore, favorable agro-ecological conditions such as increased rainy days in mid to late spring and warmer temperatures during autumn through spring can mitigate the disease's impact and support higher yields^{28–30}. Earlier disease onset, later sowing and maturity dates, lower cultivar resistance, increased rainy days, and higher wind speeds in spring are significantly associated with greater disease severity^{31–34}. Climate change projections for 2040–2070 indicate that wheat blast will primarily spread in the tropics and southern hemisphere as temperatures and relative humidity rise, potentially leading to a 13% annual global loss in wheat production. Moreover, wheat blights are expected to increase in the USA and Mexico. It can impact previously unaffected countries, including Uruguay, Japan, Italy, Spain, and New Zealand. Due to a warmer climate, leaf rust is a major cause of significant wheat yield losses, with occurrence and severity increasing in recent years. The global rise in leaf rust severity across all winter wheat production regions has been a concern³⁵. Leaf rust, a prevalent disease that can occur under diverse climatic conditions, leads to substantial yield losses across large geographic areas³⁶. Over the past decades, leaf rust has been the second most frequently reported disease, followed by powdery mildew in Asia, western and southern Europe and South America^{37,38}. A positive correlation between erratic rainfall patterns and higher temperatures during wheat growing seasons affected the severity of leaf rust in wheat producing regions globally^{39–41}.

The severity assessments of wheat stripe rust and powdery mildew are primarily conducted through visual observations⁴². However, accurately estimating and predicting the lesion area percentage on individual diseased wheat leaves remains challenging when relying solely on conventional severity grading standards for wheat stripe rust. AUDPC, disease onset time, maximum disease incidence and severity, and Gaussian model parameters (b, m, s), are significantly associated with stripe rust progression in wheat crops sown at different planting dates^{43,44}. This way, forecasting models integrating meteorological data have emerged as valuable ICT tools, comparable to mobile applications, decision support systems, and remote sensing technologies commonly employed in modern agricultural extension programs to enhance disease management^{45,46}. Informing farmers about upcoming diseases through integrating weather forecasts and crop disease severity in a particular region using machine learning models is vital to ensuring effective crop protection practices^{47,48}. Disease prediction modelling techniques utilize deep data mining and probability tools combined with weather forecasting, serving as early warning systems for predicting disease and informing farmers in advance through mobile apps and farmers' call centres to take adequate measures⁴⁹. Machine learning (ML) encompasses statistical methods that identify patterns in complex datasets. Wheat crops are critically important to the global economy and food security. Thus, ML techniques help predict disease on a regional basis^{50,51}.

Globally, climate change significantly contributes to wheat yield reduction and is influenced by diseases and pests. Moreover, changes in the plant host environment system caused diseases and pest incidence in unaffected regions globally⁵². Therefore, farmers can anticipate disease outbreaks and implement timely interventions by integrating meteorological data such as temperature, humidity, and rainfall into the current predictive model^{53–55}. However, microclimate-based pest and disease management is an important aspect of developing decision support systems (DSS) to forecast the likelihood of disease occurrence⁵⁶ enabling farmers to apply fungicides more judiciously and reduce unnecessary applications⁵⁷. This practice minimizes the economic burden and mitigates environmental impact^{58,59}. This study aims to evaluate the effectiveness of ANN and regularized regression models in explicating critical weather-disease correlations and predicting the most suitable multiple regression model to predict disease. Wheat yield loss ascribed by yellow rust and powdery mildew is a global concern. Thus, machine learning algorithms or digital technology offers a strategic approach for timely weather-related disease interventions. Disease prediction minimises crop losses and optimises resource use⁶⁰. Disease prediction using weather data is an approach that enhances disease management and supports sustainable

agricultural practices, ensuring long-term productivity and food security. Therefore, our research objective is to quantify the relationship between meteorological variables and wheat crop rust and powdery mildew epidemics.

Material and methodology

Study area

The study was conducted in two consecutive rabi growing seasons in 2023 And 2024 at the Crop Research Center of Govind Ballabh Pant University of Agriculture And Technology, Pantnagar. Geographically, it is located at a longitude of 29.0222°E And latitude of 79.4908°N at An altitude of 243.86 m above sea level in the Shivalik Himalayan foothills of the Terai region in Uttarakhand, India (Fig. 1). The region is weakly developed under the catena of mollic epipedons and grouped under Mollisols order. Parent material is characterized by loamy alluvial sediments with moderately coarse textures^{61,62}. The area is characterised by a humid sub-tropical climate with hot And dry summers from the first week of March to June, May being the hottest month, And a cold season from November to February. Maximum temperatures reach about 45 ± 1.5 °C And 46 ± 1.5 °C during summer during the Kharif season, while a minimum temperature of 2.0 ± 0.5 °C usually occurs in the Rabi season. Frost generally occurs from December to February. The average Annual precipitation is about 1574.2 mm, And a maximum rainfall of 1450 mm (80–90%) occurs during the monsoon season (July to September).

Experimental design and data curation

The field experiment followed a randomized block design with PBW-343 variety utilized with four distinct sowing dates: November 22nd, November 29th, December 6th, And December 13th, across the 2022-23 And 2023–24 growing seasons. The sowing dates were replicated in both growing seasons to provide various environmental conditions for disease progress evaluation and prediction. The study focused on two important wheat diseases: yellow rust and powdery mildew. Disease severity assessments were conducted every week by visual appearance, starting from the first appearance of the disease until maximum severity was observed. To quantify disease severity, standard rating scales were developed for yellow rust (Fig. 2) and powdery mildew (Table 1). Meteorological data includes maximum temperature ($T_{\max}$), minimum temperature ($T_{\min}$), maximum relative humidity ($RH_{\max}$), minimum relative humidity ($RH_{\min}$), sunshine hours (SSH), rainfall, wind speed (WV), and evaporation (Eva), were obtained from the Department of Agrometeorology at Pantnagar under the IMD recognition. Meteorological variables were categorized weekly using daily data to correspond with the disease severity observation dates over the two growing seasons. Four different sowing dates were incorporated in the experimental design to represent the wheat crop, which has a diverse range of environmental conditions. Thus, it is vital to observe the influence of meteorological variables on disease progression. This approach facilitated a comprehensive evaluation of the relationships between prevailing meteorological factors and the development

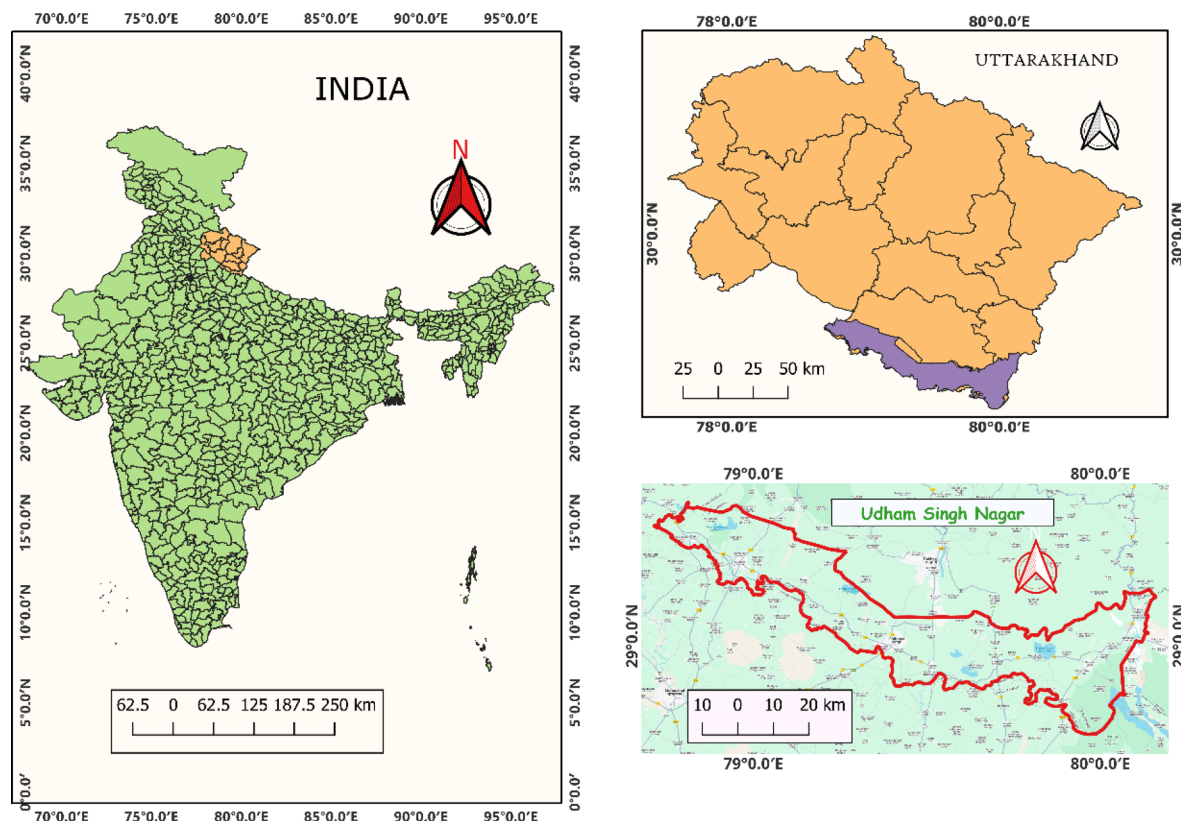


Fig. 1. Spatial layout map of the study area (prepared by QGIS software- version 3.40.6; <https://www.qgis.org>).

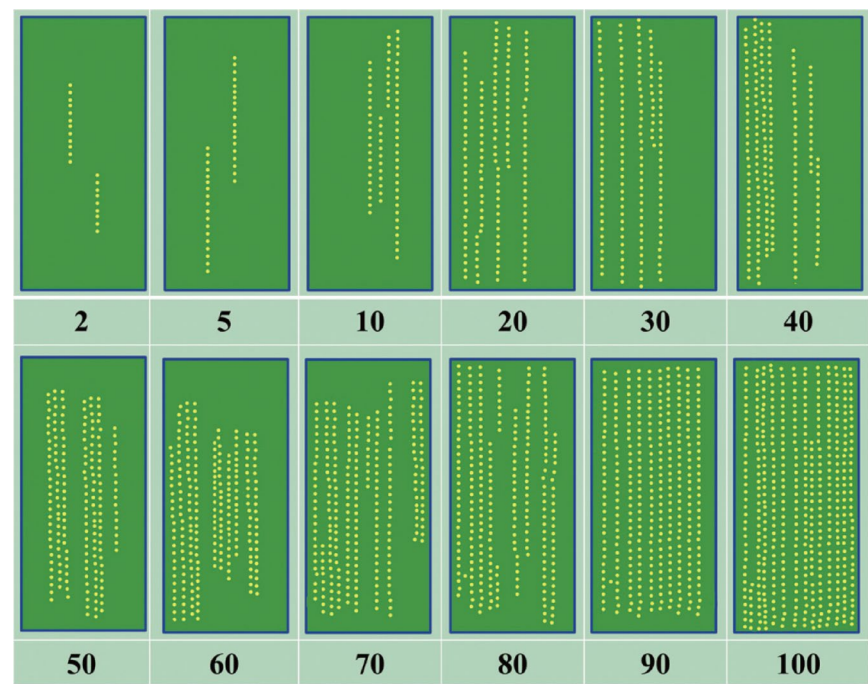


Fig. 2. Scale of rust severity (percent of leaf area infected)⁶³.

IT	Host response	Symptoms
0	Immune (I)	No visible signs or symptoms
1	Highly resistant (HR)	Small spots only
2	Resistant (R)	Chlorosis spots evident
3	Moderately resistant (MR)	Large spots with chlorosis and necrosis
4	Low intermediate (LI)	Mycelium and conidia detectable
5	Intermediate (I)	Small to moderate sized colony and conidia present
6	High intermediate (HI)	Predominance of moderate sized colony and conidia present
7	Moderately susceptible (MS)	At the least 50% of the large colony and conidia are visual
8	Susceptible (S)	75 to 80% of the leaf segment was covered with large colony and conidia
9	Highly susceptible (HS)	100% of the leaf segment covered with large colony and conidia

Table 1. Wheat powdery mildew infection type (IT) according to a (0–9) scale⁶⁴.

of yellow rust and powdery mildew diseases in wheat under climatic conditions (Fig. 3). The objective is not to predict final severity from real-time weather alone, but to capture the relationship between weather conditions in the preceding weeks And the subsequent disease development trajectory. The total dataset consisted of 128 disease scores per year, providing 256 observations per disease.

The average weather conditions during both wheat growing seasons from the onset of disease incidence (Tables 2 and 3). The standard deviation ascribed larger than the mean, indicating the high degree of variability in the severity of yellow rust disease. A wide range of disease severity was observed throughout both growing periods for yellow rust and powdery mildew.

Model development

Artificial neural network (ANN)

The ANN model was developed using the ‘nnet’ package in R. The dataset was split into 70% training And 30% testing sets. The model used 3-fold cross-validation for hyperparameter tuning. Input features were ‘preprocessed’ by ‘centering’ and ‘scaling’. The model architecture was tuned using a grid search over different combinations of hidden layer sizes (1, 3, 5 neurons) and weight decay values (0.1, 0.5, 1). The final model used a linear output layer for regression. The training was performed using the ‘train’ function from the caret package, with the method set to “nnet” (Table 4).

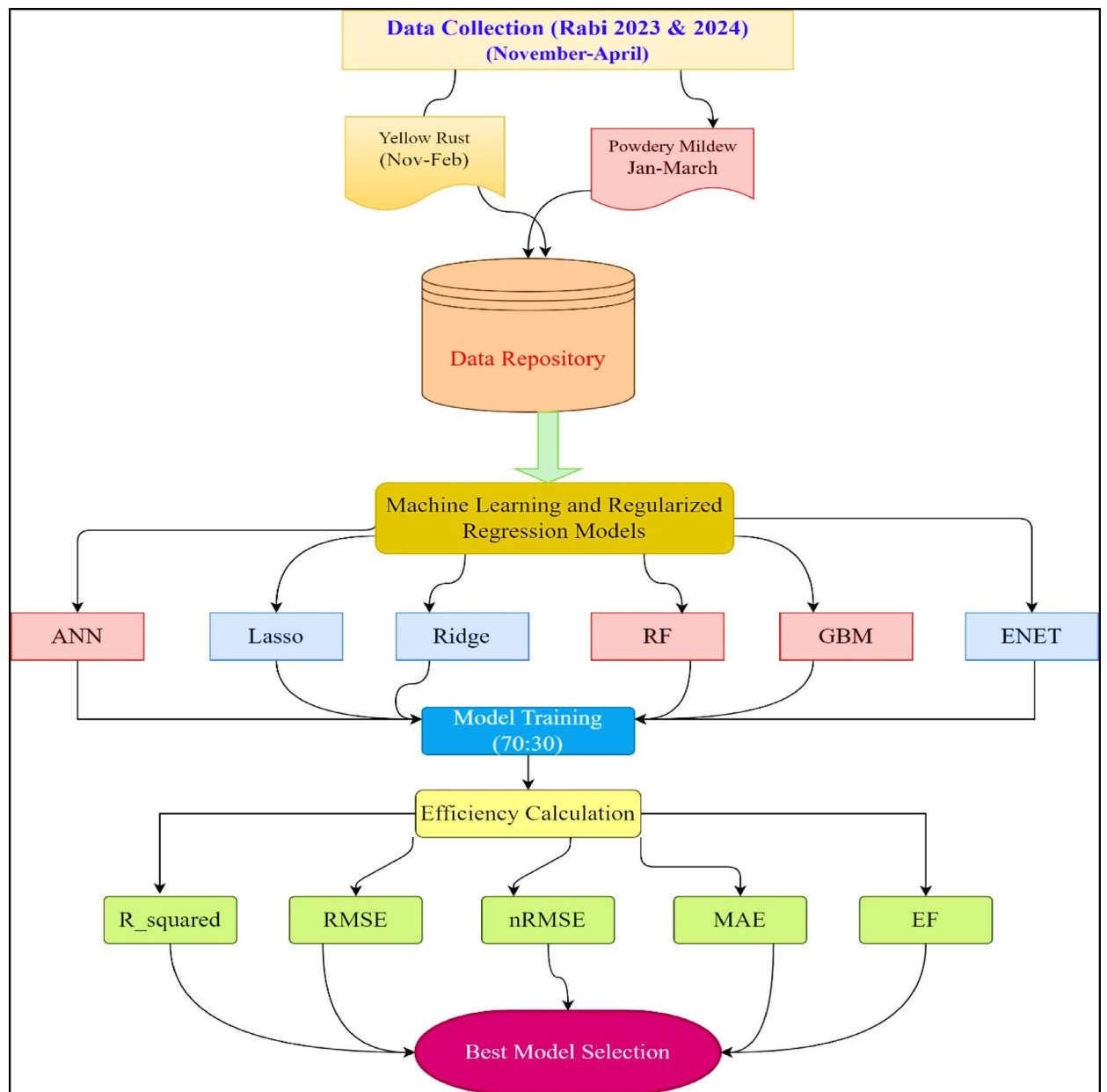


Fig. 3. Flow Chart for yellow rust and powdery mildew disease severity prediction in wheat.

Met. Var.	Mean	Std (±)	Median	Min.	Max.	Skew	Kurtosis	SE(±)
DS	31.27	32.91	15.00	0.00	100.00	0.58	-1.21	3.51
Tmax	20.79	5.04	21.31	10.97	30.00	-0.19	-0.83	0.54
T _{min}	8.16	1.90	8.09	5.06	12.09	0.32	-0.94	0.20
RH _I	91.55	3.01	92.29	85.43	96.71	-0.58	-0.53	0.32
RH _{II}	57.39	17.14	59.29	29.00	86.29	0.01	-1.37	1.83
Rainfall	4.59	12.60	0.00	0.00	60.00	3.85	14.03	1.34
SSH	4.80	2.83	5.53	0.46	9.09	-0.11	-1.51	0.30
WV	3.11	1.32	2.91	1.17	5.75	0.59	-0.67	0.14
Eva	17.98	10.29	16.35	4.60	54.00	1.18	1.64	1.10

Table 2. Meteorological condition during yellow rust progression during both growing season.

Met. Var.	Mean	Std ($\pm$)	Median	Min.	Max.	Skew	Kurtosis	SE($\pm$)
DS	29.35	33.55	12.50	0.00	95.00	0.66	-1.21	3.62
Tmax	26.67	4.58	26.59	12.66	35.24	-0.42	0.25	0.49
T _{min}	11.05	2.79	10.97	4.71	16.19	-0.30	-0.60	0.30
RH _I	87.10	6.06	87.86	70.57	96.14	-1.33	1.48	0.65
RH _{II}	42.23	14.25	40.86	15.29	84.00	0.58	0.21	1.54
Rainfall	9.33	20.07	0.00	0.00	68.80	2.12	2.94	2.16
SSH	7.31	2.09	7.81	0.96	10.31	-0.95	0.24	0.22
WV	3.52	1.09	3.50	1.00	5.31	-0.21	-0.82	0.12
Eva	25.86	6.88	26.70	4.80	37.00	-0.70	0.55	0.74

Table 3. Meteorological condition during powdery mildew progression during both growing season.

Model	Hyperparameters tuned	Values/ranges
ANN	Hidden layer size (size)	1, 3, 5
	Weight decay (decay)	0.1, 0.5, 1
Ridge	Lambda	500 values from 10^5 to 10^{-5}
	Alpha	Fixed at 0
ENET	Lambda	500 values from 10^5 to 10^{-5}
	Alpha	Fixed at 0.5
GBM	Number of trees	1000
	Interaction depth	3
	Learning rate (shrinkage)	0.01
Lasso	Lambda	500 values from 10^5 to 10^{-5}
	Alpha	Fixed at 1
Random Forest	mtry (variables per split)	Default (auto-tuned)
	ntree (number of trees)	Default

Table 4. Hyperparameter tuning for machine learning model.

Ridge regression

Ridge regression⁶⁵ was implemented using the ‘glmnet’ package. The data was split into 70:30 for training And testing. The model used 3-fold cross-validation for ‘lambda’ tuning. Features were ‘centered’ and ‘scaled’. A lambda vector was created using 500 values ranging from 10^5 to 10^{-5} . The alpha parameter was set to 0 for ridge regression. The ‘train’ function from caret was used with the method set to “glmnet”. The model was trained on the training set, using the predefined lambda vector for tuning (Table 4).

Elastic net (ENET)

The Elastic Net model⁶⁶ also used the ‘glmnet’ package. Data splitting And preprocessing were similar to the Ridge model. The key difference was setting alpha to 0.5, balancing between ridge and lasso penalties. The same lambda vector (500 values from 10^5 to 10^{-5}) was used for tuning. The model was trained using caret’s ‘train’ function with method “glmnet”, incorporating both L1 and L2 regularization (Table 4).

Gradient boosting machine (GBM)

The GBM model⁶⁷ was developed using the ‘gbm’ package. After the 70 – 30 data split, features were ‘centered’ and ‘scaled’. The model was configured with 1000 decision trees, An interaction depth of 3, and a learning rate (shrinkage) of 0.01. Five-fold cross-validation was used within the GBM training process. The distribution was set to “gaussian” for regression tasks. The model was trained directly using the ‘gbm’ function, allowing for parallel processing across available cores (Table 4).

Lasso regression

Lasso regression⁶⁸ was implemented similarly to Ridge regression, using the ‘glmnet’ package. The key difference was setting the alpha parameter to 1, which applies L1 regularization. Data splitting, preprocessing, and lambda vector creation were identical to the Ridge model. The ‘train’ function from caret was used with method “glmnet”, tuning over the same lambda vector but with alpha fixed at 1 (Table 4).

Random forest

The Random Forest model was developed using the ‘randomForest’ package^{69,70}. After the 70 – 30 data split, 3-fold cross-validation was used for model evaluation. Unlike the other models, explicit feature scaling wasn’t necessary due to the nature of random forests. The model was trained using caret’s ‘train’ function with method

“rf”. Default parameters were used for the number of trees and variables considered at each split, allowing the algorithm to optimize these based on the data (Table 4).

Principle component analysis

The Principal Component Analysis was conducted in this study to reduce the dataset’s dimensionality and explore the underlying patterns and relationships between wheat disease severity variables. By reducing the number of variables, PCA simplifies the interpretation of the data and helps identify the most important features driving the variation. Scaling was applied to the data before performing PCA to ensure that all variables were on a similar scale and to prevent variables with larger magnitudes from dominating the analysis. This study used standard scaling (z-score normalization), which subtracts the mean And divides by the standard deviation of each variable. This standardization ensures that all variables have a mean of 0 And a standard deviation of 1, making them directly comparable regarding their influence on the principal components.

The biplot was created, which visually represents the relationships between the independent variables and disease severity. The biplot was customized to include ellipses for each disease type, making it easier to visualize the separation between rust and powdery mildew.

Evaluation matrix

The performances of the different models were evaluated using several statistical criteria: coefficient of determination (R^2), Root Mean Squared Error (RMSE), normalized root mean square (nRMSE), mean absolute error (MAE) and model efficiency (EF). Values of evaluation metrics corresponding to a given variable for yellow rust and powdery mildew were computed in R. R^2 values nearer to 1 and RMSE values close to 0 indicate better model performance. The developed model is usually considered excellent, good, fair or poor if the nRMSE value lies in the range of <10%, 10–20%, 20–30% or >30%, respectively⁷¹.

$$R^2 = 1 - \frac{\sum (-\hat{y})^2}{\sum (y_i - \bar{y})^2} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2} \tag{2}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}| \tag{3}$$

$$nRMSE = \frac{RMSE}{y_{max} - y_{min}} \times 100 \tag{4}$$

$$EF = 1 - \frac{\text{Residual Sum of Squares}}{\text{Total Sum of Squares}} \tag{5}$$

where,
 y_i = predicted value of y , $\hat{y}$ = mean value of y , n = number of predictors, N = total sample size.

Result
Yellow rust prediction

The performance metrics of six different ML models: ANN, LASSO, Ridge regression, ENET, GBR and RF, to evaluate the impact of meteorological variables on disease severity of yellow rust in wheat are presented in Table 5. The metrics used for assessment include Training RMSE, Training Mean Absolute Error (MAE), Training R-squared (R^2 and Training Efficiency (EF), Testing RMSE, Testing MAE, Testing R-squared, and Testing EF.

The results showed that R^2 between weather parameters And disease severity are strongly correlated with values of 0.96 And 0.93, respectively, obtained in calibration And validation sets. Moreover, the lowest RMSE, 6.38 And 10.96, was obtained during calibration And validation, respectively. The MAE value of 4.77 for calibration And 6.92 for validation envisaged that ANN showed robustness in disease prediction accuracy. Similarly, the RF model also demonstrates excellent performance with R^2 values of 0.96 for calibration And 0.92 for validation, and RMSE (5.91 for calibration And 10.93 for validation) and MAE (4.48 for calibration And

Model	Calibration					Validation				
	R^2	RMSE	nRMSE	MAE	EF	R^2	RMSE	nRMSE	MAE	EF
ANN	0.96	6.38	6.72	4.77	0.96	0.93	10.97	10.97	6.92	0.88
ENET	0.80	14.72	15.49	12.21	0.80	0.70	18.17	18.17	15.61	0.68
GBM	0.95	0.23	8.01	0.18	0.95	0.89	0.38	12.79	0.26	0.84
Lasso	0.80	14.68	15.45	12.11	0.80	0.68	18.40	18.40	15.68	0.67
RF	0.97	5.91	6.22	4.48	0.97	0.93	10.93	10.93	7.58	0.88
Ridge	0.80	14.99	15.78	12.46	0.79	0.73	17.50	17.50	15.32	0.70

Table 5. Evaluation matrix of disease severity of yellow rust using machine learning models.

7.58 for validation), confirming the RF reliability in disease prediction using climatic variables. Moreover, the Elastic Net, Lasso and Ridge models exhibited moderate to lower predictive performance with lower R^2 values and higher RMSE and MAE, revealing lesser accuracy among the models. The GBM model revealed a high R^2 of 0.94, which showed overfitting potential, but the least R^2 of 0.88. The EF values further validate these findings, highlighting that AnN And RF are efficient models with EF values of 0.96 And 0.97 for calibration And 0.88 And 0.88 for validation, respectively. The ANN and RF models are the most effective and accurate for predicting yellow rust in wheat, while ANN slightly outperforms RF in overall evaluation metrics.

The feature importance for predicting yellow rust in wheat using various machine learning models is presented in Fig. 4. For the Ridge model, evapotranspiration and wind velocity are identified as the most significant predictors, followed by rainfall and relative humidity. The Random Forest model highlights evapotranspiration as the dominant predictor with a substantial margin, followed by Tmax and SSH. The Lasso model also underscores the importance of evapotranspiration and wind velocity, with RHII and rainfall being moderately important. The Elastic Net model aligns closely with Lasso, strongly emphasising evapotranspiration, followed by wind velocity and RHII. In the ANN model, evapotranspiration and wind velocity are most important, with minimum temperature, SSH, and rainfall contributing significantly. Finally, the GBM model, represented by relative influence (rel.inf), places the highest importance on evapotranspiration, followed by Tmax and Tmin, indicating a similar pattern of critical features. Across all models, evapotranspiration consistently emerges as the most critical feature, highlighting its pivotal role in predicting yellow rust in wheat. Wind velocity and temperature variables (Tmax and Tmin) also appear prominently across different models, suggesting their relevance in predicting disease. This comprehensive analysis of feature importance across diverse models underscores the critical environmental variables that must be monitored for effective yellow rust prediction in wheat.

3.2 Powdery mildew prediction

Table 6 presents the performance of various machine learning models in predicting powdery mildew of wheat. The AnN model demonstrated exceptional performance, with An R-squared value of 0.98 And 0.95 for the calibration And validation sets. It exhibited relatively low RMSE values of 4.98 And 6.98, along with nRMSE values of 5.24 And 7.35 And MAE values of 3.32 And 4.98 for the calibration And validation sets, respectively. The EF values were also high at 0.98 And 0.95, indicating excellent model efficiency. The GBM model also showed promising results, with An R-squared of 0.98 And 0.90, RMSE of 0.15 And 0.30, nRMSE of 5.28 And 10.86, MAE of 0.10 And 0.22, And EF of 0.98 And 0.89 for the calibration And validation sets, respectively. The RF model performed similarly well, with An R-squared of 0.98 And 0.90, RMSE of 4.52 And 10.60, nRMSE of 4.75 And 11.16, MAE of 2.82 And 6.85, And EF of 0.98 And 0.89 for the calibration and validation sets, respectively.

On the other hand, the Elastic Net, Lasso, And Ridge Regression models exhibited relatively lower performance. The Elastic Net model had An R-squared of 0.77 And 0.95, RMSE of 16.35 And 6.98, nRMSE

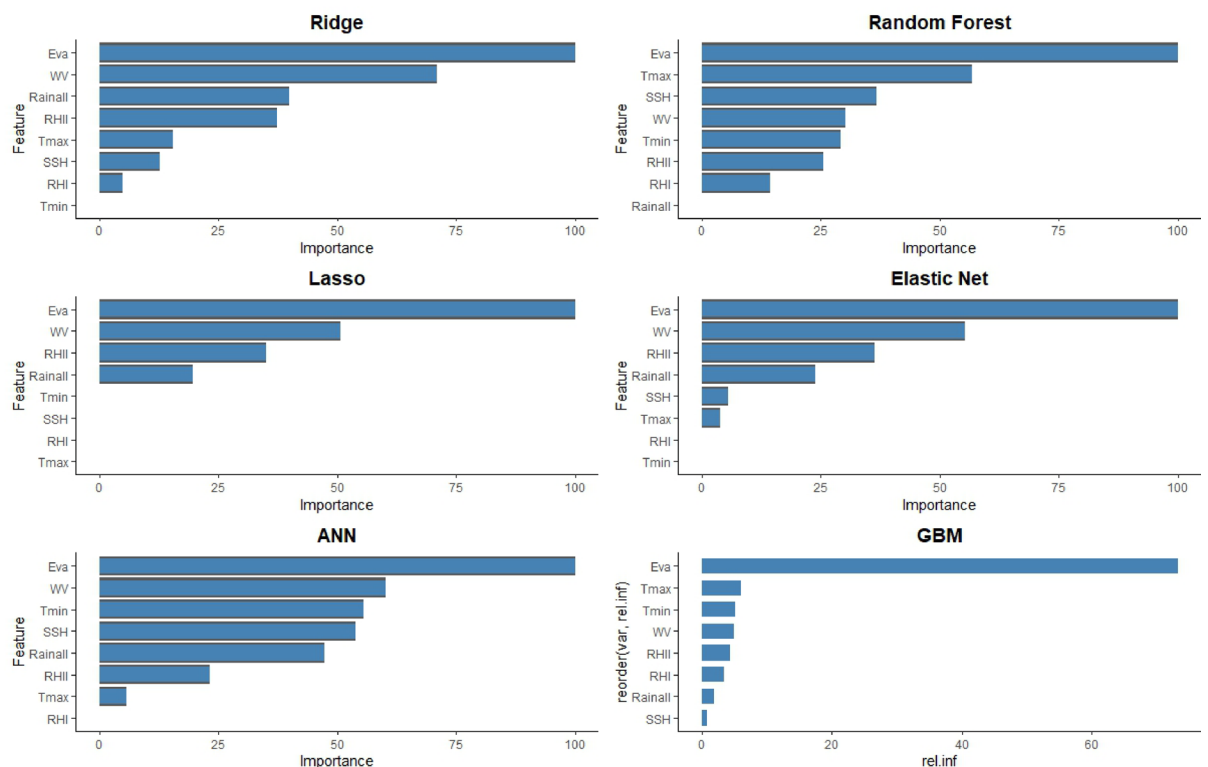


Fig. 4. Variable importance unit for different models for yellow rust severity assessment.

Model	Calibration					Validation				
	R^2	RMSE	nRMSE	MAE	EF	R^2	RMSE	nRMSE	MAE	EF
ANN	0.98	4.98	5.24	3.32	0.98	0.95	6.98	7.35	4.98	0.95
ENET	0.77	16.35	17.21	13.46	0.77	0.95	6.98	7.35	4.98	0.95
GBM	0.98	0.15	5.28	0.10	0.98	0.90	0.30	10.86	0.22	0.89
Lasso	0.77	16.38	17.24	13.54	0.77	0.74	16.36	17.22	14.55	0.73
RF	0.98	4.52	4.75	2.82	0.98	0.90	10.60	11.16	6.85	0.89
Ridge	0.74	17.41	18.32	14.66	0.74	0.73	16.69	17.57	14.24	0.72

Table 6. Evaluation matrix for disease severity of powdery mildew for machine learning models.

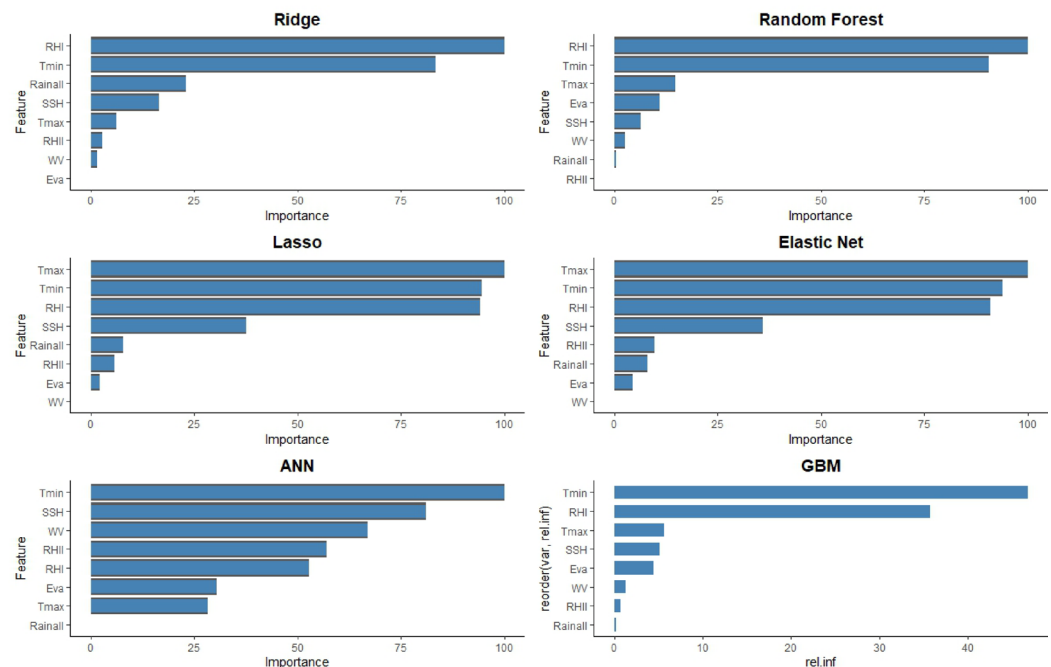


Fig. 5. Variable importance unit for different models for Powdery Mildew severity assessment.

of 17.21 And 7.35, MAE of 13.46 And 4.98, And EF of 0.77 And 0.95 for the calibration And validation sets, respectively. The Lasso model showed consistent performance with An R-squared of 0.77 And 0.74, RMSE of 16.38 And 16.36, nRMSE of 17.24 And 17.22, MAE of 13.54 And 14.55, And EF of 0.77 And 0.73 for the calibration And validation sets, respectively. The Ridge Regression model had An R-squared of 0.74 And 0.73, RMSE of 17.41 And 16.69, nRMSE of 18.32 And 17.57, MAE of 14.66 And 14.24, And EF of 0.74 And 0.72 for the calibration and validation sets, respectively.

The variable importance plot (Fig. 6) provide valuable insights into the relative importance of various meteorological factors in modeling powdery mildew in wheat, allowing for better understanding and potential improvements in the prediction models. For the Ridge Regression, the most important variables are relative humidity, Tmin, rainfall, SSH, and Tmax. The Random Forest model assigns high importance to Tmin, Tmax, relative humidity, Eva, and SSH. The Lasso Regression model highlights Tmax, Tmin, relative humidity, SSH, and rainfall as the top influential variables. The Elastic Net model also identifies Tmin, relative humidity, SSH, rainfall, and Eva as the most critical factors.

The ANN model reveals that Tmin, SSH, WV, relative humidity, rainfall, Eva, and Tmax are the most significant variables for predicting powdery mildew. The variable importance plot indicates that the GBM model assigns importance to Tmin, relative humidity, Tmax, SSH, and Eva as crucial predictors of powdery mildew based on the collective importance assigned by multiple machine learning models, the key factors that appear to be most influential for predicting powdery mildew in wheat are Tmin, relative humidity, rainfall and sunshine hours.

The density distributions of predicted values for yellow rust and powdery mildew disease severity in wheat, obtained from various machine learning models are presented in Fig. 6. The density distribution plot for yellow rust showed multiple peaks, signifying that the models predict varying levels of disease severity. The ANN and Ridge models exhibited distinct dense peaks at certain severity values. In contrast, other models such as Elastic Net, GBM, Lasso, and RF showed dispersed distributions with several local peaks. The powdery mildew density

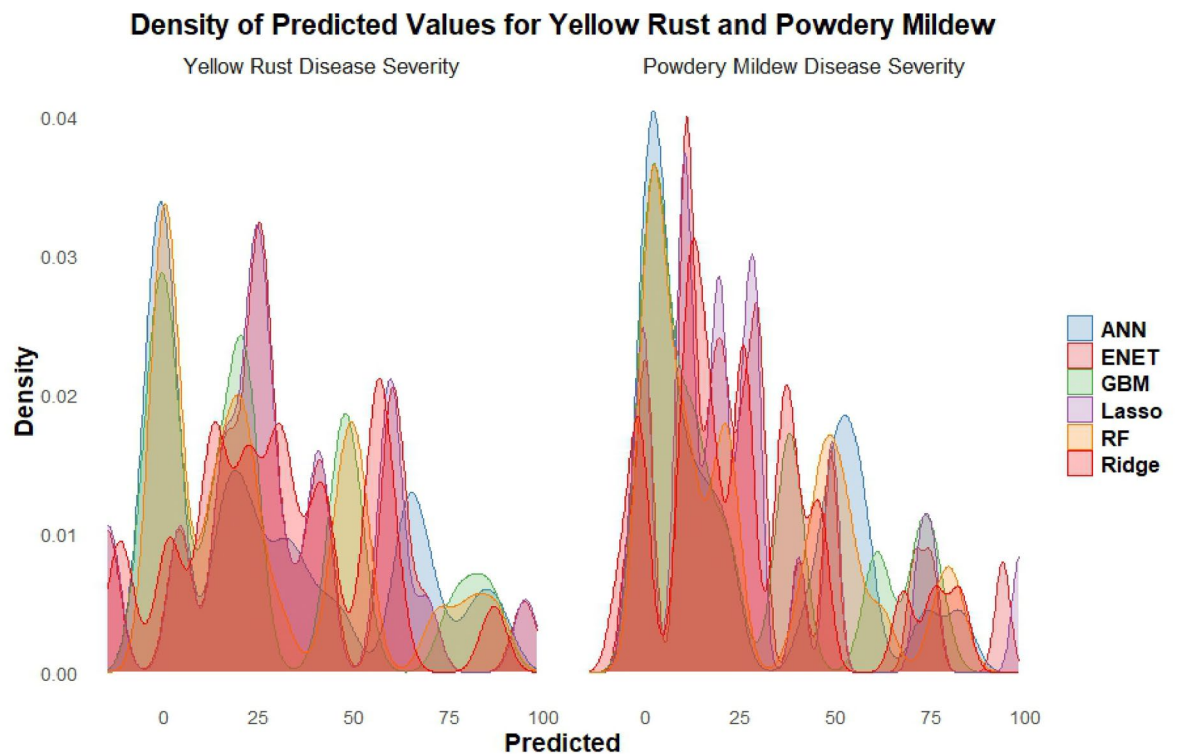


Fig. 6. Density distributions of predicted values for two different diseases affecting wheat yellow rust and powdery mildew disease.

distribution pattern is concentrated with higher peaks, indicating that the models predict specific severity levels more frequently. The ANN and Ridge models revealed sharper peaks, indicating more consistent predictions, while the other models show slightly broader distributions with multiple local maxima. The multi-modal distributions indicate that most models predict a range of severity levels, with some models tending to predict specific severity levels more frequently than others.

The figure (Fig. 7) presents a scatter plot comparison between the observed and predicted values of disease severity for yellow rust and powdery mildew in wheat obtained from various machine learning models. The scatter points align closely along the diagonal line for yellow rust disease severity, indicating a good agreement between the observed and predicted values across the different models. The points are evenly distributed, suggesting that the models can accurately predict disease severity levels across the range of observed values. However, there are a few outliers where the predicted values deviate from the observed ones.

In the case of powdery mildew disease severity, the scatter points also follow the diagonal line reasonably well, implying that the models can capture the observed disease severity levels to a reasonable extent. Nevertheless, there appears to be slightly more deviation from the diagonal, particularly at higher severity levels, where some models tend to overpredict or underpredict the observed values. Additionally, there are more scattered outliers than the yellow rust plot, indicating instances where the models struggle to predict the severity of powdery mildew accurately.

Overall, the scatter plots provide a visual assessment of the model performance, allowing for the identification of potential biases or discrepancies between the observed and predicted values. While both diseases exhibit satisfactory agreement, the powdery mildew plot suggests slightly more challenges in predicting disease severity accurately, especially at higher severity levels, compared to yellow rust.

Regression models

The regression equations (Table 7) provide insights into the relationships between meteorological factors and the severity of yellow rust disease, enabling better understanding and potential improvements in disease prediction and management strategies.

The regression equation obtained from ENET suggests that Eva, WV, rainfall, and sunshine hours positively influence the severity of yellow rust. At the same time, it is negatively affected by maximum relative humidity (RHII). Like the Elastic Net equation, the Lasso regression equation shows a positive relationship between disease severity and Eva, WV, and rainfall, while relative humidity has a negative impact. The Ridge regression equation also exhibits positive coefficients for Eva, WV, rainfall, and Tmax, suggesting that higher values of these variables are associated with increased yellow rust severity.

The regression equations obtained from different machine learning models for predicting the severity of powdery mildew disease in wheat are presented in Table 8. The regression equation obtained for ENET suggests that the severity of powdery mildew is negatively influenced by maximum temperature and relative humidity

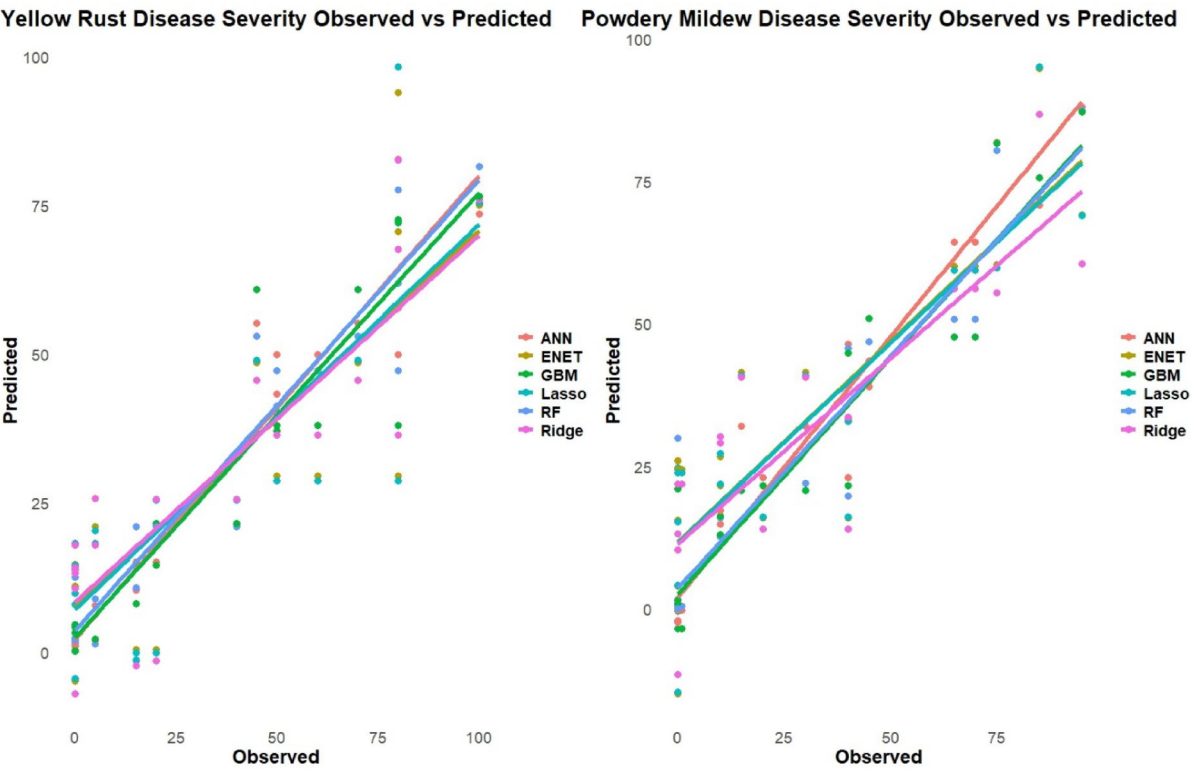


Fig. 7. Comparison between the observed and predicted values for the severity of two crop diseases: Yellow Rust and Powdery Mildew.

Model	Regression equation
Elastic Net	$DS = 31.937 + 15.215 \times Eva + 8.423 \times WV - 5.534 \times RHII + 3.645 \times Rainfall + 0.829 \times SSH$
Lasso	$DS = 31.937 + 16.681 \times Eva + 8.471 \times WV - 5.849 \times RHII + 3.3 \times Rainfall + 0 \times Tmax$
Ridge	$DS = 31.937 + 10.999 \times Eva + 8.071 \times WV + 4.964 \times Rainfall - 4.711 \times RHII + 2.495 \times Tmax$

Table 7. Regression equations obtained for disease severity prediction using different machine learning approaches for yellow rust.

Model	Regression equation
Elastic Net	$DS = 29.726 - 27.312 \times Tmax + 25.748 \times Tmin - 24.939 \times RHI + 10.553 \times SSH - 3.673 \times RHII$
Lasso	$DS = 29.726 - 26.375 \times Tmax + 25.032 \times Tmin - 24.927 \times RHI + 10.582 \times SSH + 3.075 \times Rainfall$
Ridge	$DS = 29.726 - 16.601 \times RHI + 13.925 \times Tmin + 4.192 \times Rainfall + 3.114 \times SSH - 1.46 \times Tmax$

Table 8. Regression equations obtained for disease severity prediction using different machine learning approaches for powdery mildew.

(RHI and RHII). At the same time, it is positively related to Tmin and sunshine hours. Like the Elastic Net equation, the Lasso regression equation shows a negative relationship between disease severity, Tmax, and relative humidity. At the same time, Tmin and sunshine hours have a positive impact.

Additionally, the Lasso model includes rainfall as a significant predictor, with a positive coefficient indicating that higher rainfall is associated with increased powdery mildew severity. The Ridge regression equation also exhibits a negative coefficient for maximum relative humidity (RHI) and Tmax. On the other hand, Tmin, rainfall, and sunshine hours have positive coefficients.

Based on the regression equations and variable importance plots, some notable differences exist in the influential variables for predicting yellow rust and powdery mildew severity in wheat. Eva and WV emerge as significant positive predictors across the Elastic Net, Lasso, and Ridge regression equations, and rainfall also has a positive coefficient. On the other hand, temperature variables (Tmax and Tmin) play a crucial role in powdery mildew prediction. This suggests that higher maximum temperatures may suppress powdery mildew

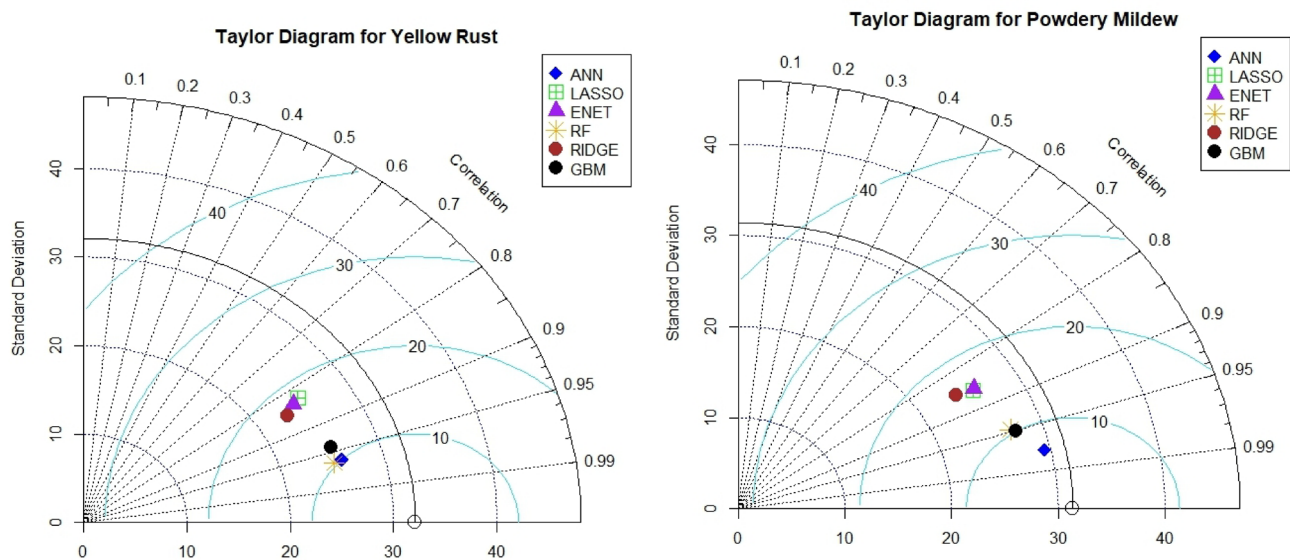


Fig. 8. Taylor diagram for performance of different machine learning models for predicting Yellow Rust and Powdery Mildew in Wheat.

Variables	PC1	PC2	PC3
Maximum temperature	−0.439	0.103	0.100
Minimum temperature	−0.332	−0.314	0.402
Relative Humidity (I)	0.340	−0.338	0.151
Relative Humidity (II)	0.413	−0.263	0.026
Rainfall	−0.134	−0.589	−0.744
Sunshine Hours	−0.426	0.167	−0.109
Wind Speed	−0.219	−0.570	0.456
Evaporation	−0.400	−0.080	−0.171
Eigenvalue	4.75	1.62	0.59
Cumulative (%)	59.43	79.77	87.18
Standard deviation	2.18	1.27	0.77

Table 9. Component matrix for the first three principal components of yellow rust.

development, while higher minimum temperatures could promote disease severity. Additionally, the plots of variable importance further highlight the differences in influential factors for the two diseases. Variables like Eva, WV, and rainfall appear to be more important for yellow rust. In contrast, for powdery mildew, temperature variables (minimum and maximum temperatures), relative humidity, and sunshine hours are among the top influential factors. These differences in the significant predictors and their relationships with disease severity suggest that the environmental conditions favouring yellow rust and powdery mildew development and spread may differ. This insight can be valuable for targeted disease management strategies and understanding each disease’s unique environmental requirements.

Taylor diagram obtained for yellow rust predictions, the ANN and Lasso models demonstrate the highest correlation with the observed data, followed by the ENET, RF, Ridge, and GBM models. The ANN model again exhibits the lowest centered root mean square error (CRMSE), suggesting superior performance in predicting yellow rust severity compared to the other models. In the case of powdery mildew predictions, the ANN and RF models show the highest correlation with the observed data, followed by the Lasso, ENET, Ridge, and GBM models. The ANN model exhibits the lowest CRMSE, indicating the best overall performance among the evaluated models (Fig. 8).

Principal component analysis

PCA for yellow rust

Factor analysis was conducted to cluster the originally recognized eight weather parameters into statistical factors based on their correlation patterns. Using the Cumulative Variance Explained approach, three principal components (PCs) with eigenvalues greater than 0.5 were selected (Table 9), collectively accounting for 87.18% of the cumulative variance. The first principal component (PC1) has a significant contribution from several variables, with Tmax (−0.4395), SSH (−0.4262), RHII (0.4137), and Eva (−0.4009) showing the highest loadings, indicating

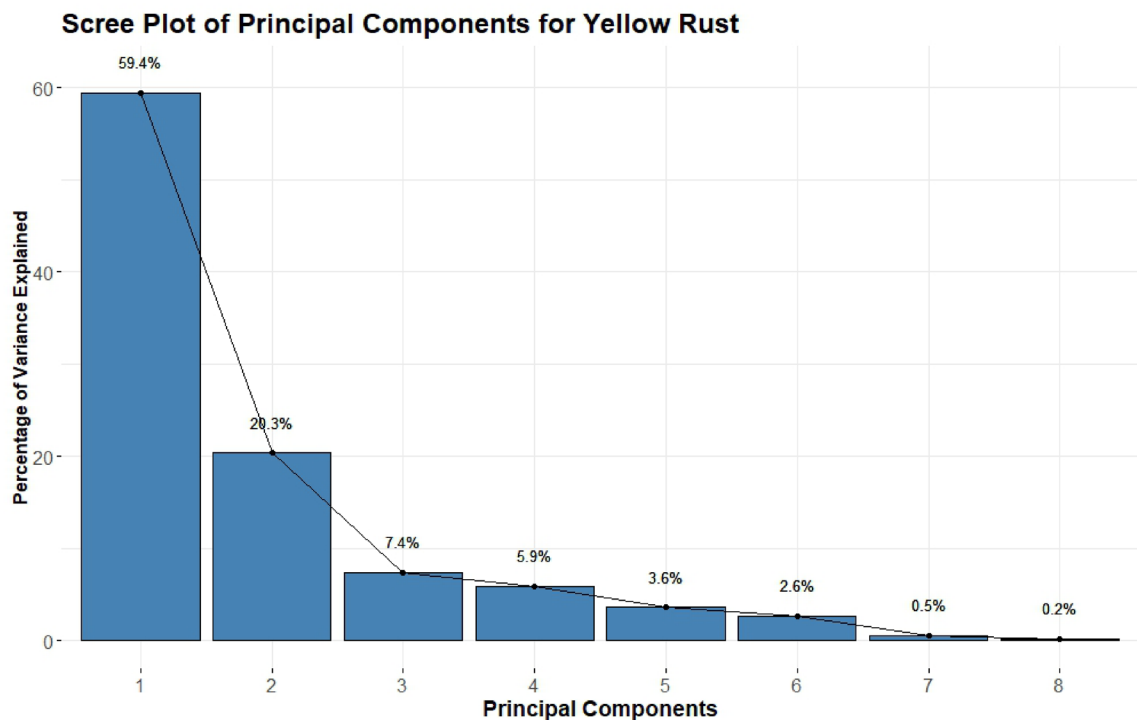


Fig. 9. Scree plot of principal component analysis for yellow rust.

Variables	PC1	PC2	PC3
Maximum Temperature	0.452	−0.034	0.092
Minimum Temperature	0.355	0.228	0.477
Relative Humidity (I)	−0.407	0.115	−0.206
Relative Humidity (II)	−0.409	0.184	0.306
Rainfall	−0.023	0.691	0.325
Sunshine Hours	0.405	−0.060	−0.359
Wind Speed	−0.003	0.599	−0.624
Evaporation	0.411	0.243	−0.044
Eigenvalue	4.65	1.56	0.84
Cumulative (%)	58.22	77.77	88.36
Standard deviation	2.15	1.25	0.92

Table 10. Component matrix for the first three principal components of powdery mildew.

these variables strongly influence PC1. This component explains the largest proportion of variance, 59.43%, with An eigenvalue of approximately 4.75, suggesting it captures a substantial portion of the data’s variability. The second principal component (PC2) has significant loadings from Rainfall (−0.5893), WV (−0.5702), and Tmin (−0.3141), explaining 20.35% of the variance with An eigenvalue of around 1.63. This suggests that these variables are primarily associated with PC2, indicating a distinct underlying pattern in the weather data. The third principal component (PC3) has the highest contributions from Rainfall (−0.7448), WV (0.4560), and Tmin (0.4025), explaining 7.41% of the variance with An eigenvalue of approximately 0.59. This component captures additional variability not explained by the first two components. These three principal components collectively explain 87.18% of the total variance in the weather data, with each component highlighting specific weather parameters that contribute most significantly to the observed patterns. Temperature-related variables and sunshine hours dominate PC1, PC2 is influenced mainly by rainfall and wind velocity, and PC3 highlights the variability due to rainfall, wind velocity, and Tmin (Fig. 9).

PCA for powdery mildew

The Factor analysis was conducted to cluster the initially recognized eight weather parameters into statistical factors based on their correlation patterns. Using the Cumulative Variance Explained approach, three principal components (PCs) with eigenvalues greater than 0.5 were selected (Table 10), collectively accounting for 88.36%

of the cumulative variance. The first principal component (PC1) has significant contributions from Tmax (0.4528), SSH (0.4056), Eva (0.4120), and RHII (−0.4097), indicating these variables strongly influence PC1. This component explains the largest proportion of variance, 58.22%, with An eigenvalue of approximately 4.66, suggesting it captures a substantial portion of the data's variability. The second principal component (PC2) has significant loadings from Rainfall (0.6912), WV (0.5991), and Tmin (0.2284), explaining 19.55% of the variance with An eigenvalue of around 1.56. This suggests that these variables are primarily associated with PC2, indicating a distinct underlying pattern in the weather data. The third principal component (PC3) has the highest contributions from Tmin (0.4773), Rainfall (0.3258), and RHII (0.3063), explaining 10.59% of the variance with An eigenvalue of approximately 0.85. This component captures additional variability not explained by the first two components. These three principal components collectively explain 88.36% of the total variance in the weather data, with each component highlighting specific weather parameters that contribute most significantly to the observed patterns. Temperature-related variables and sunshine hours dominate PC1, PC2 is influenced mainly by rainfall and wind velocity, and PC3 highlights the variability due to minimum temperature, rainfall, and relative humidity (Fig. 10).

The PCA biplot for yellow rust disease of wheat reveals that relative humidity indices (RHI and RHII) and rainfall are closely associated with the disease incidence, indicating a positive correlation. Conversely, factors like Tmin, SSH, and Eva negatively correlate.

In the case of powdery mildew disease, the biplot shows that sunshine hours (SSH), Tmax, and evaporation are positively correlated with the disease incidence. These environmental factors likely create conditions conducive to the growth and proliferation of the powdery mildew pathogen. On the other hand, relative humidity indices (RHI and RHII) and wind velocity exhibit a negative correlation (Fig. 11).

Discussion

Region-wise prediction of rust and powdery mildew based on meteorological data is crucial for effective disease management and optimizing crop yields^{72–74}. Specific weather conditions influence these fungal diseases; meteorological data is crucial for forecasting outbreaks⁵⁵. The study results indicate rust and powdery mildew infestations increase with average temperature and humidity. Similarly, previous studies also reported that powdery mildew flourishes in high humidity and moderate temperatures (15–20 °C), while rust is favoured by warm temperatures (20–25 °C) and prolonged leaf wetness^{75,76}. A strong correlation between disease incidence and severity of winter wheat rust with weather data ($r = 0.89$ and $r = 0.87$, respectively) was reported using CART or decision tree analysis⁷⁷. However, in our study, the coefficient of determination between weather parameters And disease severity strongly correlates with values of 0.96 And 0.93 during calibration and validation, respectively. Hence, it is best to describe the relationship between weather variables and winter wheat rust epidemics with some differences in weather variables^{32,78}. Stepwise regression analysis for wheat rust showed that maximum relative humidity positively correlated with rust severity, while solar radiation and minimum relative humidity

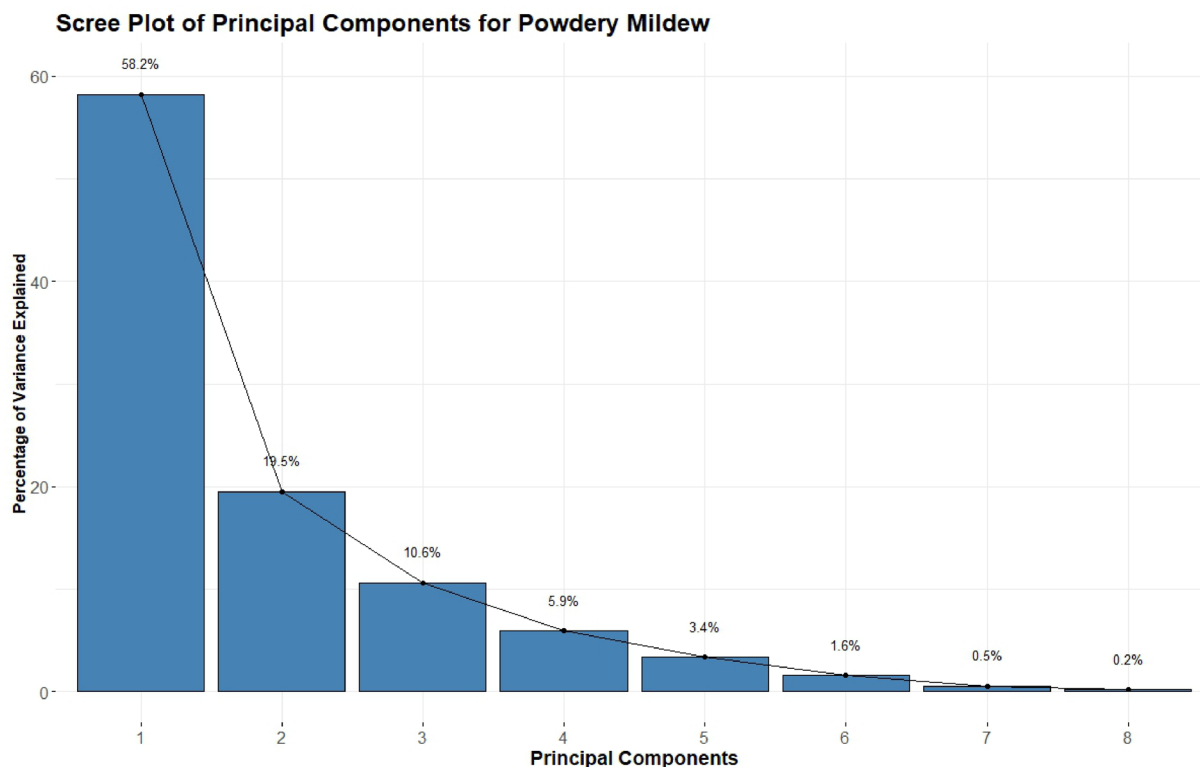


Fig. 10. Scree Plot of Principal Component Analysis for Powdery Mildew.

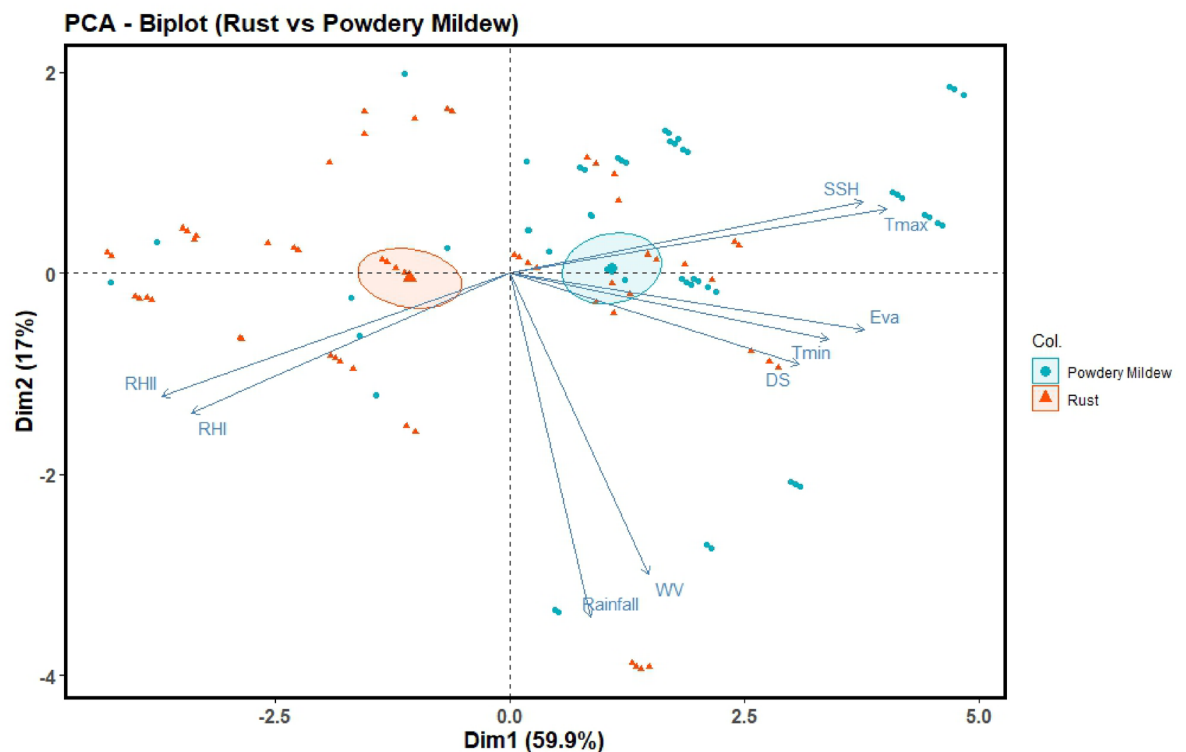


Fig. 11. Biplot of principal component analysis for Yellow Rust and Powdery Mildew.

were negatively correlated in disease prediction⁷⁹. Temperature-related variables are the most crucial for explaining and forecasting the occurrence of leaf rust, whereas a humid and warm environment better explains stripe rust. However, evapotranspiration is primarily a function of maximum air temperature^{80,81} whereas wind velocity and temperature also appear as important variables across different models⁸². Similarly, our results align with these findings as the most critical feature across all machine learning models.

Linear regression analysis relating leaf rust severity to maximum air temperature and humidity thermal ratio showed r^2 values of 0.68 And 0.71, respectively⁸³. The relationship between wheat leaf rust and weather variables was analysed by⁸⁴ using the stochastic approach, including the Dennis model and Monte Carlo simulation. This approach emphasized the study of the combined effects of temperature, relative humidity, and rainfall on the infection of wheat by *P. striiformis* during continuous Hours over the critical period for the development of stripe rust incidence And severity in wheat. The results indicated optimal conditions for developing wheat stripe rust included a combination of relative humidity greater than 92%, temperatures between 4 °C And 16 °C for at least 4 continuous Hours, And rainfall of 0.1 mm or less. Shrinkage regression and machine-learning approaches envisaged that the ANN model is outperformed in calibration and validation to derive models for predicting epidemiological characteristics of white rust in mustard crops⁸⁵. A decision tree classifier is utilized to predict stripe rust in wheat using limited weather variables And achieved a training And test accuracy of 75.70%. However, stem rust developed a prediction accuracy of 90.05% using 10-fold cross-validation⁸⁶ In the present study, a six linear regression model showed the highest accuracy of $r^2 = 90.60$, revealing that WV and relative humidity are important variables in predicting disease⁸⁷. Similarly, features importance by our best-performing model ANN gives a high importance/weightage to WV for predicting yellow rust severity. In contrast, regression analysis by Lasso and Ridge shows that minimum relative humidity is negatively aligned with the prediction of yellow rust.

An empirical model study⁸⁷ shows that wind is the primary weather factor influencing the powdery mildew epidemic, followed by sunshine and radiation. However, temperature is the main weather factor influencing disease severity⁸⁸. Clam weather conditions create a humid environment that favors the development of powdery mildew severity^{89,90}. The effect of wind on humidity is stronger than the effect of wind on dispersal. Low wind runs may still be sufficient to move the initial inoculum between fields. Our best-performing model, ANN also shows the importance of minimum temperature and sun shine hours for prediction of powdery mildew.

The stepwise selection method in logistic regression revealed that weekly mean temperature, relative humidity and sunshine are the key predictor variables for powdery mildew assessment at the high risk⁹¹. Odds ratio point estimates indicate that increasing weekly temperature (from 11.5 to 21.9 °C) and decreasing weekly relative humidity (from 65.6 to 41.8%) and sunshine hours (from 8.7 to 5.4 h) were key drivers for high powdery mildew severity. Cumulative logit models relating the development of wheat powdery mildew to six weather variables produced R^2 of 84.1%. Rainfall and vapor pressure deficit were statistically significantly related to the disease score. The best model with temperature and humidity-related produced a generalized R^2 of 72.4%.

Greater average wind speed positively impacts the mildew epidemic^{29,92}. Hence, the positive effect of wind on spore dispersal outweighs its negative effect on reducing humidity^{90,93}. However, precipitation has a negative effect on spore survival due to the leaching of spores into deep soil.

Accurate disease prediction aids in implementing integrated pest management (IPM) strategies, combining cultural practices, resistant varieties, and targeted chemical applications^{94,95}. Early warnings of disease incidence allow for proactive measures, such as adjusting irrigation schedules to reduce leaf wetness or optimizing planting dates to avoid peak disease periods^{96,97}. Leveraging meteorological data for predicting wheat powdery mildew and rust is vital for enhancing disease management strategies^{98,99}. It empowers farmers to make informed decisions, safeguarding crop health, ensuring food security, and promoting sustainable agricultural practices.

PCA conducted by³² to associate wheat leaf rust with weather And planting date revealed a tripartite structure in the leaf rust progress curves, collectively explaining 73% of the total variance. The first principal component (PC1) accounted for 33.8% of the variability, PC2, explaining 21.6% of the total variance, was predominantly associated with the mean monthly temperature And the frequency of days characterized by minimum temperatures ranging from 5 °C to 25 °C, coupled with maximum relative humidity exceeding 60%. The third principal component (PC3) elucidated 17.9% of the total variance and demonstrated a moderate positive relation for the mean monthly temperature.

Conclusion

The integration of ANN and regularized regression models with meteorological data demonstrates a powerful approach to predicting the severity of wheat diseases. The ANN model emerged as the most effective. Results indicate superior predictive accuracy and robustness, with low RMSE and MAE values for both diseases. The RF model also showed strong predictive performance, although slightly less accurate than the ANN model. The principal component analysis (PCA) identified evapotranspiration, temperature, and maximum relative humidity as the most critical variables influencing disease incidence, accounting for a significant portion of the variance in the data (87.18% for yellow rust And 88.36% for powdery mildew). The findings emphasize the importance of monitoring key meteorological variables, particularly evapotranspiration, temperature, and humidity, for effective disease prediction and management in wheat. The ANN and RF models provide valuable tools for farmers and researchers, enabling more accurate and timely predictions of disease outbreaks, which can inform proactive management strategies to mitigate the impact of these diseases on wheat production. Future research could explore integrating additional environmental factors and applying these models to other crops and regions to further enhance their predictive capabilities and generalizability. One limitation of our current methodology is the reliance on visual assessment of disease severity, which may not precisely capture the subtle or early-stage changes in disease progression and may introduce observer bias. Furthermore, while our models incorporate real-time weather data, they may not fully account for the latent period between infection and visible symptom development, potentially weakening the temporal alignment between weather events and disease severity. In future studies, integrating lagged weather variables, spore dispersal data, and high-resolution remote sensing or hyperspectral imaging for disease detection could enhance the temporal sensitivity and predictive power of disease models. Moreover, integrating mechanistic disease models with machine learning models that simulate the biological development of the pathogen in response to cumulative weather thresholds could offer a more robust framework for predicting disease outbreaks and severity.

Data availability

Data will be made available on request from the lead author Amit Bijlwan (amitbijlwan07@gmail.com).

Received: 15 February 2025; Accepted: 2 September 2025

Published online: 06 October 2025

References

1. Giraldo, P., Benavente, E., Manzano-Agugliaro, F. & Gimenez, E. Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy* **9**, 352 (2019).
2. OECD-FAO. *Agricultural Outlook 2022–2031*. (2022).
3. Zhang, D. et al. Integration of spectroscopy and image for identifying fusarium damage in wheat kernels. *Spectrochim Acta Mol. Biomol. Spectrosc.* **236**, 118344 (2020).
4. Randhawa, M. S., Bhavani, S., Singh, P. K., Huerta-Espino, J. & Singh, R. P. Disease Resistance in Wheat: Present Status and Future Prospects. In *Disease Resistance in Crop Plants*. pp. 61–81, Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-20728-1_4
5. Liu, T., Wan, A., Liu, D. & Chen, X. Changes of races and virulence genes in *Puccinia striiformis* f. sp. *tritici*, the wheat Stripe rust pathogen, in the united States from 1968 to 2009. *Plant. Dis.* **101**, 1522–1532 (2017).
6. Cat, A., Tekin, M., Akan, K., Akar, T. & Catal, M. Virulence characterization of the wheat Stripe rust pathogen, *Puccinia striiformis* f. sp. *tritici*, in Turkey from 2018 to 2020. *Can. J. Plant Pathol.* **45**, 158–167 (2023).
7. Kusch, S. et al. Long-term and rapid evolution in powdery mildew fungi. *Mol Ecol* **33**(10), e16909 (2024). <https://doi.org/10.1111/mec.16909>.
8. Zhang, G. et al. Yr5-virulent races of *Puccinia striiformis* f. sp. *tritici* possess relative parasitic fitness higher than current main predominant races and potential risk. *J. Integr. Agric.* **23**, 2674–2685 (2024).
9. Green, A. J. et al. Genetic resistance to and effect of leaf rust and powdery mildew on yield and its components in 50 soft red winter wheat cultivars. *Crop Prot.* **64**, 177–186 (2014).
10. Zhang, Q., Men, X., Hui, C., Ge, F. & Ouyang, F. Wheat yield losses from pests and pathogens in China. *Agric. Ecosyst. Environ.* **326**, 107821 (2022).
11. Draz, I. S., Esmail, S. M., Abou-Zeid, M. A. E. H. & Essa, T. A. E.-M. Powdery mildew susceptibility of spring wheat cultivars as a major constraint on grain yield. *Annals Agricultural Sci.* **64**, 39–45 (2019).

12. Dwivedi, S. L., Scheben, A., Edwards, D., Spillane, C. & Ortiz, R. Assessing and exploiting functional diversity in germplasm pools to enhance abiotic stress adaptation and yield in cereals and food legumes. *Front Plant. Sci* **8**, 1461 (2017). <https://doi.org/10.3389/fpls.2017.01461>.
13. Naseri, B. Estimating yield in commercial wheat cultivars using the best predictors of powdery mildew and rust diseases. *Front Plant. Sci* **13**, 1056143 (2022). <https://doi.org/10.3389/fpls.2022.1056143>.
14. Hatfield, J. L. & Dold, C. Water-Use efficiency: advances and challenges in a changing climate. *Front Plant. Sci* **10**, 103 (2019). <https://doi.org/10.3389/fpls.2019.00103>.
15. Zulfikri, A., Ningsih, E. M. N., Harsono, I. & Susanto, H. Agricultural adaptation strategies to weather fluctuations for improved agribusiness sustainability in West Java. *West. Sci. Nat. Technol.* **2**, 17–23 (2024).
16. USDA. Grain and Feed Annual – 2022. (2024).
17. Rosenzweig, C. & Parry, M. L. Potential impact of climate change on world food supply. *Nature* **367**, 133–138 (1994).
18. Fan, M. et al. Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. *J. Exp. Bot.* **63**, 13–24 (2012).
19. Lobell, D. B. & Schlenker, W. Costa-Roberts, J. Climate trends and global crop production since 1980. *Sci.* (1979). **333**, 616–620 (2011).
20. Fei, L. et al. Maize, wheat and rice production potential changes in China under the background of climate change. *Agric. Syst.* **182**, 102853 (2020).
21. Hou, P. et al. Quantifying maize grain yield losses caused by climate change based on extensive field data across China. *Resour. Conserv. Recycl.* **174**, 105811 (2021).
22. Challinor, A. J. et al. A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Chang.* **4**, 287–291 (2014).
23. Zhang, H., Tang, Y., Chandio, A. A. & Sargani, G. R. Ankrah twumasi, M. Measuring the effects of climate change on wheat production: evidence from Northern China. *Int. J. Environ. Res. Public Health.* **19**, 12341 (2022).
24. Zhao, C. et al. Temperature increase reduces global yields of major crops in four independent estimates. *Proc. Natl. Acad. Sci.* **114**, 9326–9331 (2017).
25. Singh, J., Chhabra, B., Raza, A., Yang, S. H. & Sandhu, K. S. Important wheat diseases in the US and their management in the 21st century. *Front Plant. Sci* **13**, 1010191 (2023). <https://doi.org/10.3389/fpls.2022.1010191>.
26. Kaur, B. et al. Omics for the improvement of abiotic, biotic, and agronomic traits in major cereal crops: applications, challenges, and prospects. *Plants* **10**, 1989 (2021).
27. Saharan, M. S. et al. Wheat blast disease—an overview. *J. Wheat Res.* **8**(1), 1–5 (2016).
28. Naseri, B., Sasani, S. & Fareghi, S. Wheat productivity as a function of cultivar, planting date, powdery mildew and weather interplay. *Res. J. Bot.* **20**, 47–54 (2025).
29. Naseri, B. & Sheykholeslami, M. Powdery mildew development is highly associated with a combination of sowing date, weather, wheat cultivar, and maturity. *J. Agricultural Sci. Technol.* **23**, 1367–1378 (2021).
30. Naseri, B. & Sheykholeslami, M. Progression of powdery mildew in susceptible-resistant wheat (*Triticum aestivum*) cultivars sown at different dates. *J. Phytopathol.* **169**, 640–647 (2021).
31. Naseri, B. & Jalilian, F. Characterization of leaf rust progress in wheat cultivars with different resistance levels and sowing dates. *Eur. J. Plant. Pathol.* **159**, 665–672 (2021).
32. Naseri, B. & Sasani, S. Cultivar, planting date and weather linked to wheat leaf rust development. *Cereal Res. Commun.* **48**, 203–210 (2020).
33. Naseri, B. & Sharifi, F. Predicting wheat Stripe rust epidemics according to influential Climatic variables. *J. Plant. Prot. Res.* <https://doi.org/10.24425/jppr.2019.131260> (2023).
34. Naseri, B., Sasani, S. S. & Rust Planting date, wheat maturity and genetic resistance, weather and productivity. *World J. Environ. Biosci.* **13**, 1–6 (2024).
35. Singh, R. P. et al. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annu. Rev. Phytopathol.* **49**, 465–481 (2011).
36. Jevtić, R. & Župunski, V. The challenge of managing yellow rust (*Puccinia striiformis* f.sp. *tritici*) in winter wheat: how combined climate and pathogen stressors impact variability in genotype reactions. *Front Plant. Sci* **14**, 1270087 (2023). <https://doi.org/10.3389/fpls.2023.1270087>.
37. Castrillo-Sequeira, D., Jiménez-Robles, R. & Granados-Montero, M. Current situation of viticulture in Costa Rica and management strategies for downy mildew (*Plasmopara viticola*). *Revista Mexicana De Fitopatología Mexican J. Phytopathology* **42**(1), 7 (2023). <https://doi.org/10.18781/r.mex.fit.2309-3>.
38. Ullah, A. et al. Worldwide wheat diseases their current status and mode of resistance: A review. *Asian J. Biotechnol. Genetic Eng.* **7**, 139–149 (2024).
39. Aoun, M. et al. Genome-Wide association mapping of leaf rust response in a durum wheat worldwide germplasm collection. *Plant Genome* **9**, (2016). <https://doi.org/10.3835/plantgenome2016.01.0008>.
40. Morgounov, A. et al. Global incidence of wheat rusts and powdery mildew during 1969–2010 and durability of resistance of winter wheat variety Bezostaya 1. *Eur. J. Plant. Pathol.* **132**, 323–340 (2012).
41. Figueroa, M., Hammond-Kosack, K. E. & Solomon, P. S. A review of wheat diseases—a field perspective. *Mol. Plant. Pathol.* **19**, 1523–1536 (2018).
42. Yao, Z., Lei, Y. & He, D. Early visual detection of wheat Stripe rust using Visible/Near-Infrared hyperspectral imaging. *Sensors* **19**, 952 (2019).
43. Naseri, B. & Kazemi, H. Structural characterization of Stripe rust progress in wheat crops sown at different planting dates. *Heliyon* **6**, e05328 (2020).
44. Naseri, B. & Safaei, D. Rust, weather & wheat yield yellow rust, wheat cultivar, and weather conditions are influencing crop yield. *World J. Environ. Biosci.* **12**, 20–26 (2023).
45. Wang, A. et al. A rapid tool for quantification of latent infection of wheat leaves by powdery mildew1. *J. Integr. Agric.* <https://doi.org/10.1016/j.jia.2024.06.004> (2024).
46. An, L. et al. Classification of wheat powdery mildew based on hyperspectral: from leaves to canopy. *Crop Prot.* **177**, 106559 (2024).
47. Gaurav, S. & Chaudhary, V. Do farmers care about basis risk? Evidence from a field experiment in India. *Clim. Risk Manag.* **27**, 100201 (2020).
48. Fenu, G. & Mallocci, F. M. Forecasting plant and crop disease: an explorative study on current algorithms. *Big Data Cogn. Comput.* **5**, 2 (2021).
49. Gouache, D., Léon, M. S., Duyme, F. & Braun, P. A novel solution to the variable selection problem in window pane approaches of plant pathogen – Climate models: development, evaluation and application of a Climatological model for brown rust of wheat. *Agric. Meteorol.* **205**, 51–59 (2015).
50. Hayat, T., Erbay, H., Varçın, F., Hayat, F. & Akci, N. The classification of wheat yellow rust disease based on a combination of textural and deep features. *Multimed Tools Appl.* **82**, 47405–47423 (2023).
51. Sperschneider, J. Machine learning in plant–pathogen interactions: empowering biological predictions from field scale to genome scale. *New Phytol.* **228**, 35–41 (2020).
52. Huerta-Espino, J. et al. Global status of wheat leaf rust caused by *puccinia triticina*. *Euphytica* **179**, 143–160 (2011).
53. Giroto, C. D. et al. A critical review of digital technology innovations for early warning of water-related disease outbreaks associated with Climatic hazards. *Int. J. Disaster Risk Reduct.* **100**, 104151 (2024).

54. Delfani, P., Thuraga, V., Banerjee, B. & Chawade, A. Integrative approaches in modern agriculture: iot, ML and AI for disease forecasting amidst climate change. *Precis Agric.* <https://doi.org/10.1007/s11119-024-10164-7> (2024).
55. González-Domínguez, E., Caffi, T., Rossi, V., Salotti, I. & Fedele, G. Plant disease models and forecasting: changes in principles and applications over the last 50 years. *Phytopathology* **113**, 678–693 (2023).
56. Lindell, C. A., Irish-Brown, A., Rothwell, N. L. & Wallis, A. E. Pest and disease risk and management in high-density perennial crops: current knowledge and areas of future research. *Crop Prot.* **165**, 106150 (2023).
57. Wakhungu, C. Loss of soil biodiversity through judicious use of synthetic pesticides; A case study of trans Nzoia County Kenya. *Scientific Rep. Life Sciences* **3**(4), 1–24. (2023). <https://doi.org/10.5281/zenodo.10492764>.
58. Venugopal, E., Krishna, S. B. N., Golla, N., Patil, S. J. & Nanofungicides, A promising solution for Climate-Resilient plant disease management. In *Plant Quarantine Challenges Under Climate Change Anxiety* 513–532 (Springer Nature Switzerland, 2024). https://doi.org/10.1007/978-3-031-56011-8_16.
59. Boonupara, T., Udomkun, P., Khan, E. & Kajitvichyanukul, P. Airborne pesticides from agricultural practices: A critical review of pathways, influencing factors, and human health implications. *Toxics* **11**, 858 (2023).
60. Jain, H., Dhupper, R., Shrivastava, A., Kumar, D. & Kumari, M. Leveraging machine learning algorithms for improved disaster preparedness and response through accurate weather pattern and natural disaster prediction. *Front Environ. Sci* **11**, 1194918 (2023). <https://doi.org/10.3389/fenvs.2023.1194918>.
61. Singh, M. et al. Growth, yield and economic evaluation of early mint technology under Poplar based silvi-medicinal system in the foothills of himalaya. *Agroforest Syst.* **96**, 205–219 (2022).
62. Deshpande, S. B., Fehrenbacher, J. B. & Ray, B. W. Mollisols of Tarai region of Uttar pradesh, Northern india, 2. Genesis and classification. *Geoderma* **6**, 195–201 (1971).
63. Loegering, W. Q. *Methods for Recording Cereal Rust Data in International Spring Wheat Rust Nursery (IRN)*. (1959).
64. Saari, E. E. & Prescott, J. M. A Scale for Appraising the Foliar Intensity of Wheat Diseases. vol. 59 (Plant disease reporter) (1975).
65. Hoerl, A. E. & Kennard, R. W. Ridge regression: biased Estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
66. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R Stat. Soc. Ser. B Stat. Methodol.* **67**, 301–320 (2005).
67. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Ann Stat* **29**(5), 1189–1232 (2001). <https://doi.org/10.1214/aos/1013203451>.
68. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R Stat. Soc. Ser. B Stat. Methodol.* **58**, 267–288 (1996).
69. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
70. Liaw, A. & Wiener, M. Breiman and Cutler's random forests for classification and regression. R package version, (4.6). 6–14 Preprint at (2015).
71. Jamieson, P. D., Porter, J. R. & Wilson, D. R. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in new Zealand. *Field Crops Res.* **27**, 337–350 (1991).
72. Verreet, J. A., Klink, H. & Hoffmann, G. M. Regional monitoring for disease prediction and optimization of plant protection measures: the IPM wheat model. *Plant. Dis.* **84**, 816–826 (2000).
73. Ali, Y. et al. Predicting Stripe rust severity in wheat using meteorological data with environmental response modeling. *J. King Saud Univ. Sci.* **35**, 102591 (2023).
74. Singh, R., Krishnan, P., Singh, V. K. & Das, B. Estimation of yellow rust severity in wheat using visible and thermal imaging coupled with machine learning models. *Geocarto Int* **38**(1), 2160831 (2023). <https://doi.org/10.1080/10106049.2022.2160831>.
75. Ahlawat, Y. S. *PLANT PATHOLOGY: Crop Diseases and Their Management* Vol. 110012 (Division of Plant Pathology Indian Agricultural Research Institute, 2008).
76. Rana, V., Batheja, A., Sharma, R., Rana, A. & Priyanka Springer Nature Singapore, Singapore., Powdery Mildew of Wheat: Research Progress, Opportunities, and Challenges. In *New Horizons in Wheat and Barley Research* 133–178 (2022). https://doi.org/10.1007/978-981-16-4134-3_5.
77. Rodríguez-Moreno, V. M., Jiménez-Lagunes, A., Estrada-Avalos, J., Mauricio-Ruvalcaba, J. E. & Padilla-Ramírez, J. S. Weather-data-based model: an approach for forecasting leaf and Stripe rust on winter wheat. *Meteorological Applications* **27**, e1986 (2020). <https://doi.org/10.1002/met.1896>.
78. El-Orabey, W. M. & Elkot, A. F. Prediction of leaf rust severity and yield loss in wheat based on environmental factors. *J. Plant Dis. Prot.* **127**, 507–519 (2020).
79. El Jarroudi, M. et al. Weather-Based predictive modeling of wheat Stripe rust infection in Morocco. *Agronomy* **10**, 280 (2020).
80. Tan, Z. H. et al. Surface conductance for evapotranspiration of tropical forests: calculations, variations, and controls. *Agric. Meteorol.* **275**, 317–328 (2019).
81. Valipour, M. & Guzmán, S. M. Identification of the meteorological variables influencing evapotranspiration variability over Florida. *Environ. Model. Assess.* **27**, 645–663 (2022).
82. Venter, Z. S., Brousse, O., Esau, I. & Meier, F. Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sens. Environ.* **242**, 111791 (2020).
83. Vijaya Kumar, P. Development of weather-based prediction models for leaf rust in wheat in the Indo-Gangetic plains of India. *Eur. J. Plant. Pathol.* **140**, 429–440 (2014).
84. El Jarroudi, M. et al. A Threshold-Based weather model for predicting Stripe rust infection in winter wheat. *Plant. Dis.* **101**, 693–703 (2017).
85. Singh, M., Savary, S., Setiya, P., Tewari, A. K. & Nain, A. S. Multivariate analysis of the effects of weather variables on white rust epidemics and yield reduction of mustard over multiple growing seasons. *Plant. Pathol.* **73**, 791–809 (2024).
86. Mulatu, W. B., Bedasa, M. F. & Terefa, G. K. Prediction of Wheat Rust Diseases Using Data Mining Application. *OALib* **07**, 1–27 (2020).
87. Te Beest, D. E., Paveley, N. D., Shaw, M. W. & van den Bosch, F. Disease–Weather relationships for powdery mildew and yellow rust on winter wheat. *Phytopathology* **98**, 609–617 (2008).
88. Garrett, K. A. et al. Elsevier, Plant pathogens as indicators of climate change. In *Climate Change* 499–513 (2021). <https://doi.org/10.1016/B978-0-12-821575-3.00024-4>.
89. Friedrich, S. Modelling infection probability of powdery mildew in winter wheat by meteorological input variables. *Zeitschrift Fur Pflanzenkrankheiten Und Pflanzenschutz- J. Plant Dis. Prot.* **108** 354–365 (1995).
90. Wiese & Maurice Victor. Compendium of wheat diseases. *American Phytopathological Society* 112 (1987).
91. Bhardwaj, N. R., Banyal, D. K. & Roy, A. K. Prediction model for assessing powdery mildew disease in common oat (*Avena sativa* L.). *Crop Prot.* **146**, 105677 (2021).
92. Misra, A. K., Yadav, S. B., Mishra, S. K. & Tripathi, M. K. Impact of Meteorological Variables and Climate Change on Plant Diseases. in *Plant Pathogens* 313–327 Apple Academic Press, Includes bibliographical references and index., (2020). <https://doi.org/10.1201/9780429057212-18>.
93. Janowski, D. & Leski, T. Factors in the distribution of mycorrhizal and soil fungi. *Divers. (Basel)* **14**, 1122 (2022).
94. Ahmad, A. & Saraswat, D. El gamal, A. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agricultural Technol.* **3**, 100083 (2023).
95. Tonle, F. B. N. et al. A road map for developing novel decision support system (DSS) for disseminating integrated pest management (IPM) technologies. *Comput. Electron. Agric.* **217**, 108526 (2024).
96. Valleggi, L. & Stefanini, F. M. A mini-review on data science approaches in crop yield and disease detection. *Frontiers Agronomy* **6**, 1352219 (2024). <https://doi.org/10.3389/fagro.2024.1352219>.

97. Araújo, S. O., Peres, R. S., Ramalho, J. C., Lidon, F. & Barata, J. Machine learning applications in agriculture: current trends, challenges, and future perspectives. *Agronomy* **13**, 2976 (2023).
98. Nguyen, C., Sagan, V., Skobalski, J. & Severo, J. I. Early detection of wheat yellow rust disease and its impact on terminal yield with Multi-Spectral UAV-Imagery. *Remote Sens. (Basel)*. **15**, 3301 (2023).
99. Gao, R., Jin, F., Ji, M. & Zuo, Y. Research on the method of identifying the severity of wheat Stripe rust based on machine vision. *Agriculture* **13**, 2187 (2023).

Author contributions

A.B. and R.R. performed Data curation, Formal analysis, Methodology, Investigation, Writing – original draft, Writing – review & editing. M.S. and R.P. wrote original draft, review & editing, Formal analysis. S.C. and A.J.N. reviewing & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Competing interests

The authors declare no competing interests.

Conflict of interest

The authors declare there is no financial interests/personal relationships which may be considered as potential competing interests that could have appeared to influence the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to R.R. or S.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025