



Revolutionizing wheat crop disease prediction: A novel framework on integrating nature-inspired random forest optimization and explainable artificial intelligence (XAI) in Morocco

Satiprasad Sahoo ^{*} , Chiranjit Singha , Ajit Govind

International Center for Agricultural Research in the Dry Areas (ICARDA), 2 Port Said, Victoria Sq, Ismail El-Shaer Building, Maadi, 15A, Cairo, 11728, Egypt

ARTICLE INFO

Keywords:

Wheat crop diseases
Explainable artificial intelligence
CMIP6
Sustainable
Development goals

ABSTRACT

Precision agriculture has substantial hurdles in wheat crop disease because of the complicated environmental variability. This study presents a novel method by fusing Random Forest models optimized using nature-inspired algorithms like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), Differential Evolution (DE), and Honey Badger Algorithm (HBA) with explainable artificial intelligence (XAI). Nine important climatic and biophysical factors are assessed, such as precipitation (Pr), maximum and minimum temperatures (Tmax and Tmin), the Green Chlorophyll Index (CGI), the Normalized Difference Chlorophyll Index (NDCI), the Modified Soil-Adjusted Vegetation Index 2 (MSAVI2), the Normalized Difference Vegetation Index (NDVI), soil organic carbon (SOC), and total nitrogen (TN). Multicollinearity analysis and Boruta were used to evaluate feature importance. SSP2-4.5 and SSP5-8.5 scenarios from CMIP6 were used to predict climate projections using GEE (1990–2030) across three GCMs (EC-Earth3, NorESM2-LM, and MIROC6). Blight, rust, fusarium wilt, and powdery mildew disease predictions were verified using 10-fold cross-validation and field data observed by farmers. Based on current research, the best AUC for powdery mildew illness (AUC = 0.946) was represented by RF-GWO, the highest AUC for rust (AUC = 0.897) was recorded by the RF-DE model, and the highest AUC values for blight (AUC = 0.833) and fusarium wilt (AUC = 0.836) diseases were created by RF-GA. RF-GA performs the best on average. The method's importance for sustainable agriculture and SDG accomplishment was supported by the XAI interpretation, which identified temperature and precipitation as major disease-promoting factors. Furthermore, it is a cutting-edge decision-support system that transforms Morocco's wheat disease management by fusing XAI with cutting-edge nature-inspired random forest optimization (NIRFO).

1. Introduction

Biotic stresses such as pests and diseases have a significant impact on global agricultural production, resulting in projected losses of up to \$220 billion each year. Cereals, like all main crops in Africa, are susceptible to several plant diseases [1]. Since 1987, disease monitoring efforts in Morocco have identified several wheat-related diseases, with the most severe and destructive being the septoria/tan spot complex, leaf rust, stripe (yellow) rust, and root rot [2]. Globally, wheat ranks among the three primary staple crops, alongside rice and maize [3]. Wheat farming, both internationally and in Morocco, is posing considerable challenges due to several diseases. These diseases have the potential to negatively affect both the yield and quality of crops. While

they may result from bacterial or fungal pathogens, fungal infections generally present a more significant threat to grain production and quality. Among these, Fusarium is considered one of the most harmful fungal diseases affecting cereal crops, particularly wheat [4]. Early and precise information on disease outbreaks is critical for successful fungicide use and economic benefits in wheat harvests [3]. Thus, the current study focuses on predicting wheat illnesses in Morocco using a novel framework that combines nature-inspired random forest optimization (NIRFO) with explainable artificial intelligence (XAI). This research can help integrate genetic advancements, precision agriculture, and adaptive cultural practices for sustainable crop management.

Numerous studies have identified various wheat crop diseases that pose a threat to agriculture in Morocco. In a study conducted by

^{*} Corresponding author. International Center for Agricultural Research in the Dry Areas (ICARDA), 2 Port Said, Victoria Sq, Ismail El-Shaer Building, Maadi, 15A, Cairo, 11728, Egypt.

E-mail address: satisps@gmail.com (S. Sahoo).

<https://doi.org/10.1016/j.jafr.2026.102924>

Received 11 June 2025; Received in revised form 20 November 2025; Accepted 10 April 2026

Available online 14 April 2026

2666-1543/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

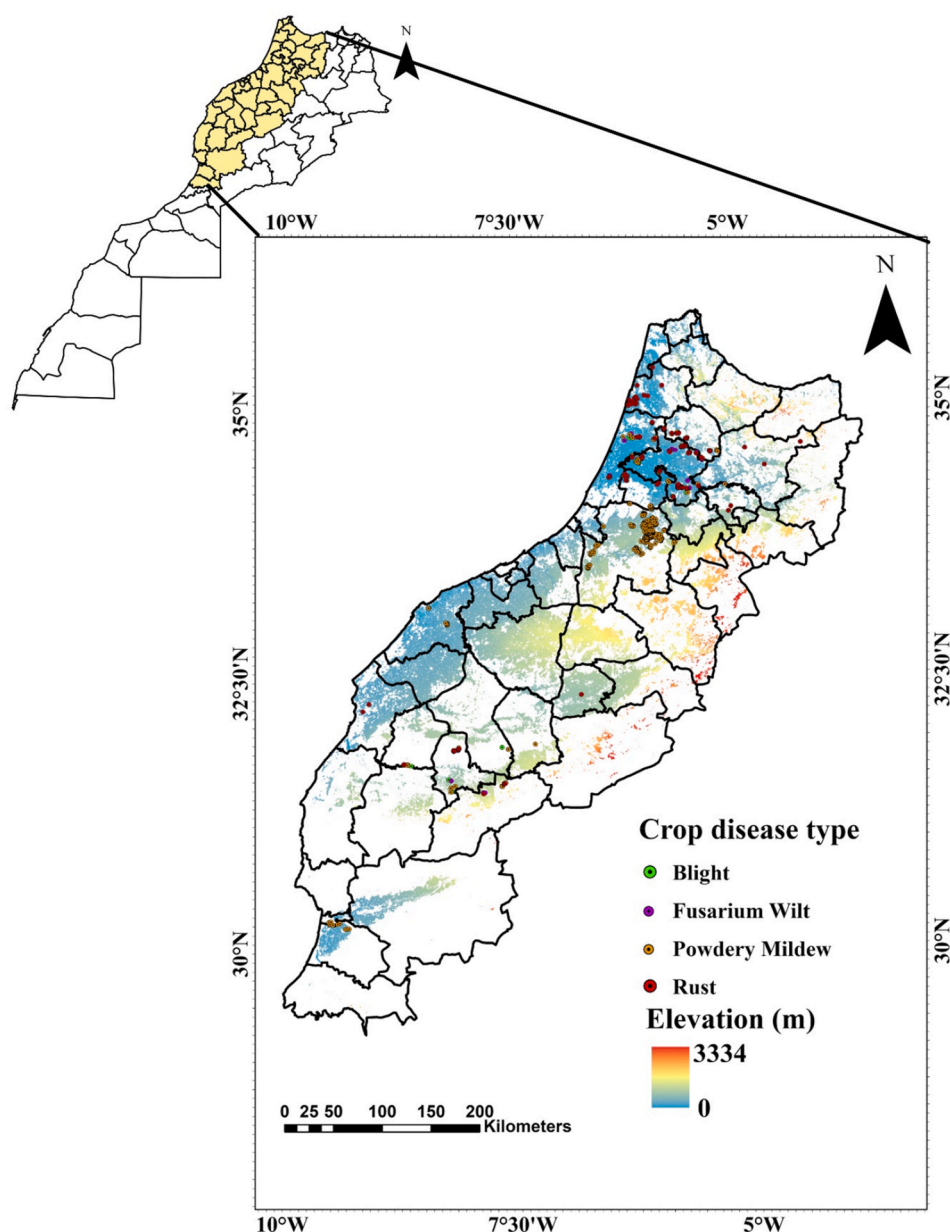


Fig. 1. Location of the study area.

Bozalmat et al. [4], the effectiveness of various farming techniques in controlling three fungal diseases affecting both leaves and ears of durum wheat was evaluated. This research was carried out over two consecutive years 2019 and 2020 at two different locations: an organic-certified farm in the Benslimane region in 2019 and the experimental farm of the National School of Agriculture in Meknes in 2020. Additionally, El Jarroudi et al. [5] developed a weather-based predictive model for infections caused by *Puccinia striiformis* f. sp. *tritici*. This model plays a critical role in assessing the timing and necessity of fungicide applications to curb wheat stripe rust (WSR), which can lead to significant yield losses. El Jarroudi et al. [5] described the use of an artificial intelligence algorithm, i.e., the artificial neural network, in a weather-based model strategy to forecast the advancement of *Septoria tritici* blotch (STB) in wheat in Luxembourg. Bouarda et al. [6] aimed their research on Fusarium root rot (FRR), one of the common diseases of the Mediterranean drylands caused by *F. culmorum*, which enormously affects the yield of durum wheat. The objective of this study is to enhance understanding of the fungus's virulence and the interactions between various durum wheat lines and fungal isolates. Additionally, the research

identified several potential resistance sources for breeders, though the mechanisms are not yet understood. Zahraoui et al. [7] explored an eco-friendly alternative for managing helminthosporium leaf spot, caused by *Bipolaris sorokiniana*, in four varieties of durum wheat (*Triticum durum*) and common wheat (*Triticum aestivum*). The researchers assessed the potential of essential oils derived from medicinal plants to both control the disease and enhance crop yield. In an earlier study, Zahraoui et al. [8] examined the impact of essential oils extracted from *Thymus satureioides* and *Origanum compactum* on root rot in wheat, specifically targeting infections caused by *Fusarium culmorum* and *Bipolaris sorokiniana*.

Various constraints have led to a notable decline in wheat production, adversely affecting the country's economic and social stability. As a result, it is important to revise the current production methods by adopting a more organized and comprehensive strategy, especially with regard to disease control practices.

However, the research questions are:

- (a) How may optimization algorithms inspired by nature improve a basic Random Forest (RF) model's ability to predict various wheat diseases?
- (b) Which particular biophysical and climatic factors, as determined by XAI and robust feature selection, are most strongly linked to disease outbreaks?
- (c) By 2030, how will the regional and temporal risk of major wheat diseases in Morocco be affected by the future climatic scenarios (SSP2-4.5 and SSP5-8.5)?
- (d) How much can an interpretable AI framework help sustainable wheat production and disease management with practical, field-level insights?

However, more researchers have tried several machine learning (ML) and deep learning (DL) models to identify agricultural diseases, and the results reveal that the RF ML model outperforms the others. Thus, the current study's innovation lies in its integration of NIRFO with XAI for sustainable wheat crop disease control. The current work investigates various forms of RF models for agricultural diversification utilizing a geospatial setting.

The following sections focus on the important diseases of wheat in Morocco and provide detailed information using ML algorithms. Additionally, we considered the geographical distribution, impact, and disease management strategies and briefly addressed the current status along with anticipated future scenarios.

Therefore, the objectives of our research were to (1) collect wheat crop disease data and create distribution maps for 2024 and 2030 using different ML algorithms for disease detection; (2) compare the accuracies of different algorithms and identify the algorithm with the most satisfactory performance.

2. Study area

The study area extends from 3°53'06" W to 10°26'21" W longitude and 29°27'68" N to 35°54'62" N latitude, encompassing approximately 172,521.5 km² (Fig. 1). Situated in northern Morocco, the region includes 50 administrative provinces and supports a population of around 28.6 million. The map overlays field-surveyed data highlighting the occurrence of major crop diseases such as Blight, Fusarium Wilt, Powdery Mildew, and Rust. Elevation across the region ranges from 0 to 3334 m, displayed using a continuous color gradient to reflect variations in topography, which can influence disease prevalence. Climate records show an average yearly temperature of approximately 17.5 °C, with monthly temperatures varying from around 9.4 °C in the winter months (December–January) to about 26 °C during the summer (July–August). Annual rainfall averages nearly 318.8 mm, with most precipitation occurring between October and April, while the summer months (June to August) experience very little rainfall (<https://climateknowledgeportal.worldbank.org/>).

3. Materials and methods

3.1. Crop disease inventory

Data on wheat crop diseases (WCD) across major wheat-producing regions of Morocco were obtained in 2024 through geotagged field surveys. A total of 781 agricultural plots were assessed, with systematic recording of disease presence and their spatial patterns. From the wide spectrum of diseases observed, four key types powdery mildew, rust, fusarium wilt, and blight were prioritized for further investigation due to their high incidence and agronomic impact. To achieve a representative and unbiased dataset, stratified sampling was implemented, ensuring proportional representation of these disease categories across varying field conditions and locations. All WCDs recorded a range of wheat varieties, including *Triticum durum*, Ough, Marzouki, Blé tendre Sids, Marrouane, Baledi, Blé Marzouki, Sebou, Massira, Kanz, Blé dur,

Karim, Annouar, Amjad, and Rihane. By adopting this approach, the study effectively reduced the risk of systematic biases and variances, which could otherwise result in misleading or distorted outcomes. This method not only improved data precision but also provided a solid foundation for analyzing disease patterns across the surveyed areas. Comprehensive data were collected across multiple dimensions to build a robust framework for wheat yield optimization and disease management. This included critical factors such as crop-growing periods, fertilizer application rates, supplemental nutrient usage, irrigation practices, farmers' socio-economic conditions, and the total expenditure costs associated with wheat cultivation. This multifaceted dataset provides valuable insights into the interplay between agricultural practices and crop performance, offering a foundation for enhancing crop resilience and addressing disease-related challenges faced by local farming communities. By leveraging this diverse information, researchers can conduct a thorough analysis to identify patterns and correlations that influence wheat health and productivity. Furthermore, the datasets on crop disease types serve as key dependent parameters, enabling the classification and detailed analysis of disease occurrences. This approach not only supports the development of targeted strategies for disease mitigation but also empowers farmers with data-driven solutions to improve yield sustainability and economic outcomes.

In this study, WCDs were systematically categorized into distinct labels to facilitate analysis: label 1 represented blight, label 2 denoted powdery mildew, label 3 indicated rust, and label 4 was also assigned to fusarium wilt. To ensure robust evaluation, a 10-fold cross-validation methodology was implemented during the ML process. The datasets for each disease category were partitioned into training and testing subsets, with 80% allocated for training and the remaining 20% reserved for testing. In the dataset, a total of 80 instances of blight were identified, with approximately 64 instances allocated for training and the remaining 16 reserved for testing. For powdery mildew, out of 545 instances, around 436 were used for training, while 109 were designated for testing. The rust category comprised 111 instances, with 88 assigned to training and 23 to testing. Similarly, for fusarium wilt, there were 45 instances in total, of which 36 were used for training and 9 were set aside for testing. This structured approach ensured a comprehensive evaluation of the ML model's performance across different disease categories.

3.2. Crop diseases influencing factors (CDIFs)

The creation of thematic maps to visualize crop disease occurrences is a complex process driven by the intricate interplay of various environmental and agricultural factors (SF 1). These include climatic conditions, soil nutrient levels, and vegetation health, all of which interact in non-linear ways to influence disease patterns. Accurately mapping and classifying crop diseases requires careful consideration of these multifaceted influences, each contributing distinct effects on disease prevalence. To ensure the accuracy and relevance of such mappings, input from a diverse group of stakeholders is essential. This includes experts in the field, agronomists, and local farmers, whose collective insights help identify the most suitable factors for analysis. Previous research has already identified and outlined the critical elements that shape crop disease classification mapping [9]. These studies have provided valuable guidance on the factors to incorporate into classification models, laying the groundwork for more precise and effective disease mapping strategies.

In this study, CDIF layers were precisely generated using ArcGIS v.10.7 software. These layers were constructed from a comprehensive geospatial database that incorporated nine key CDIFs. Each factor was carefully selected to act as an indicator of environmental and agricultural conditions that influence the suitability for crop disease development. This approach enabled the creation of detailed and accurate maps, providing valuable insights into areas prone to crop diseases based on the interplay of these critical factors. Soil nutrient factors, including Soil

Table 1
Data description and statistical overview of all the CDIF layers.

Index	Description	Rationale for selection	Formula	Min	Max	Mean	Std.Dev	Source
Precipitation (mm)	IDAHO_EPSCOR/TERRACLIMATE, 4638.3 m, (1990-2024)	key driver of wheat disease occurrence, as it controls soil moisture and leaf wetness essential for pathogen development. Higher rainfall increases disease incidence, while low precipitation intensifies crop stress and vulnerability [11].		106.51960	894.73040	366.75301	123.3319	[10]
Tmax (°C)		temperatures critically affect wheat disease outbreaks by regulating spore germination and host susceptibility. High Tmax accelerates disease development, while elevated Tmin enhances leaf wetness and favors foliar infections [12].		14.508333	28.36347	25.05796	1.44095	
Tmin (°C)		temperatures critically affect wheat disease outbreaks by regulating spore germination and host susceptibility. High Tmax accelerates disease development, while elevated Tmin enhances leaf wetness and favors foliar infections [12].		−1.56029	16.23480	11.99926	1.84844	
Soil Organic Carbon (SOC) (%)	ISRIC — World Soil Information, SoilGrids250m, (0-30 cm depth)	key factor as it enhances soil fertility, microbial activity, and plant defense mechanisms. Low SOC weakens crop resilience, making plants more susceptible to pathogen attack. Thus, SOC reflects the soil's health and capacity to support healthy wheat growth and disease resistance [13].		3.4000	78.09999	11.441433	4.050507	URL: https://data.isric.org
Total Nitrogen (TN) (cg/kg)		key variables influencing by regulating plant vigor and susceptibility to pathogens. Excessive TN promotes dense canopies and humidity favorable to foliar pathogens, while deficiency weakens plant defenses, increasing infection risk [14].		4.82999	31.7199	9.67241	1.534859	
NDVI	COPERNICUS/S2_SR, 10m (2024)	Vegetation indices are used for wheat disease detection as they capture canopy and chlorophyll changes linked to stress.	$\frac{(NIR - RED)}{(NIR + RED)}$	−0.25497	0.71420	0.18603	0.08803	URL: https://browser.dataspace.copernicus.eu/
GCI		NDVI reflects crop vigor, while GCI and NDCI indicate chlorophyll reduction due to disease.	$\frac{NIR}{GREEN} - 1$	−0.6026	3.39515	0.76836	0.265126	URL: https://browser.dataspace.copernicus.eu/
MSAVI2		MSAVI2 minimizes soil background effects, improving early stress detection.	$MSAVI2 = \frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)})}{2}$	−0.68407	0.833267	0.30485	0.11595	
NDCI		These indices reveal physiological and structural alterations before visible symptoms appear. Hence, they serve as reliable indicators for monitoring wheat crop health through RS [15, 16].	$\frac{(Red\ Edge\ 1 - RED)}{(Red\ Edge\ 1 + RED)}$	−0.07051	0.36203	0.06495	0.02766	URL: https://browser.dataspace.copernicus.eu/
Future Precipitation (mm), Tmax (°C), Tmin (°C)	NASA/GDDP-CMIP6, 27830 m, (1990-2030), (SSP2-4.5, SSP5-8.5)	Global warming affects pathogen survival, virulence, and geographic distribution, thereby threatening wheat productivity. Increased CO ₂ and erratic precipitation modify host–pathogen interactions, intensifying disease risks. Modeling these climatic shifts helps predict future disease patterns and supports adaptive management strategies [17, 18].	(URL: https://registry.opendata.aws/nex-gddp-cmip6/)					URL: https://registry.opendata.aws/nex-gddp-cmip6/

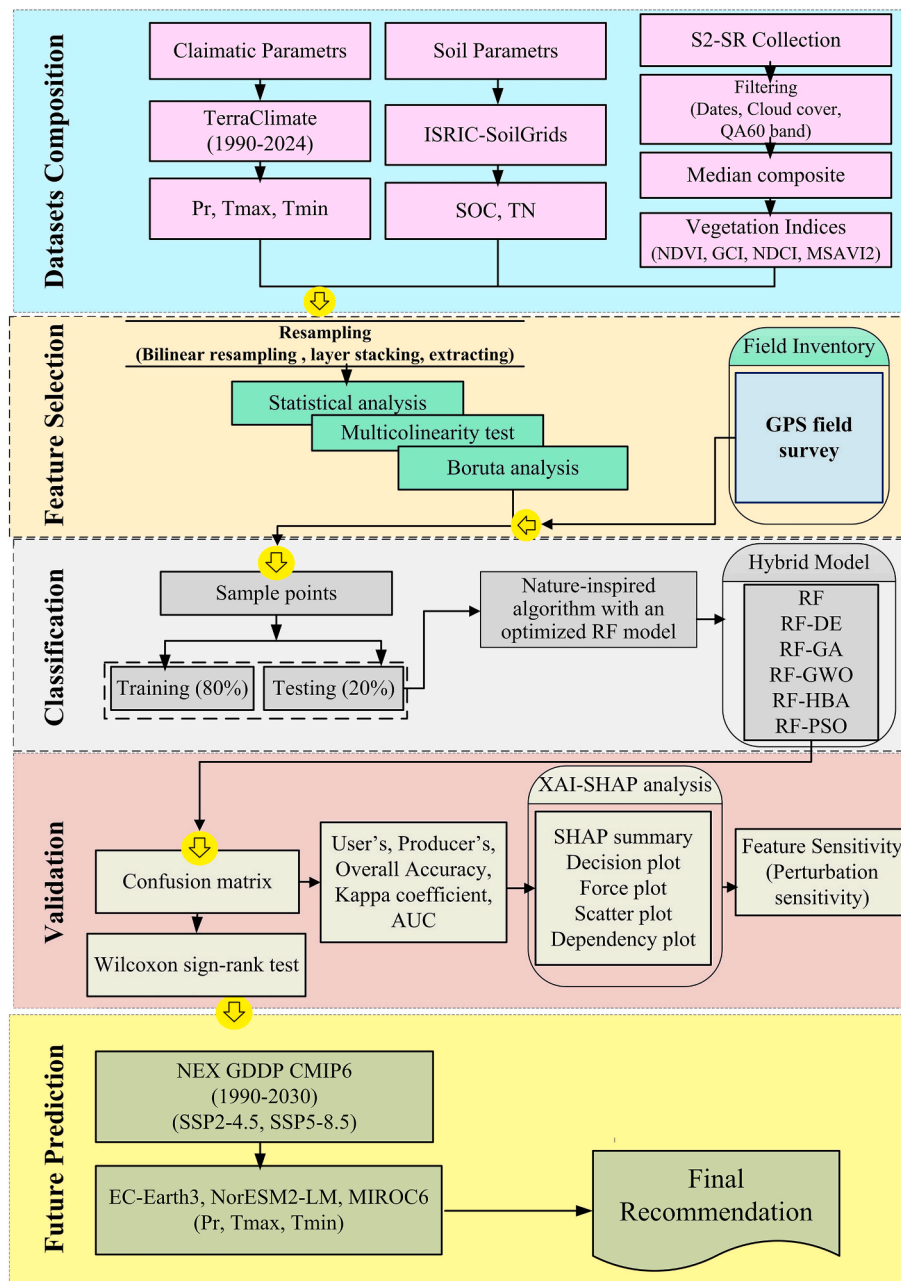


Fig. 2. System architecture and methodology.

Organic Carbon (SOC) and Total Nitrogen (TN), were extracted from the SoilGrids database (250m × 250m), covering a depth range of 0 to 30 m (source: <https://soilgrids.org/>). These factors were analyzed and mapped at the individual pixel level. Climatic factors such as precipitation (Pr), maximum temperature (Tmax), and minimum temperature (Tmin) were obtained from the TerraClimate dataset, which has a spatial resolution of approximately 4638.3 m by 4638.3 m [10]. Remote sensing (RS)-based vegetation indices, including the Green Chlorophyll Index (CGI), Normalized Difference Chlorophyll Index (NDCI), Modified Soil-Adjusted Vegetation Index 2 (MSAVI2), and Normalized Difference Vegetation Index (NDVI), were processed using cloud filtering, radiometric correction, median composite, and mosaicking techniques. These steps were carried out on the Google Earth Engine (GEE) platform using multispectral Sentinel-2 imagery with a spatial resolution of 10 m by 10 m (source: <https://sentiwiki.copernicus.eu/web/s2-applications>).

In this study, the future prediction of crop disease patterns within the region was analyzed by examining various climatic factors, including

Tmax, Tmin, and Pr. These predictions were derived using downscaled CMIP-6 datasets accessible through GEE, as detailed in Table 1. The dataset includes downscaled climate projections on a global scale, created through simulations based on the General Circulation Model (GCM). The analysis focused on two combined Shared Socioeconomic Pathway (SSP) scenarios SSP-245 and SSP-585 which highlight the potential vulnerability of crop disease under varying future conditions. To minimize the uncertainty associated with projected climate scenarios, ensemble averages were calculated using datasets from three GCMs: EcEarth3, NorESM2 LM, and MIROC6, all sourced from GEE. These datasets spanned the period from 1990 to 2030 (see Table 1).

Subsequently, all CDIF layers were resampled using a bilinear resampling function in R software v.4.3.2, achieving a spatial resolution of 30 m by 30 m. These layers were then integrated and optimized through ML algorithms to enhance the accuracy of crop disease predictions for the region. This approach provides a robust framework for understanding and mitigating future crop disease risks in the context of

changing climatic conditions. Table 1 provides a comprehensive summary and statistical overview of all the CDIF layers.

3.3. Methodology

Fig. 2 illustrates the detailed methodological workflow for the WCD mapping analysis. The process is divided into two key stages:

I) Data Preparation Phase: In the initial stage, the WCD inventory data, combined with nine input CDIFs, were utilized to generate training and testing samples for the ML model. This step ensured the availability of structured data for model development and evaluation. II) Feature Assessment and Selection: The subsequent stage involved evaluating the nonlinearity of each factor, assessing multicollinearity, and performing feature selection. The Boruta technique was employed to identify the most relevant features, ensuring the model's efficiency and accuracy in predicting WCD. III) Development of WCD classification Maps: In this phase, WCD classification maps were generated using a novel approach that integrates a NIRFO model. The performance of all models was evaluated using various metrics, including accuracy, producer accuracy, user accuracy, kappa coefficient, and AUC score. These metrics ensured the reliability and robustness of the model predictions. IV) Exploration of Future WCD Distribution Patterns: The subsequent stage focused on analyzing future patterns of WCD distribution (1990-2030) under different climate scenarios. This was achieved using CMIP6 datasets, specifically the SSP2-4.5 and SSP5-8.5 scenarios, to project potential changes in disease suitability and severity over time. V) Sensitivity Analysis and Feature Explainability: Additionally, crop disease sensitivity was assessed through perturbation sensitivity analysis, which provided insights into the impact of varying input factors on disease predictions. Finally, the explainability of the model's features was achieved using the SHAP (SHapley Additive exPlanations) XAI technique. This step enhanced the interpretability of the model, offering a clear understanding of the factors driving WCD predictions. This structured methodology ensures a comprehensive analysis of WCD patterns, combining advanced modeling techniques with interpretable AI to support informed decision-making.

3.4. Multicollinearity test

This study employed a multicollinearity test (MCT) to evaluate the correlation strength among the CDIFs using the Variance Inflation Factor (VIF) and tolerance scores. The analysis determines whether features can be included in the WCD assessment prior to implementing ML models. A VIF threshold exceeding 10 signifies multicollinearity concerns. From an initial set of features, nine CDIFs were selected based on multicollinearity diagnostics.

3.5. Boruta feature selection

The Boruta method is a popular feature selection algorithm that determines significant predictors by systematically comparing original attributes with artificially generated shadow features. This study implements the Boruta feature selection algorithm in conjunction with an RF model. The method iteratively assesses variable significance through standardized Z-scores, computed with Eq. 1

$$Z = \frac{\text{Mean Decrease in Accuracy}}{\sigma} \quad (1)$$

where σ represents the standard deviation. Features exhibiting statistically higher Z-scores than their shadow counterparts are selected as essential, while insignificant ones are discarded. Unlike conventional methods that prioritize minimal feature sets, Boruta adopts an all-relevant strategy, ensuring comprehensive retention of impactful predictors. This approach proves especially valuable for high-dimensional data, improving model transparency and predictive power by

removing irrelevant or redundant variables.

3.6. WCD modeling techniques

3.6.1. Random Forest (RF)

RF is an ensemble-based ML technique that builds several decision trees during the training phase and uses the majority vote of these trees to make predictions in classification problems. It efficiently processes high-dimensional datasets while demonstrating strong resilience to noise and outliers. Additionally, it provides interpretable feature importance rankings, enhancing model transparency. Given a dataset with N samples and M features, an RF model F aggregates predictions from T decision trees $\{h_1, h_2, \dots, h_T\}$. For a multiclass problem with K classes, the predicted class $\hat{y}$ for an input x is determined via majority voting (Eq. (2)):

$$\hat{y} = \arg \max_{k \in \{1, \dots, K\}} \sum_{t=1}^T \mathbb{I}(h_t(x) = k) \quad (2)$$

where $\mathbb{I}(\cdot)$ is the indicator function. Each tree h_t is trained on a bootstrapped subset of data with random feature selection at each split, enhancing robustness against overfitting. The Gini impurity criterion is commonly used for node splitting (Eq. (3)):

$$\text{Gini}(D) = 1 - \sum_{k=1}^K p_k^2 \quad (3)$$

where p_k is the proportion of class k in node D . RF's feature importance is derived from the mean decrease in impurity (MDI), quantifying each feature's contribution to predictive accuracy [19].

The hybridization of five distinct metaheuristic optimization algorithms within the NIRFO framework represents a significant methodological advancement beyond conventional RF modeling. Each optimizer i.e., Genetic algorithm (GA), Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), Differential Evolution (DE), and Honey Badger Algorithm (HBA) were strategically selected for its unique exploration and exploitation dynamics and complementary search mechanisms. When combined with the RF architecture, these algorithms enable dynamic hyperparameter tuning that enhances model adaptability, convergence stability, and predictive accuracy across complex, nonlinear datasets [20]. Unlike traditional RF models with static parameterization, the NIRFO-inspired RF framework introduces a self-adaptive learning environment that simultaneously strengthens interpretability through feature importance and SHAP analyses while ensuring robustness against noise and data imbalance. Thus, this integrated optimization-learning paradigm establishes a novel and intelligent approach to ensemble modeling, transcending the limitations of conventional machine learning and single-optimizer frameworks.

3.6.2. Genetic algorithm (GA)

The GA is an evolutionary metaheuristic that emulates biological selection processes to solve complex optimization problems. Through successive generations, GA evolves candidate solutions using genetic operators selection, crossover, and mutation to progressively approximate optimal solutions [21]. This study employs a GA to optimize RF parameters, including the number of trees (nestimators), maximum depth (dmax), and minimum samples per leaf (mleaff). The rationale behind selecting the RF optimized by GA (RF-GA) lies in the Genetic Algorithm's robust capacity to explore and exploit complex, nonlinear search spaces efficiently. The RF model, while powerful, is highly sensitive to hyperparameter tuning, and conventional grid or random search methods often fail to capture optimal parameter combinations due to their limited exploration ability. The GA, inspired by the principles of natural evolution, selection, crossover, and mutation, offers a dynamic and adaptive optimization mechanism that effectively avoids local minima and enhances global search performance. This makes it particularly suitable for fine-tuning RF parameters such as the number of trees, maximum depth, and feature subsets. Within the NIRFO

framework, RF-GA serves as a foundational model that leverages evolutionary intelligence to improve predictive accuracy, stability, and generalization in wheat disease prediction under diverse agro-climatic conditions in Morocco.

GA algorithm's stochastic optimization efficiently explores hyperparameter spaces, avoiding local optima while enhancing RF generalization through optimal parameter selection [22]. The GA evolves a population of candidate solutions, where each chromosome encodes RF hyperparameters (Eq. (4)).

$$\text{Fitness}(\theta) = \frac{1}{C} \sum_{i=1}^C \text{Accuracy}_i(\text{RF}_\theta) \quad (4)$$

where θ represents the hyperparameters, C is the number of cross-validation folds, and accuracy i is the model's performance on the i -th fold.

Genetic operators selection, crossover, and mutation iteratively improve solutions. Selection favors high-fitness individuals, while crossover and mutation introduce diversity (Eq. (5)):

$$P_{\text{mutation}}(\theta) = \begin{cases} \theta + \Delta\theta & \text{with probability } p_m \\ \theta & \text{otherwise} \end{cases} \quad (5)$$

where $\Delta\theta$ is a small perturbation and p_m is the mutation probability.

3.6.3. Particle Swarm Optimization (PSO)

PSO is a population-based stochastic optimization method that models collective intelligence observed in biological swarms [23]. The selection of the RF optimized by PSO (RF-PSO) is grounded in PSO's exceptional ability to balance exploration and exploitation within multidimensional search spaces. The algorithm iteratively refines a set of potential solutions through coordinated movement patterns derived from avian flocking dynamics, enabling efficient convergence toward optimal solutions. PSO, inspired by the social behavior of bird flocking and fish schooling, enables a population of candidate solutions ("particles") to collaboratively search for the global optimum by sharing positional and velocity information. This collective intelligence mechanism allows PSO to converge rapidly toward promising regions of the solution space while maintaining sufficient diversity to avoid premature convergence. In the context of RF modeling, PSO efficiently tunes critical hyperparameters such as the number of estimators, maximum depth, and feature selection by dynamically adjusting them based on fitness evaluations that reflect model performance. Integrated within the NIRFO system, the RF-PSO model harnesses swarm intelligence to enhance predictive performance, computational efficiency, and model generalization, providing a biologically inspired alternative to traditional optimization methods for wheat disease prediction in Morocco.

The PSO algorithm iteratively adjusts a swarm of particles, where each particle position $\mathbf{x} = (\text{n_estimators}, \text{max_depth}, \text{min_samples_split})$ represents a potential RF configuration [24].

The particle velocity $\mathbf{v}$ and position update equations (Eq. (6)):

$$\begin{aligned} \mathbf{v}_i(t+1) &= w\mathbf{v}_i(t) + c_1r_1(\mathbf{pbest}_i - \mathbf{x}_i(t)) + c_2r_2(\mathbf{gbest} - \mathbf{x}_i(t)) \\ \mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \end{aligned} \quad (6)$$

where w is inertia weight, c_1, c_2 indicate acceleration coefficients, $r_1, r_2 \sim U(0,1)$, $\mathbf{pbest}_i$ is particle's best solution and $\mathbf{gbest}$ represents global best solution.

The fitness function maximizes cross-validated accuracy:

$$f(\mathbf{x}) = 1 / K \sum_{k=1}^K \text{Accuracy}_K(\text{RF}_\mathbf{x}) \quad (7)$$

for K classes.

In PSO configuration through swarm intelligence. Experimental results demonstrate PSO optimized RF achieves 3-5% higher accuracy than grid search on benchmark multiclass datasets, with particular

improvement in handling class imbalance.

3.6.4. Grey Wolf Optimization (GWO)

The GWO represents a nature-inspired metaheuristic approach that mimics the social structure and collective hunting mechanisms of grey wolf packs to address complex optimization challenges. The integration of the GWO algorithm with the RF model was motivated by GWO's unique capability to simulate intelligent leadership dynamics and adaptive hunting strategies found in grey wolf packs. Unlike conventional optimizers, GWO employs a socially structured search process that enables efficient global exploration while refining the best candidate solutions through leader-follower interactions. This natural coordination allows the algorithm to effectively locate optimal hyperparameter settings for the RF, such as the number of estimators, depth of trees, and minimum samples per split, without becoming trapped in local optima. The RF-GWO configuration within the NIRFO framework thus capitalizes on this adaptive intelligence to enhance prediction robustness, stability, and convergence speed. By leveraging the wolves' collective decision-making behavior, the model achieves improved precision and generalization in wheat disease forecasting under Morocco's diverse environmental conditions, offering a distinct, behavior-driven optimization perspective compared to other algorithms in the study.

This algorithm mathematically models the dominant hierarchy (alpha, beta, delta, omega) and cooperative prey encircling behavior observed in wolf populations to efficiently explore solution spaces. The GWO algorithm mimics the leadership hierarchy and hunting mechanism of grey wolves, comprising alpha (α), beta (β), delta (δ), and omega (ω) wolves that represent candidate solutions [25]. GWO offers key advantages: balanced exploration-exploitation via adaptive parameters, prevention of premature convergence through hierarchical research, and fewer iterations than conventional methods.

The hunting (optimization) process follows these equations (Eqs (8) and (9)):

Distance to prey:

$$D = |C \cdot X_p(t) - X(t)| \quad (8)$$

Position update:

$$X(t+1) = X_p(t) - A \cdot D \quad (9)$$

where A is $2a \cdot r_1 - a$ (control parameter), C represents $2 \cdot r_2$ (randomness factor), a linearly decreases from 2 to 0, and $r_1, r_2 \in [0,1]$ are random vectors.

The fitness function evaluates classification accuracy using k -fold cross-validation (Eq. (10)):

$$\text{Fitness} = 1 / K \sum_{i=1}^K \text{Accuracy}_i(\text{RF}_\theta) \quad (10)$$

where θ represents the hyperparameters ($\text{n_estimators}, \text{max_depth}, \text{min_samples_split}$).

3.6.5. Differential Evolution (DE)

The DE, a population-based optimization algorithm [26], overcomes the limitations of conventional hyperparameter search methods by efficiently exploring large parameter spaces. We employ DE for RF optimization due to its: (I) effective global search capability, (II) minimal parameter requirements, and (III) consistent performance across diverse multiclass datasets. The incorporation of DE into the RF framework (RF-DE) was driven by DE's strong global optimization capability and its efficient handling of complex, nonlinear, and multimodal search spaces. Differential Evolution operates through simple yet powerful mechanisms of mutation, crossover, and selection, which collectively promote both exploration and exploitation during optimization. This enables DE to identify optimal hyperparameter combinations for RF,

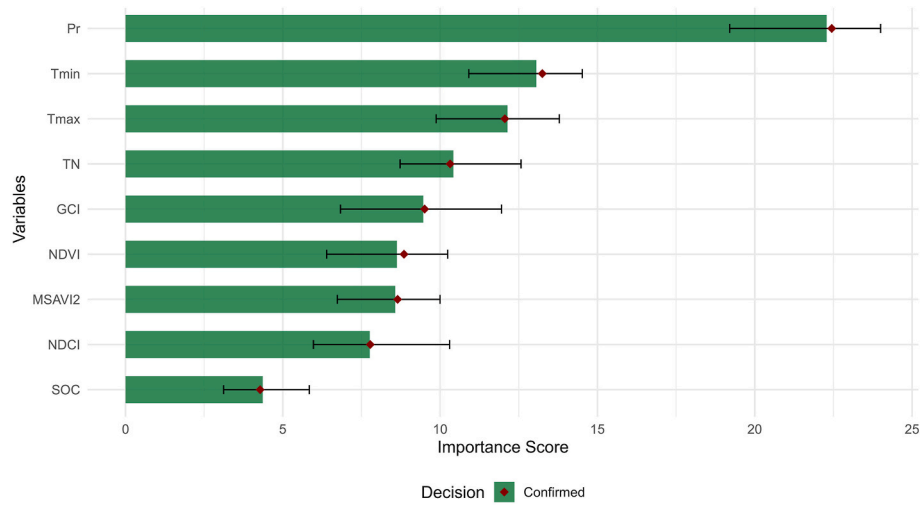


Fig. 3. Baruta results for wheat crop disease sensitivity analysis.

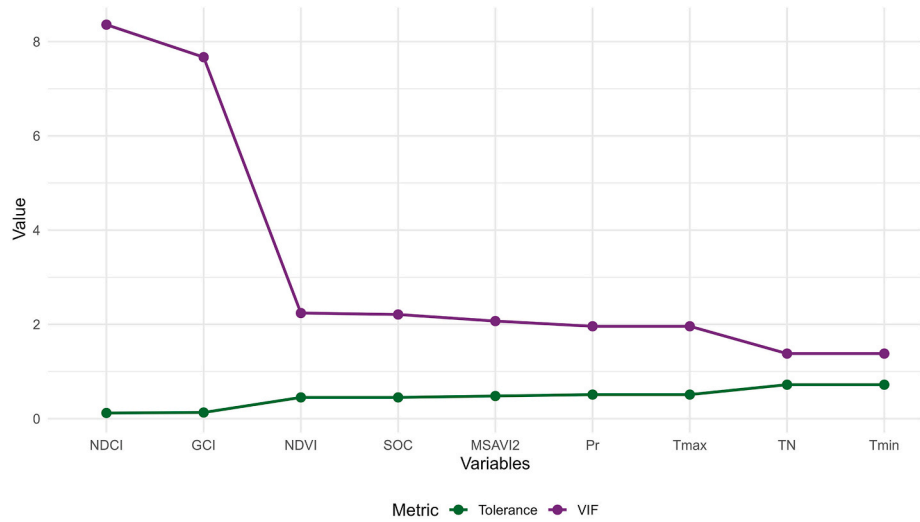


Fig. 4. VIF results for wheat crop disease predictions.

Table 2

Comparison results of the models for the Paired *t*-test and Wilcoxon signed-rank test.

Comparison	Mean Difference	t-stat	Paired t-test p-value	Wilcoxon signed-rank test p-value
RF-GA vs RF	3.19	3.18	0.034***	0.043***
RF-PSO vs RF	1.91	1.72	0.15	0.109
RF-GWO vs RF	1.91	1.72	0.15	0.088
RF-DE vs RF	2.55	2.55	0.07	0.066
RF-HBA vs RF	1.27	1.28	0.25	0.127

Note: Significance at the 1% level.

such as the number of trees, depth, and feature subsets, while minimizing overfitting and enhancing generalization. The novelty of integrating DE within the NIRFO environment lies in its ability to dynamically adapt search directions based on population diversity, leading to faster convergence toward global optima and improved model stability. By embedding DE's self-adaptive evolutionary strategy into the RF, the RF-DE model introduces a more flexible and data-driven approach to hyperparameter tuning, ultimately enhancing predictive

accuracy and robustness for wheat disease detection across Morocco's variable agro-climatic conditions.

DE optimizes RF hyperparameters by iteratively improving candidate solutions based on mutation and crossover strategies. Given a population of candidate parameter sets, DE updates each vector X_1^g in generation g as follows (Eq. (11)):

$$V_i^{g+1} = X_r^g + F \cdot (X_b^g - X_r^g) + F \cdot (X_1^g - X_2^g) \quad (11)$$

where X_r^g , X_1^g and X_2^g are randomly chosen vectors, F is the scaling factor, and V_i^{g+1} is the mutant vector. The best-performing solution updates the RF parameters, leading to an optimized model. Experimental results demonstrate that DE-tuned RF outperforms conventional tuning methods, achieving higher accuracy and efficiency in multiclass classification tasks.

3.6.6. Honey Badger Algorithm (HBA)

The HBA represents a novel metaheuristic approach that emulates the intelligent foraging strategies of honey badgers, incorporating mathematical formulations to model their distinctive digging and hunting behaviors [27]. The hybridization of the HBA with the RF model (RF-HBA) introduces a novel, behavior-driven optimization strategy. Inspired by the intelligent foraging behavior of honey badgers, HBA

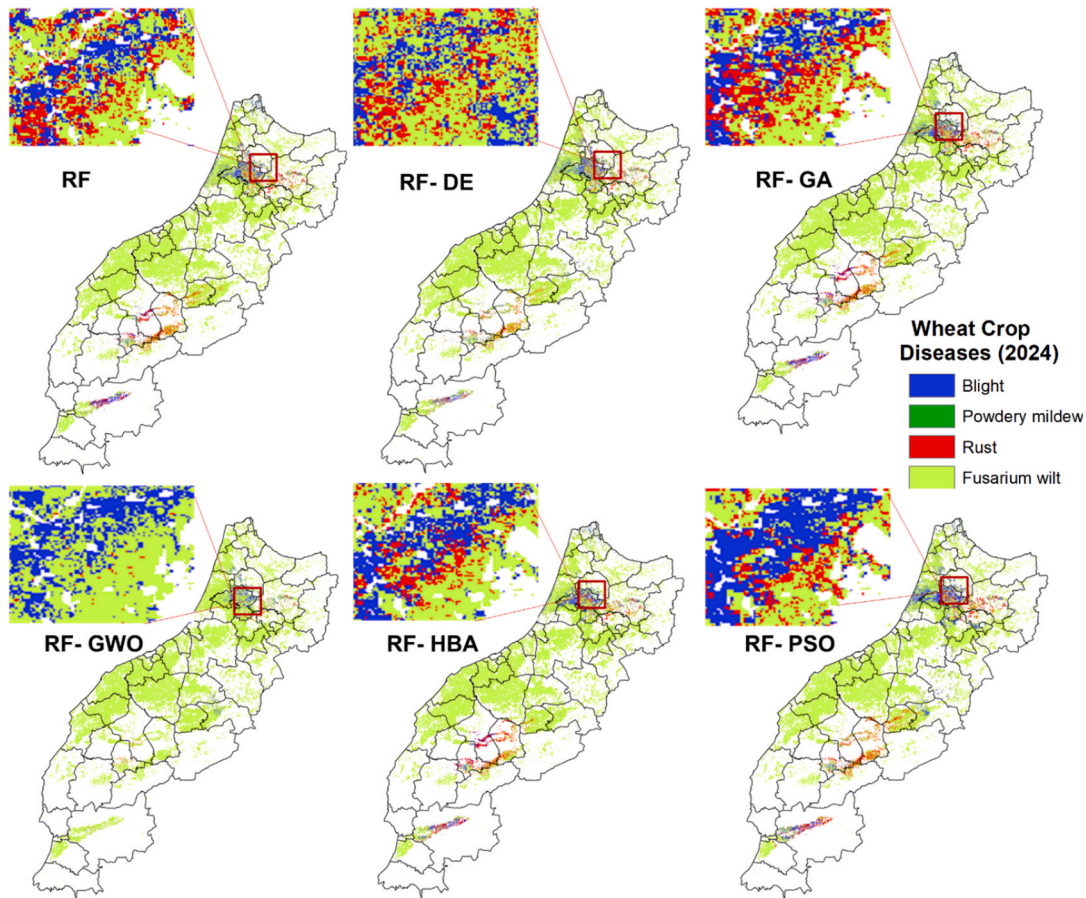


Fig. 5. Wheat diseases mapping by nature-inspired random forest optimization techniques of 2024.

alternates between intensive local search (“digging” phase) and broad global exploration (“honey-guiding” phase), enabling adaptive control over the search dynamics. This optimization technique demonstrates three key advantages: (I) a dynamic balance between exploration and exploitation through its olfactory intensity mechanism, which adjusts search intensity over iterations; (II) superior computational efficiency, achieving convergence with fewer evaluations compared to traditional optimizers such as PSO and GA; and (III) robust performance in high-dimensional search spaces, effectively preventing premature convergence. By embedding this adaptive intelligence into RF optimization, the RF-HBA model achieves higher predictive accuracy, faster convergence, and stronger generalization for wheat disease prediction under Morocco's complex agro-climatic conditions. HBA mimics the intelligent foraging behavior of honey badgers through exploration and exploitation phases. The algorithm updates candidate solutions (RF hyperparameters: $n_estimators$, max_depth , $min_samples_split$) using (Eq. (12)):

$$\begin{aligned} Intensity(I_i) &= r_2 \times (x_i - x_{\{best\}}) / 4md_i(2) \\ Density\ factor(\alpha) &= 1 - \exp(-t/t_{max}) \end{aligned} \quad (12)$$

Position update combines digging (exploitation) and honey-finding (exploration) (Eq. (13)):

$$\begin{aligned} x_{new} &= x_{prey} + F \times \beta \times 1 \times x_{prey} + F \times r_3 \times \alpha \times d_i \\ &\times |\cos(2\pi r_4) \times [1 - \cos(2\pi r_5)]| \end{aligned} \quad (13)$$

Where, $r_i \in [0, 1]$ are random numbers, β controls exploration-exploitation balance and $F \in [-1, 1]$ controls search direction.

The fitness function maximizes k-fold cross-validated accuracy (Eq. (14)):

$$Fitness = 1 / K \sum_{i=1}^K MacroF1_i(RF_{\theta}) \quad (14)$$

All the WCD model tuning parameters are summarized in **ST1**.

3.6.7. Future WCD mapping

To analyze wheat disease dynamics in Morocco, this study used an average of three CMIP6 Global Climate Models: EC-Earth3, NorESM2-LM, and MIROC6, sourced from the IPCC AR6, the NEX-GDDP-CMIP6 dataset (see details ST.2). These models, accessible through GEE, provide high-resolution (0.25°) downscaled projections for 2015–2030 under SSP2-4.5 and SSP5-8.5 pathways, representing moderate and high emission scenarios. Their proven accuracy in capturing regional climate variability and agroclimatic patterns makes them suitable for assessing localized wheat diseases. Using multiple models enhances mean ensemble robustness by reducing individual model bias and improving the reliability of temperature and precipitation projections essential for disease risk analysis (Talukder et al., 2021; [28]). Derived from CMIP6 global climate simulations with localized projections, making it well-suited for evaluating climate-driven wheat disease dynamics in Morocco [29].

3.6.8. Validation analysis

The model's effectiveness was rigorously assessed through 10-fold cross-validation, with performance evaluation focusing on WCD classification. Multiple statistical measures were employed, including: Cohen's kappa coefficient, overall accuracy (OA), producer's (PA) and user's accuracies (UA), and receiver operating characteristic (ROC) analysis. These metrics were quantitatively analyzed using the formulations presented in Eqs. (15)–(18).

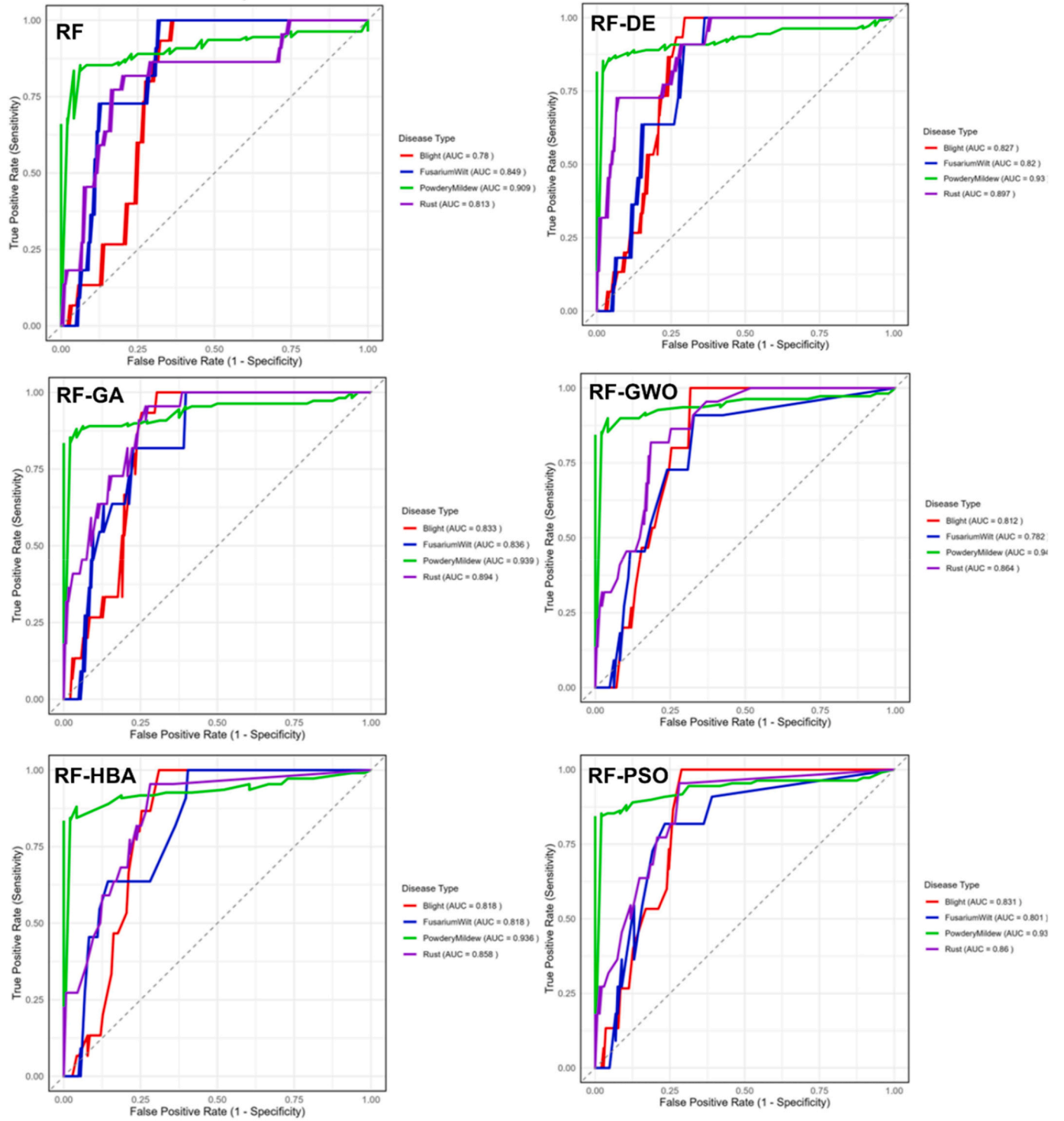


Fig. 6. Receiver operating characteristic (ROC) curves for validation of wheat crop diseases.

$$\text{Kappa coefficient} = \frac{P_{\text{obs}} - P_{\text{exp}}}{1 - P_{\text{exp}}} \quad (15)$$

$$UA(k) = \frac{S(kk)}{\sum_{p=1}^9 S(kp)} \quad (16)$$

$$PA(k) = \frac{S(kk)}{\sum_{p=1}^9 S(pk)} \quad (17)$$

$$OA = \frac{\sum_{p=1}^9 S(pp)}{\sum_{k=1}^9 \sum_{p=1}^9 S(kp)} \quad (18)$$

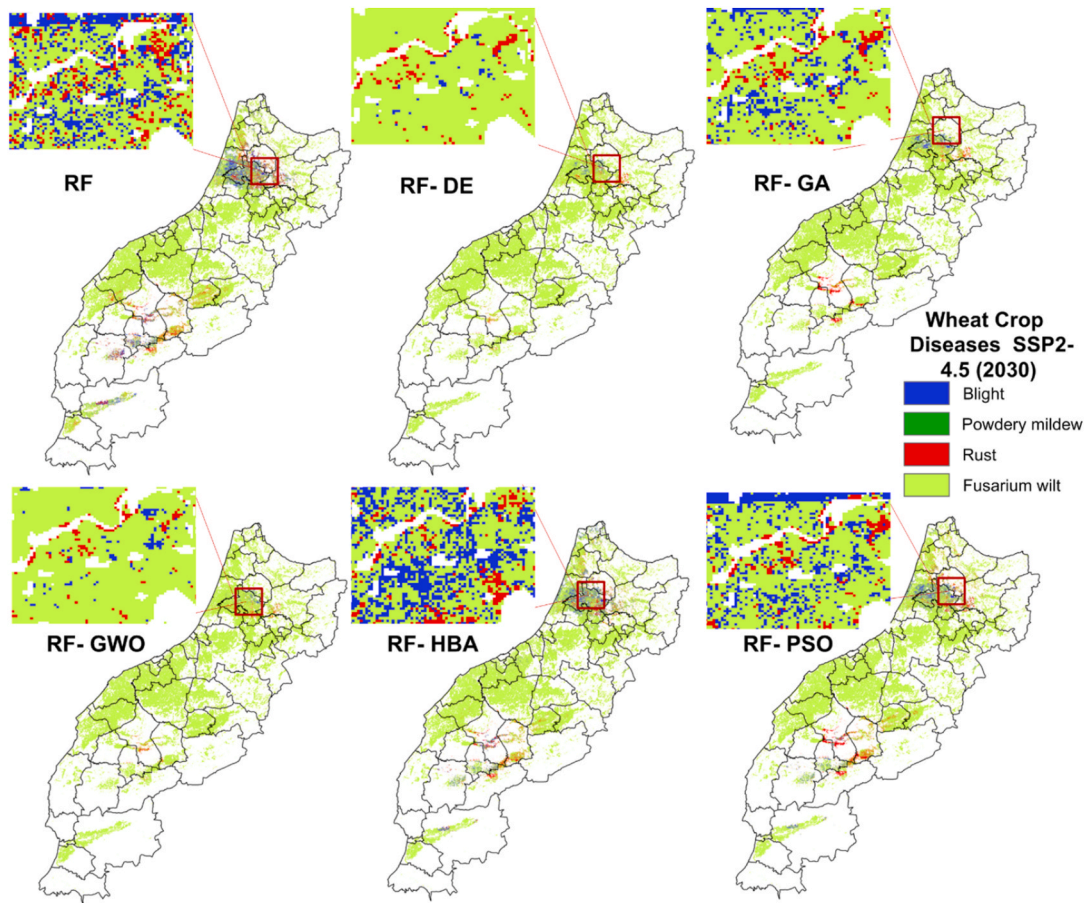


Fig. 7. Future wheat diseases mapping by nature-inspired random forest optimization techniques for 2030 under SSP2-4.5 of the integrated three GCMs model (EC-Earth3, NorESM2-LM, and MIROC6). of CMIP6.

where UA(k): User's accuracy for disease class k (commission errors), PA (k): Producer's accuracy for class k (omission errors), OA: Overall classification accuracy across all disease categories, S(kk): Correctly classified instances of disease k, S(kp): Misclassified instances of disease k assigned to class p.

To statistically assess the performance variations among the developed models (RF-GA, RF-PSO, RF-GWO, RF-DE, RF-HBA, and the baseline RF) for predicting wheat crop diseases in Morocco, both parametric and non-parametric statistical approaches were applied. Specifically, the paired *t*-test and the Wilcoxon signed-rank test were used to determine whether the observed differences in model accuracy were statistically significant. These complementary tests were selected to ensure a reliable evaluation of model performance under different data distribution conditions. The paired *t*-test, a standard parametric method, was employed to compare the mean performance differences between two related models under the null hypothesis (H_0) that their mean values are identical. When the p-value is less than 0.05, the null hypothesis is rejected, signifying a statistically significant difference in predictive performance [30]. Because model performance data often deviate from normality or violate the homogeneity of variance assumption required for the *t*-test, the non-parametric Wilcoxon signed-rank test was also applied. This test analyzes paired samples without relying on specific distributional assumptions, making it more robust for skewed or non-normally distributed data [31]. Similar to the *t*-test, the Wilcoxon test rejects the null hypothesis when $p < 0.05$ and the computed t-statistic and mean difference exceed their respective thresholds, confirming that the differences in model performance are statistically significant.

3.6.9. Sensitivity analysis

Perturbation sensitivity analysis evaluates the robustness of multi-class classification models by measuring performance changes under input perturbations [32]. Given a classifier f and input x , the sensitivity S to a perturbation δ is (Eq. (19)):

$$S(f, x, \delta) = \mathbb{E}_{\delta \sim \mathcal{D}} [||f(x + \delta) - f(x)||] \quad (19)$$

where $\mathcal{D}$ is a perturbation distribution (e.g., Gaussian noise). This quantifies the expected deviation in predictions due to small input variations. High sensitivity indicates vulnerability to adversarial attacks or noise. Studies Goodfellow et al. [33] highlight its importance in ensuring model reliability. Techniques like adversarial training can reduce sensitivity, enhancing robustness.

3.6.10. SHAP analysis

This study employs a XAI approach, specifically the SHAP method, to identify critical factors influencing prediction accuracy and assess their individual contributions at a per-prediction level. It provides interpretable feature importance for multiclass classification by leveraging cooperative game theory. Given a model f with K classes, the SHAP value $\phi_{i,k}$ for feature i and class k is computed as (Eq. (20)):

$$\phi_{i,k} = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_k(S \cup \{i\}) - f_k(S)] \quad (20)$$

where F is the feature set, S is a subset, and f_k is the model's output for class k . SHAP values quantify each feature's contribution to class predictions, enhancing model transparency [34].

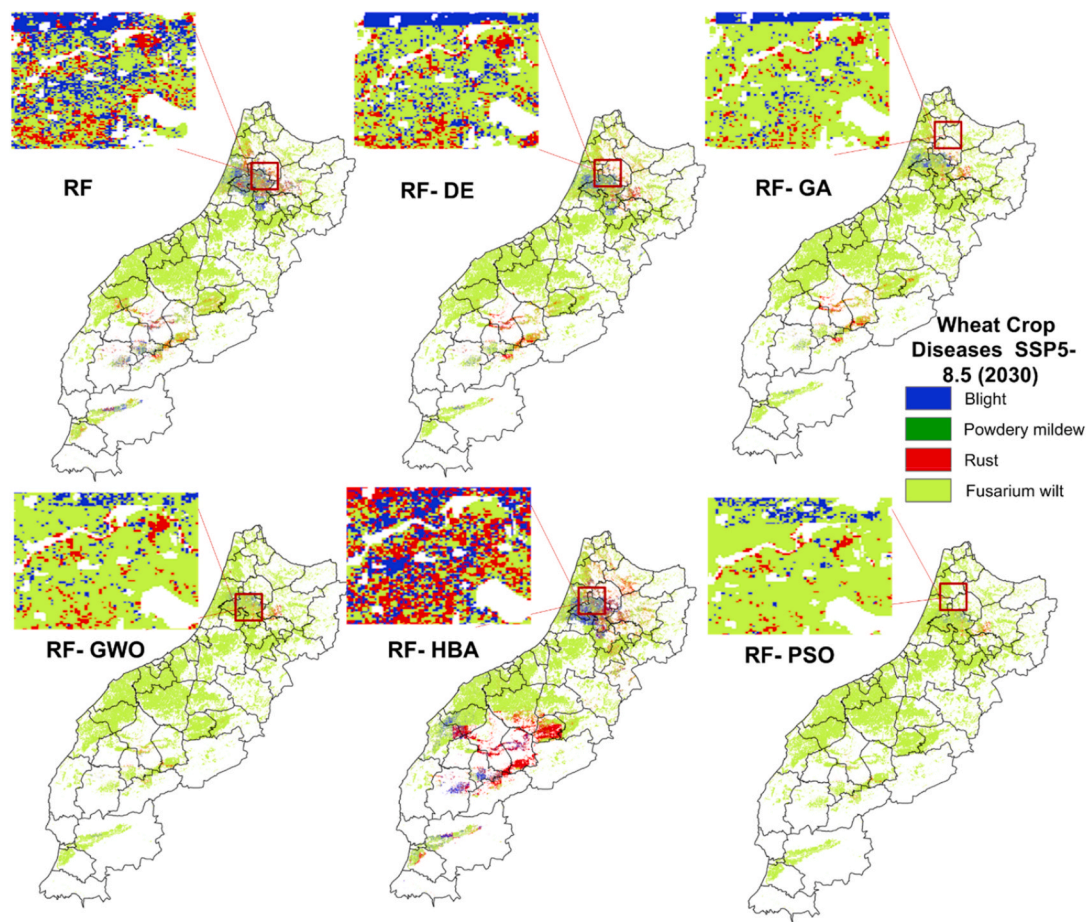


Fig. 8. Future wheat diseases mapping by nature-inspired random forest optimization techniques for 2030 under SSP5-8.5 of the integrated three GCMs model (EC-Earth3, NorESM2-LM, and MIROC6.) of CMIP6.

4. Results

This study combines explainable AI, statistical validation, and feature selection to provide a strong methodology for wheat disease prediction. To find the most pertinent predictors, the workflow starts with feature selection based on the Boruta method. Next, a multicollinearity test (like VIF) is performed to guarantee independence and model stability. Machine learning classifiers are used to assess predictive performance, and the findings show important patterns in disease susceptibility. Lastly, model decisions are interpreted using SHAP analysis, which offers useful information about feature contributions. When combined, these strategies guarantee an open, data-driven framework for managing crop diseases that strikes a balance between precision and interpretability for use in agricultural settings.

4.1. Boruta feature selection

The Boruta algorithm was used with nine variables: TN, GCI, MSAVI2, NDCI, NDVI, Pr, SOC, and temperature. All variables were confirmed regarding their impact on wheat diseases (Fig. 3). Precipitation and temperature were identified as the most influential determinants of wheat disease severity. Pr had the highest mean significance value (22.3), the median importance value (22.4), the lowest importance value (19.2), and the highest importance value (24). Conversely, SOC had the lowest mean significance value of 4.36, a median importance value of 4.28, a minimum importance value of 3.12, a maximum importance value of 5.84, and a normHits value of 0.86, indicating that this variable is less dependable.

4.2. Multicollinearity test

In the multicollinearity diagnosis, TN and Tmin had the highest Tolerance score of 0.72 and a lower VIF of 1.38. Meanwhile, TN and Tmin show minimal multicollinearity and high significance (Fig. 4). On the other hand, the NDCI has the lowest tolerance value of 0.12 and the largest VIF of 8.36. It was also revealed that the GCI has a tolerance of 0.13 and a VIF of 7.67. As a result, our findings imply that both NDCI and GCI have substantial multicollinearity concerns, posing considerable challenges for wheat disease investigation.

4.3. Model performance

The evaluation of six RF based hybrid ensemble models RF-GA, RF-PSO, RF-GWO, RF-DE, RF-HBA, and baseline RF revealed distinct variations in classification accuracy across the four major wheat diseases: Blight, Fusarium wilt, Powdery mildew, and Rust (see ST.3). The RF-GA model achieved the highest performance with an overall accuracy of 80.89%, Kappa coefficient of 0.626, and strong user's and producer's accuracies across all diseases (Blight: 46.67%, Fusarium wilt: 50%, Powdery mildew: 95.10%, Rust: 61.54%). The RF-PSO model followed with an overall accuracy of 79.62% and Kappa of 0.591, showing reliable performance particularly for Powdery mildew (91.43%) and Fusarium wilt (66.67%). Similarly, RF-GWO achieved 79.62% accuracy and Kappa of 0.588, maintaining consistency across classes with notably high accuracy for Powdery mildew (93.39%) and Rust (52%). The RF-DE model performed moderately well with an overall accuracy of 80.25% and Kappa of 0.611, showing improved accuracy for Powdery mildew (95.15%) and Rust (60.87%). Meanwhile, RF-HBA attained

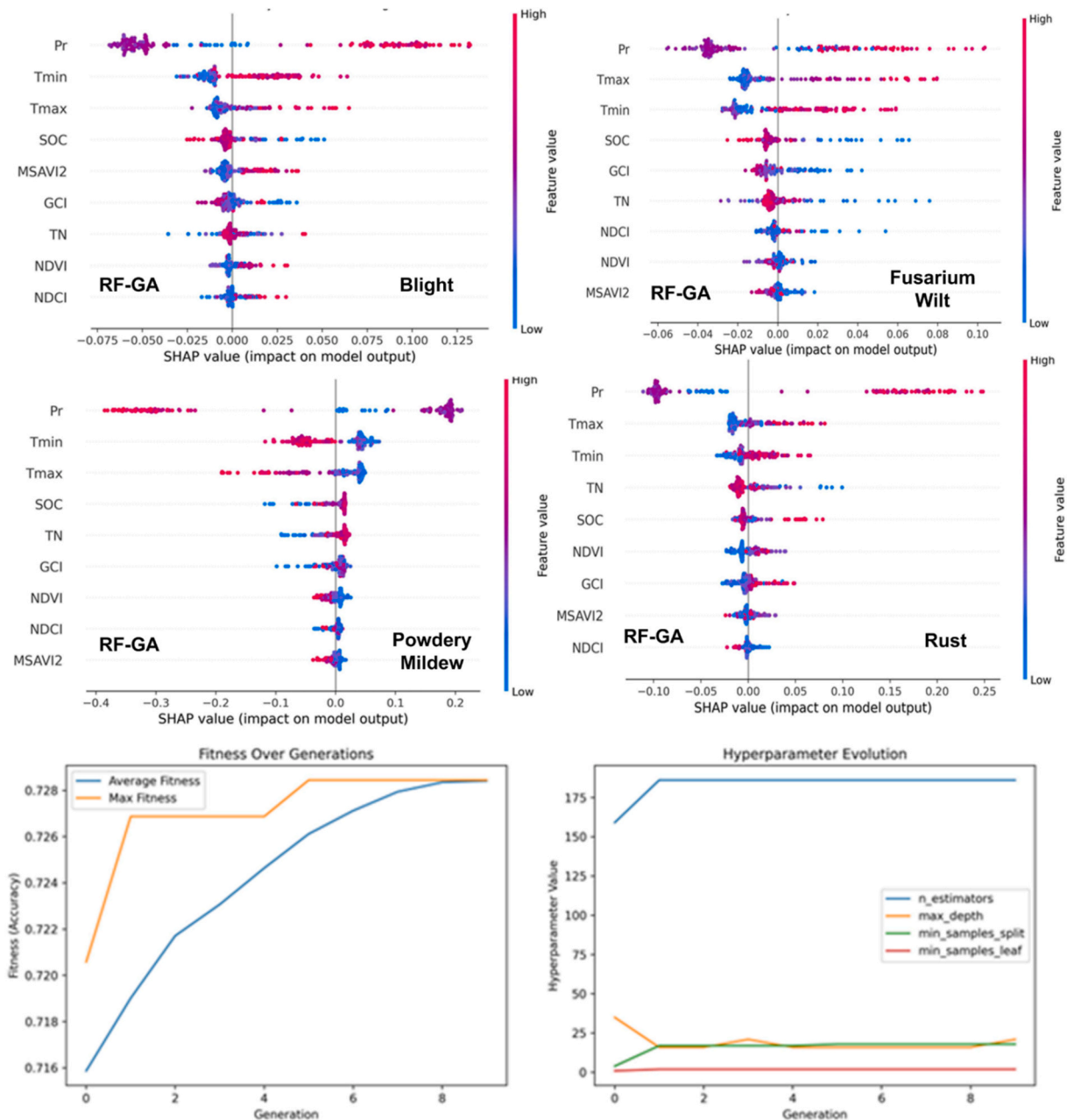


Fig. 9. SHAP summary representation of best ML model of RF-GA for four types of wheat diseases and showing of hyperparameter evolution of the model.

78.98% overall accuracy and Kappa of 0.589, maintaining balanced classification across disease types. The baseline RF model, though reliable, recorded the lowest overall accuracy (77.71%) and Kappa (0.571), particularly struggling with Blight detection (38.46%) but performing well for Powdery mildew (95.96%).

In the wheat crop disease classification analysis, the RF-GA hybrid model demonstrated superior performance compared to all other RF-based variants. It achieved the highest overall accuracy, Kappa coefficient, and AUC values, indicating strong predictive capability and robustness.

To statistically evaluate the performance differences between the hybrid RF models and the baseline RF model for Morocco wheat crop disease prediction, both the paired *t*-test and the Wilcoxon signed-rank test were employed. The results (Table 2) revealed that the RF-GA model demonstrated a statistically significant improvement over the baseline RF, with a mean difference of 3.19 and *p*-values of 0.034 and 0.043 for the paired *t*-test and Wilcoxon test, respectively. This indicates

that the integration of the Genetic Algorithm (GA) effectively enhanced model accuracy. In contrast, the RF-PSO, RF-GWO, RF-DE, and RF-HBA models exhibited positive mean differences ranging from 1.27 to 2.55, but their *p*-values exceeded 0.05 in both tests, suggesting that these improvements were not statistically significant. Notably, the RF-DE and RF-GWO hybrids showed near-significant enhancements (*p* between 0.06 and 0.09), indicating potential but inconclusive gains in predictive performance. Overall, the statistical validation confirmed that among all hybrid models, RF-GA achieved the most reliable and significant improvement over the conventional RF model, establishing it as the most effective approach for wheat disease detection in Morocco.

4.4. Analysis of wheat disease predictions

This section examines the distribution of four types of Moroccan wheat crop diseases, namely, blight, powdery mildew, rust, and fusarium wilt, in various regions, particularly in the north, such as Douar

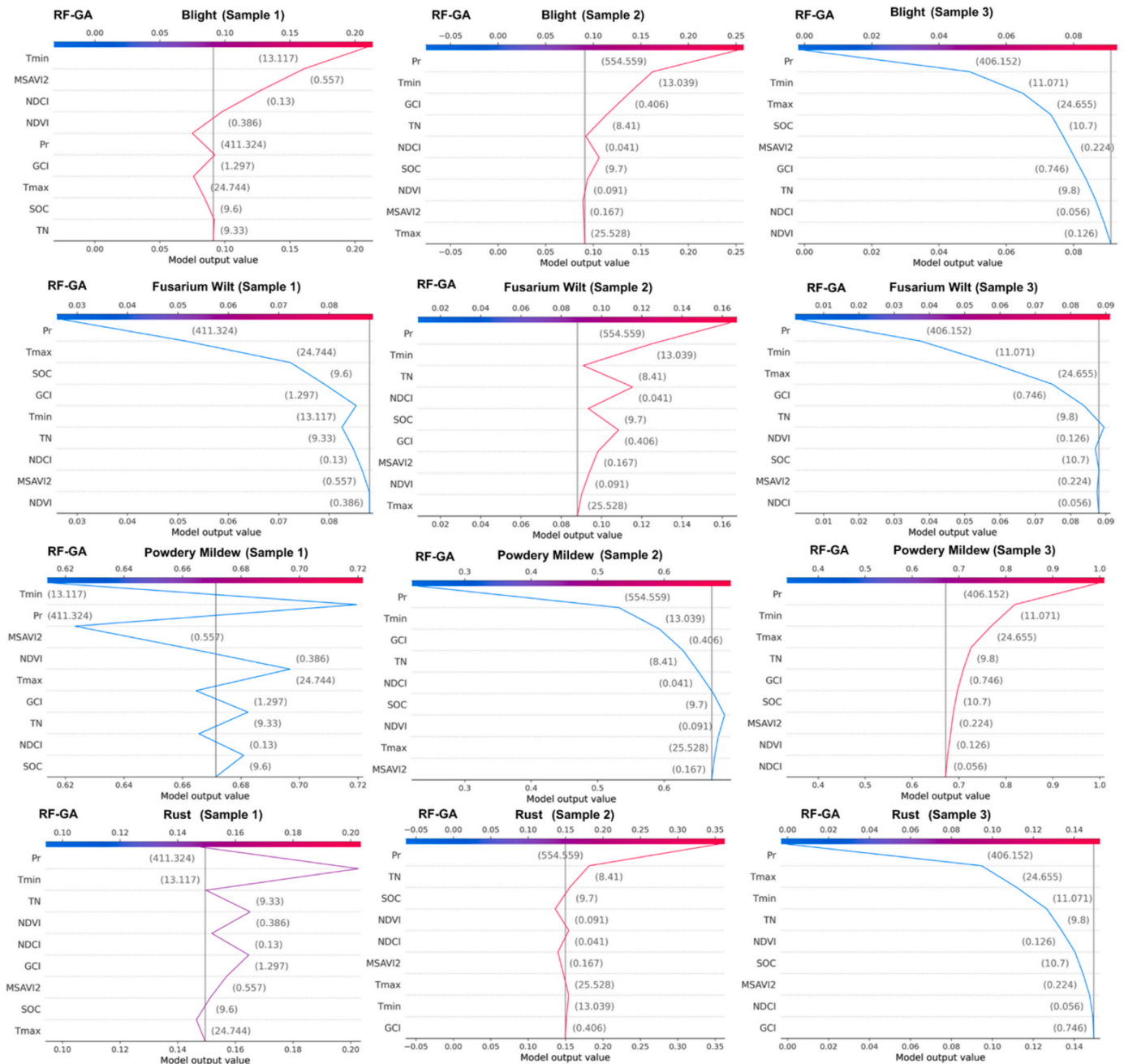


Fig. 10. Sample-wise decision plotting for the best ML model of RF-GA by four types of wheat diseases.

Khreb Ennaine, Dbichat Ghouazi, Quled Allal Ghouazi, and Douar Zehiryene regions, with some found in the southern region during 2024 (Fig. 5). The study uses RF models to compare the performance of various RF-based approaches, including RF-DE, RF-GA, RF-GWO, RF-HBA, and RF-PSO. Blight disease is primarily found in the north, with occasional cases occurring in the south, and is caused by the fungal diseases *Fusarium culmorum* and *Fusarium graminearum* in warm and humid environments. Powdery mildew wheat crop disease is quite rare in the southern and eastern regions, according to RF models. Notably, RF-DE, RF-GA, RF-HBA, and RF-PSO indicated the presence of Rust in the northern, central, and southern regions, although RF-GWO identified just a small amount in the northern zone, caused by several species of the Puccinia fungus.

From March to May, the most favorable weather conditions for the occurrence of rust epidemics included relative humidity of more than 90%, rainfall of 0.1 mm, and temperatures ranging between 8 °C and

16 °C lasting at least four consecutive hours 5% to 10% of the time [3]. In contrast, Fusarium Wilt is widespread and caused by fungi from the genera *Fusarium* and *Microdochium*. Fusarium wilt is a soil-borne disease that spreads by seeds and affects all parts of the plant, including the roots and ears. Among the RF-based models, RF-GA performs exceptionally well. Nonetheless, according to the percentage findings of RF-GA wheat crop diseases, the current scenario is blight (4.39%), powdery mildew (4.44%), fusarium wilt (71.50%), and rust (19.50%).

ROC curves were used to examine the sensitivity and specificity of four types of wheat crop diseases- Blight, Fusarium Wilt, Powdery Mildew, and Rust, with estimated Area under the curve (AUC) values (Fig. 6). The curves juxtapose the performances of nature-inspired RF models, including RF-DE, RF-GA, RF-GWO, RF-HBA, and RF-PSO, to diagnose these diseases. The x-axis represents the False Positive Rate (1-Specificity) and the y-axis the True Positive Rate (Sensitivity). The diagonal line from the bottom left to the top right corner represents a

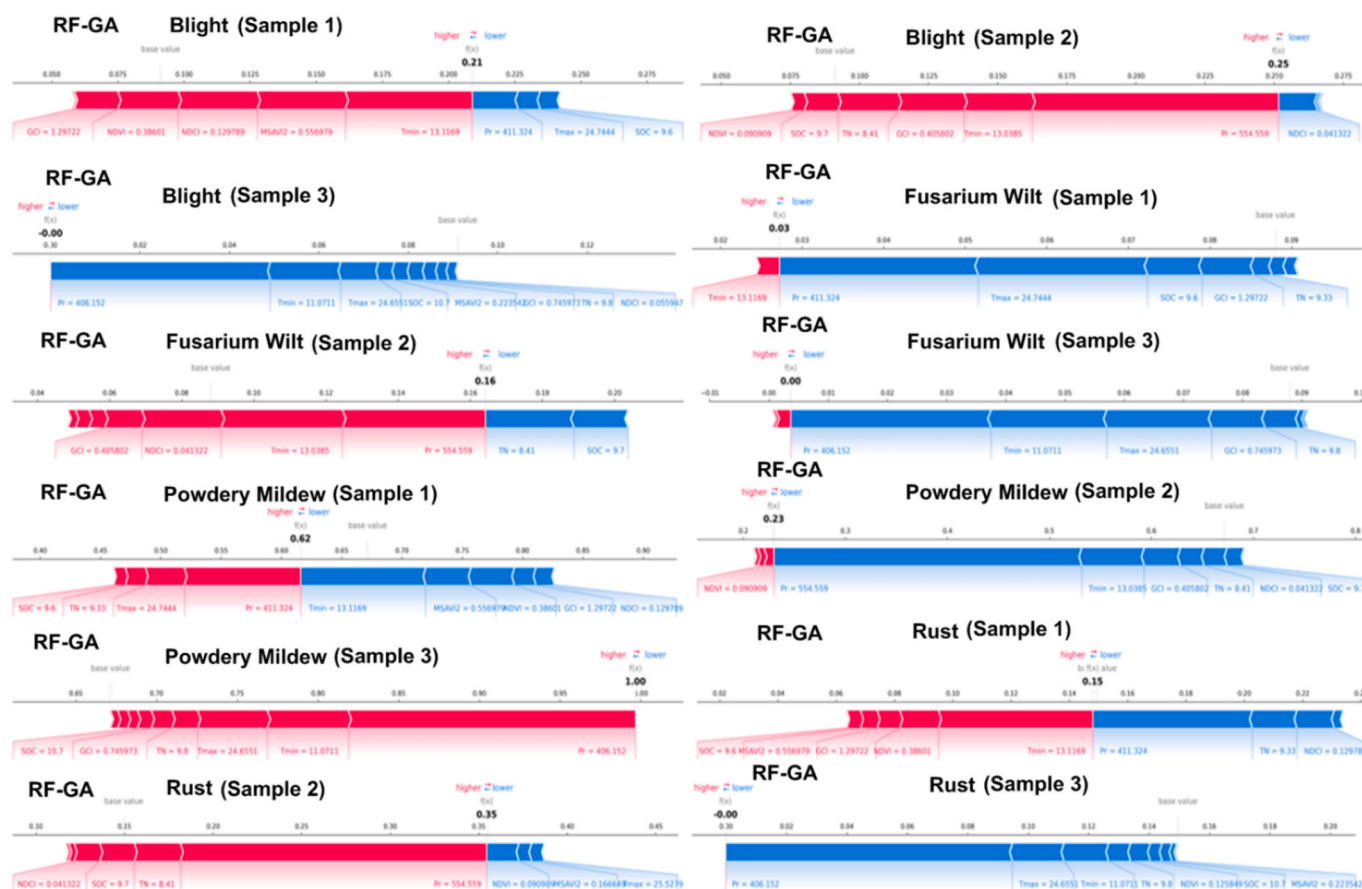


Fig. 11. Sample-wise force plotting for the best ML model of RF-GA by four types of wheat diseases.

random classifier, where $TPR = FPR$. AUC values range from 0 to 1, where a value of 1 shows that the classifier is good and 0.5 shows a random performance. The result observed that all the diseases have an AUC of more than 0.5, which implies that they are good classifiers. Among them, Powdery Mildew had the highest AUC, indicating that it was better classified than the remaining diseases. Also, RF-GA produced the highest AUC values for Blight ($AUC = 0.833$) and Fusarium wilt ($AUC = 0.836$) diseases, whereas RF-GWO represents the best AUC for Powdery mildew disease ($AUC = 0.946$), and the RF-DE model recorded the highest AUC for Rust ($AUC = 0.897$). On average, RF-GA is the best performer.

4.4.1. Analysis of future wheat disease predictions

4.4.1.1. Analysis of future SSP2-4.5 predictions. This image depicts the prediction of numerous wheat crop diseases in the research regions, as well as a performance comparison of nature-inspired random forest optimization (NIRFO) models (Fig. 7). The RF models employed include RF-DE, RF-GA, RF-GWO, RF-HBA, and RF-PSO, which have been used to diagnose and categorize wheat crop diseases such as blight, powdery mildew, rust, and fusarium wilt for the year 2030 under the SSP2-4.5 scenario. The findings show that the RF-HBA, RF-PSO, and RF-GA models can efficiently recognize the dense and detailed Blight disease in the northern regions, which is caused by a variety of pathogens, including fungi and bacteria that impact the leaves. Furthermore, all RF models detect very little powdery mildew. Rust is also commonly diagnosed in the northern and middle-southern regions, which can cause environmental stress in infected wheat crops due to excessive humidity, drought, heat stress, and low temperatures. As a result, cold and moist circumstances can cause Rust to grow on leaves, stems, and glumes, spreading via wind and causing widespread infestations.

Fusarium wilt is widely distributed throughout the region, as evidenced by all models, notably the RF-DE model. Fusarium wilt is the most common of the four diseases in the affected areas, caused by attacks from *Fusarium* species, a fungus that infects the plant's roots. These wheat crop diseases can result in a significant reduction in wheat production.

4.4.1.2. Analysis of future SSP5-8.5 predictions. The spatial study of the wheat crop disease map displays numerous illnesses, notably Blight, Powdery mildew, Rust, and Fusarium wilt, under the SSP5-8.5 scenario anticipated for 2030, which were discovered by five distinct NIRFO algorithms: RF-DE, RF-GA, RF-GWO, RF-HBA, and RF-PSO (Fig. 8). Blight disease of wheat crops was discovered to be more widespread in northern Morocco, which is more humid than the center. This disease is primarily caused by various fungal diseases that develop in cold and humid settings, leading to leaf damage and lower yield. Some of the blight illnesses identified by the RF-HBA model also occur in the central, western, and southern regions. In contrast, only a minor amount of Powdery mildew illness was found in the research area, which was produced by excessive rainfall and a coastal or semi-continental climate. Rust disease, particularly stem rust, yellow rust, and leaf rust, was caused by the fungus *Puccinia graminis* f.sp. tritici (Pgt), *Puccinia striiformis*, and the pathogen *Puccinia triticina*, which were identified in the study region's northern, central, and western areas.

Climate change and rising temperatures have a substantial impact on the epidemiology of rust disease. Fusarium wilt is also common in the study area, and it is associated with fungi from the *Fusarium* and *Microdochium* genera. Fusarium wilt infections in the research areas are caused by three species: *Fusarium Culmorum*, *Fusariumavenaceum*, and the subspecies *Microdochiumnivalemajus*. As a result, it may be concluded that Fusarium wilt is widespread throughout the region.

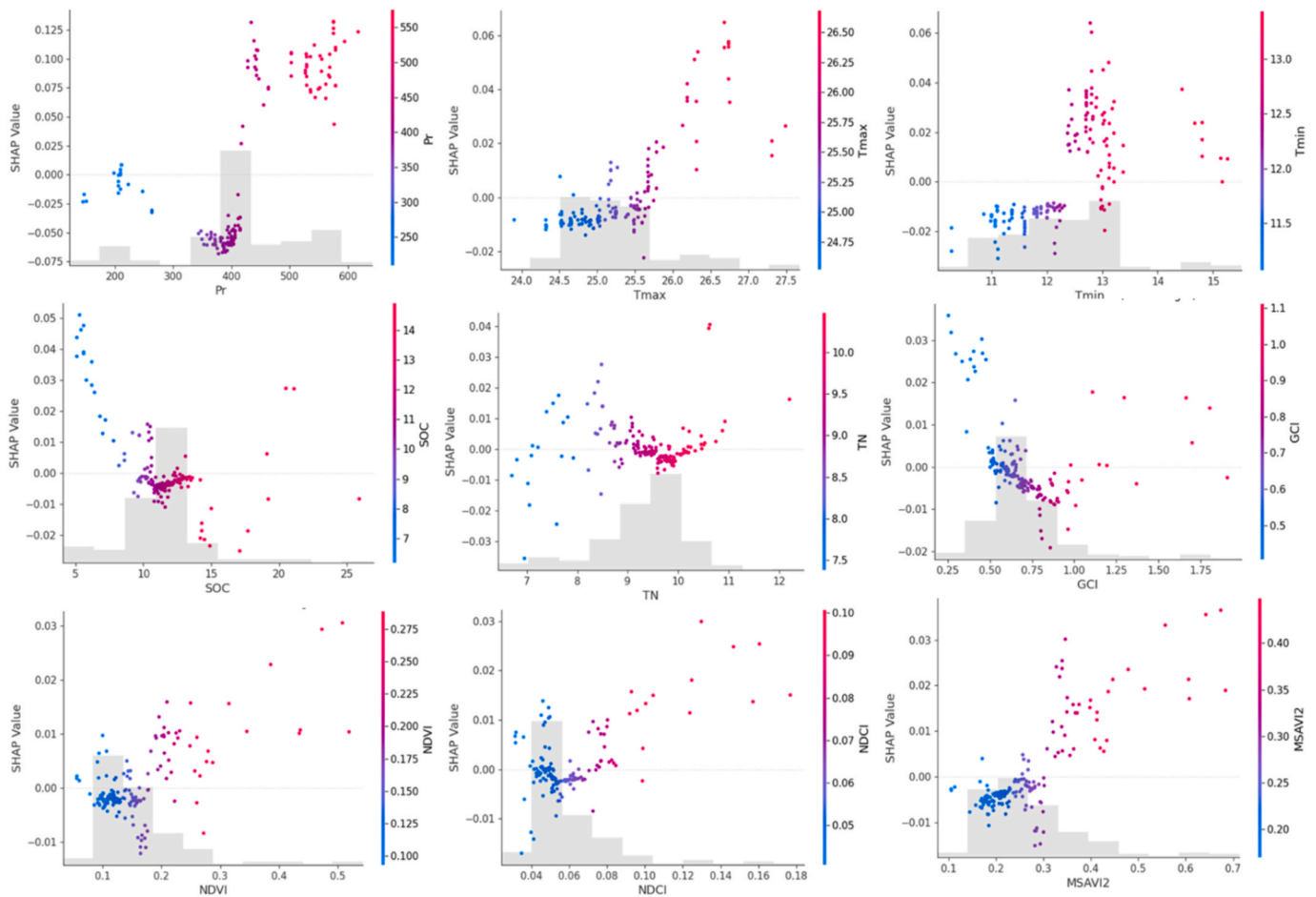


Fig. 12. Dependency graphs showing the factors that affect *Blight* wheat disease.

4.5. SHAP analysis

4.5.1. SHAP summary

The Shapley additive explanation (SHAP) summary graphic of the RF-GA model for four wheat crop diseases blight, Fusarium Wilt, Powdery Mildew, and Rust shows the SHAP values and the relevance of each feature concerning model predictions. Features are arranged vertically, from most to least important. A greater range of SHAP values indicates more differentiation in model output, and hence, the feature is more significant. This figure also shows how each feature influences the model's predictions. The color gradient is pink for high values and blue for low values. Pr and temperature are shown to have a positive impact in the sections on Blight, Fusarium Wilt, and Rust, showing that these variables increase the likelihood of these diseases occurring. In the Powdery Mildew section, however, Pr and temperature have a mostly negative influence, implying a major negative effect on the transmission of the disease (Fig. 9).

4.5.2. Decision plot

Using the RF-GA model, it offers SHAP decision plots for diseases rust, fusarium wilt, powdery mildew, and blight, illustrating how feature importance differs for each (Fig. 10). It emphasizes how SHAP's interpretability helps make it easier to understand the factors influencing the model's decisions for each of the wheat crop illnesses. Three samples are used to represent each disease. Having a strong positive SHAP value, the minimum temperature is the most important criterion in the selection of Blight sample 1. Pr has a modest positive influence on Blight sample 3, but it has a large positive effect on Blight sample 2, according to the SHAP plot. These results imply that both high and low rainfall levels, as

well as higher minimum temperatures, promote the spread of blight disease. On the other hand, Tmin has a greater influence than Pr on Blight samples 2 and 3, whereas MSAVI2 has a smaller effect than Tmin on Blight sample 1. The growth of Fusarium wilt is often supported by both high and low rainfall, as seen by sample 2's strong positive SHAP value and samples 1 and 3's low positive influence from precipitation. Tmin has a low positive influence on Powdery Mildew sample 1, just like Blight, and samples 2 and 3 have low and extremely high precipitation levels, which can be detrimental, as indicated by the positive SHAP values. For Rust wheat crop disease, precipitation generally exerts a positive influence on the model's selection across samples 1, 2, and 3, with sample 1 showing moderate levels, sample 2 indicating high levels, and sample 3 revealing low levels of precipitation, all of which could positively affect Rust growth. Ultimately, the rising minimum temperature and precipitation are crucial factors in the development of these wheat crop diseases.

4.5.3. Force plot

The force plots illustrate the Shapley values of various wheat crop diseases, including rust, fusarium wilt, powdery mildew, and blight, on various samples, are displayed in Fig. 11. By demonstrating how the features influenced the model forecast for a particular observation, this improves the interpretability of an RF-GA model prediction. This plot is highly helpful for error analysis and in-depth case comprehension. Prediction is increased by feature values in pink. The magnitude of the characteristic's effect is indicated by the bar's size. Blue feature values result in a lower prediction. The difference between the model's forecast and the baseline is justified by the sum of all SHAP values. With low confidence, the force plot shows that the anticipated probabilities of

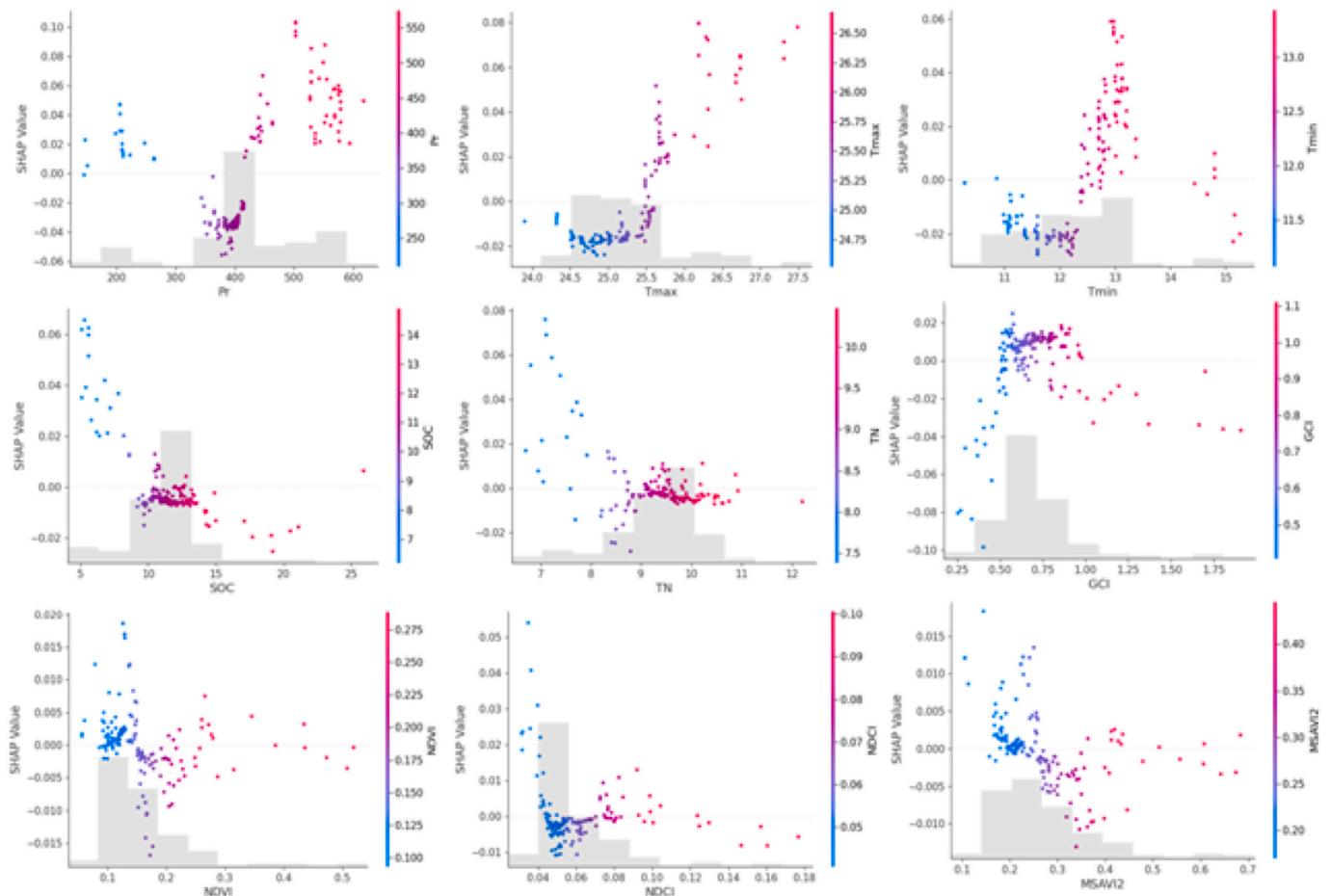


Fig. 13. Dependency graphs showing the factors that affect *Fusarium Wilt* wheat disease.

Blight samples 1 and 2 are 0.21 and 0.25, respectively. With SHAP values of 13.12 and 554.56, respectively, the Tmax and Pr are important variables that positively impact Blight samples 1 and 2. A long arrow pointing to the left indicates that instances with greater values of these parameters often have a higher possibility of Blight samples 1 and 2. On the other hand, the Pr has a positive effect on the likelihood; examples with higher Pr values have a lower probability of Blight sample 3, as shown by its SHAP value of 406.15, which is shown by a long arrow pointing to the right. Furthermore, there is a lack of confidence as the anticipated chance of Blight sample 3 is -0.00 .

Comparably, the projected probabilities $f(x)$ for Fusarium Wilt samples 1 and 3 are 0.03 and 0.00, respectively, indicating a very low degree of confidence for sample 1 and a total lack of confidence for sample 3. A long arrow pointing to the right indicates that the feature Pr pushes the model's output toward lower values, while a long arrow pointing to the left indicates that Pr pushes the model's output toward higher values for Fusarium Wilt sample 2, with a predicted probability of 0.16. For Powdery Mildew sample 1, the calculated probability $f(x)$ is 0.62, indicating a moderate degree of confidence. The most significant factors, both favorably and negatively, were the Pr and Tmin. Pr improves the model's prediction in comparison to the baseline value for Powdery Mildew sample 3, however it has a decreasing influence on the model's prediction for Powdery Mildew sample 2. For Powdery Mildew samples 2 and 3, the expected probability $f(x)$ is 0.23 and 1.00, corresponding to low and high confidence, respectively. Tmin and Pr are the main contributors to the likelihood of 0.15 based on Rust sample 1. The model's prediction is being pushed higher by Tmin and lowered by Pr. Compared to Rust sample 3, which Pr had less predictive control, Rust sample 2 had stronger predictive control due to its larger feature value.

It was intriguing to note that Rust samples 2 and 3 have respective probabilities of 0.35 and -0.0 .

4.5.4. Scatter plot

The SHAP dependence graphs suggested in this study show the correlation between a variable and its SHAP value related to the diseases rust, fusarium wilt, powdery mildew, and blight. Every subplot represents a distinct feature, with the color on the right denoting the feature value's size (Figs. 12–15). According to the analysis results, the SHAP value tends to grow when Pr, Tmax, and Tmin rise. This suggests that temperature and precipitation are important factors that affect the health of wheat crops, increasing the possibility of these diseases, aside from powdery mildew occurring. But in other cases, increasing Tmin can cause the SHAP value linked to Blight disease to drop. Furthermore, Tmin and its SHAP value have a detrimental interaction with regard to Fusarium Wilt disease. According to early observations, high SHAP values are correlated with low SOC levels; however, when SOC levels increase, the SHAP value eventually rises as well. This suggests that high SOC levels could increase the occurrence of several fungal infections, including those that cause rust, fusarium wilt, and blight. The Powdery Mildew SOC graph study showed that although the SHAP value rises as SOC levels rise, it finally falls after hitting a particular threshold. Furthermore, a decrease in SHAP values is observed for Fusarium wilt and Rust disease, but an increase in TN is associated with a rise in SHAP values for Powdery Mildew. This suggests that diseases like powdery mildew that affect wheat crops could be exacerbated by excessive nitrogen fertilizer application. Although an increase in GCI only causes a decrease in SHAP values for Fusarium Wilt and Powdery Mildew, the results show a favorable influence on Rust illnesses and a modest effect

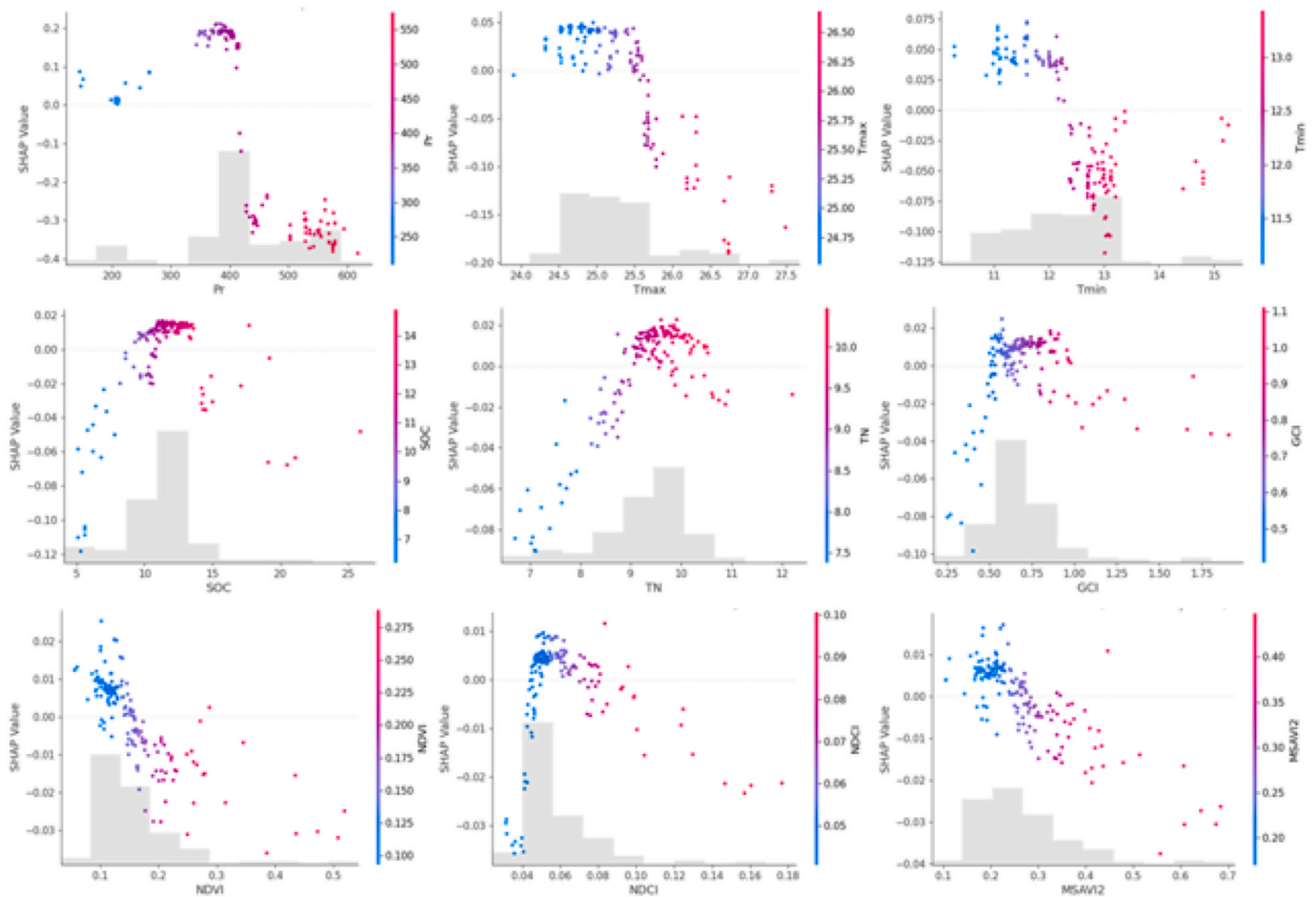


Fig. 14. Dependency graphs showing the factors that affect *Powdery Mildew* wheat disease.

on model projections of the GCI graph for Blight. Additionally, GCI has been shown to be useful in determining the severity of wheat rust disease. It is evident that whereas NDVI, NDCI, and MSAVI2 show negative SHAP values for Powdery Mildew, they show positive SHAP values for Blight. The SHAP values of Fusarium Wilt were near zero when NDVI and MSAVI2 increased, whereas the SHAP value decreased as NDCI grew. Similarly, NDVI, NDCI, and MSAVI2 showed notable oscillations and stayed near zero in the setting of Rust disease.

4.5.5. Partial dependence plot

Several illustrations with partial dependence plots that show how different biological or environmental factors affect the probability or severity of distinct wheat diseases are presented in the text that is provided (Figs. 16–19). The findings specifically look at factors that affect wheat crops' susceptibility to rust, fusarium wilt, powdery mildew, and blight. These illustrations provide insights into the dynamics of the disease and its contributing components by illuminating the relationship between changes in individual parameters and the anticipated outcome for each specific ailment. As a result, the data offers an illustration of the main factors influencing these various wheat diseases. The results of the partial dependence plots showed the main variables that affect the wheat crop diseases of rust, fusarium wilt, powdery mildew, and blight, and they had a significant impact on the prediction at different levels. Pr, Tmax, Tmin, SOC, TN, GCI, NDVI, NDCI, and MSAVI2 are among the variables that have a significant impact on the projected diseases of wheat harvests. The two most important factors that have a beneficial impact on Blight, Fusarium Wilt, and Rust diseases, and a negative impact on Powdery Mildew disease are precipitation and temperature. This suggests that rust, fusarium wilt, and blight diseases are

exacerbated by increasing temperatures and precipitation. Powdery mildew thrives best in temperatures between 13 and 30 °C, and these conditions are often the worst in areas with little to no precipitation [35]. Additionally, there is an inverse relationship between growing soil organic matter and Fusarium wilt and blight, suggesting that a high concentration of soil organic matter can lessen these diseases. A certain amount of SOC seems to be sufficient to improve Powdery Mildew, but after a certain point, additional SOC increases have a detrimental effect on Powdery Mildew since too much SOC may reduce its prevalence. Conversely, an increase in SOC has an inverse relationship with Rust as well, but it starts to have a positive effect beyond a certain point because too much SOC might cause this disease. It was then observed that, with the exception of the greatest incidence, blight disease similarly rose with N application. The disease known as wheat blight declined as GCI rose. Similarly, GCI and TN have a favorable effect on powdery mildew but a negative effect on fusarium wilt. Rust reduction is associated with the use of TN. On the other hand, GCI has a negative correlation with Rust, although it starts to have a positive effect on Rust after hitting a certain threshold (about 0.5). The development of wheat rust disease may be caused by a significant amount of green chlorophyll. Additionally, a positive increase in disease prevalence was correlated with a noteworthy rise in blight as measured by NDVI, NDCI, and MSAVI2. Up until a certain time, Fusarium Wilt's NDVI, NDCI, and MSAVI2 exhibited a fall; after that, they started to rise favorably. Powdery mildew was negatively impacted by the NDVI and MSAVI2, although the NDCI had a beneficial relationship that eventually reduced. Rust demonstrated that after a certain point, the NDVI and MSAVI2 displayed an ascending tendency before declining. In addition, Rust decreased as NDCI increased.

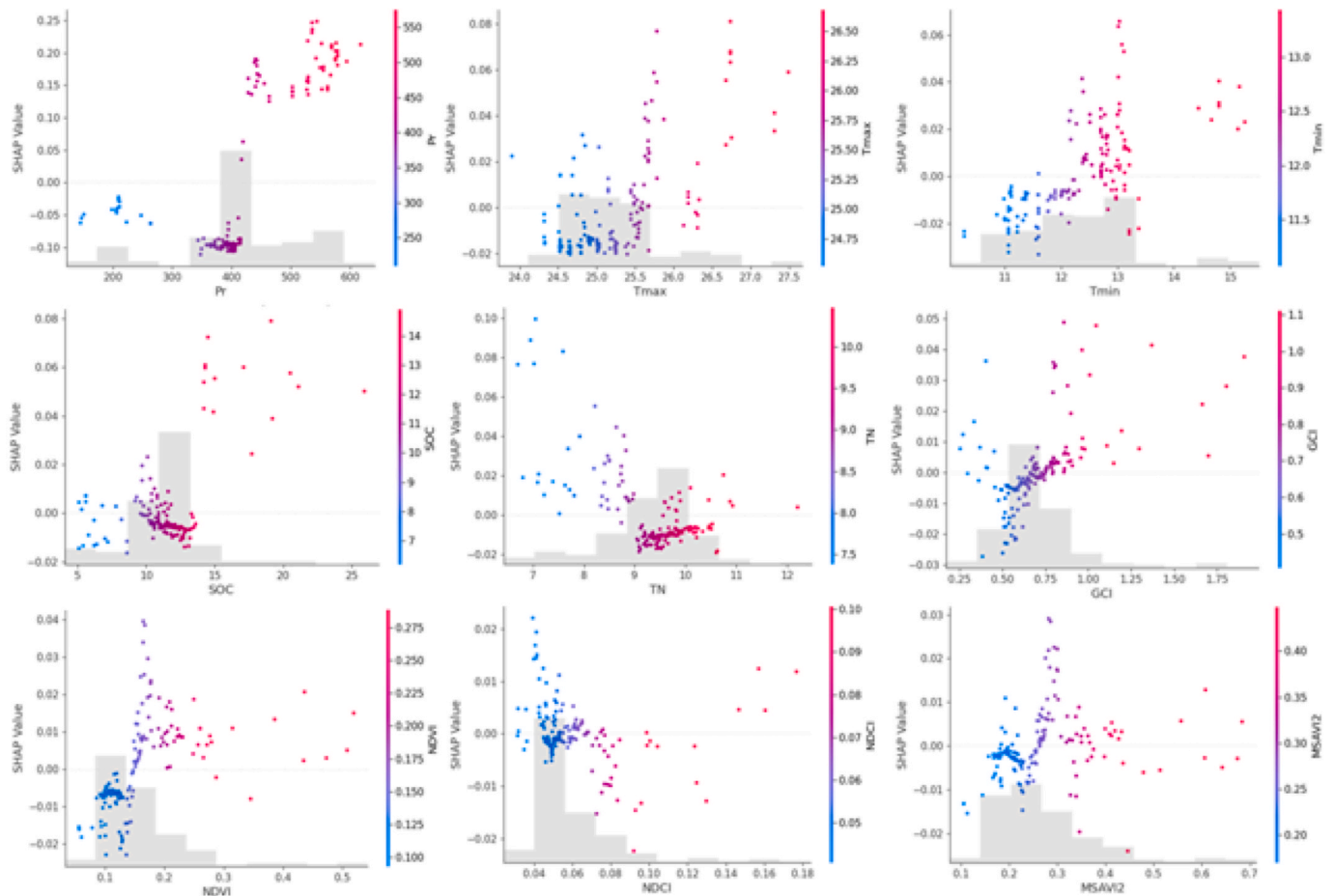


Fig. 15. Dependency graphs showing the factors that affect *Rust* wheat disease.

4.6. Feature sensitivity analysis

This section used a bar graph to calculate the average sensitivity of several variables, such as GCI, MSAVI, NDCI, NDVI, Pr, SOC, Tmax, Tmin, and TN, for each disease type, such as blight, Fusarium wilt, powdery mildew, and rust (Fig. 20). The GCI bar chart shows a strong positive sensitivity to rust, meaning that GCI is particularly good at detecting or modeling rust. The MSAVI bar chart shows a strong positive sensitivity to blight and a moderate negative sensitivity to powdery mildew disease. Similarly, the bar graphs of NDCI, NDVI, Pr, and Tmax show a strong negative correlation with Powdery Mildew. This shows that NDCI, NDVI, Pr, and Tmax are particularly useful in diagnosing Powdery Mildew disease. NDCI, NDVI, Pr, and Tmax values decrease in the presence of this disease, which is more common in arid areas. SOC and Tmin demonstrate increased positive susceptibility to Powdery Mildew. Simply said, Powdery Mildew thrives in locations with more organic carbon in the soil and greater minimum temperatures, as it appears to favor warmer nights. Finally, the TN section is extremely sensitive to Blight in a negative way, since TN levels are lower in places afflicted by Blight, implying that Blight is more common in soils with low nitrogen content.

5. Discussions

Regional prediction of crop diseases using climatological data is essential for improving disease management strategies and maximizing crop productivity [12,36]. In order to treat three fungal diseases (foliar and ear) on durum wheat, the study sought to evaluate the efficacy of this method that promotes sustainability in agriculture [4]. Our research

aims to improve wheat output by identifying several illnesses that impact wheat crops in 2024 and forecasting into 2030. In the context of CMIP6, Jung et al. [37] investigated the SSP2-4.5 and SSP5-8.5 scenarios. According to Nouri and Veysi [38], an ensemble of ten CMIP6 climate models downscaled using the NEX-GDDP-CMIP6 dataset was applied to evaluate crop-related water and food security. Their projections show that ETo may rise by around 4.8% and 5.3% during 2030–2049, and by roughly 8.0% and 13.3% in 2080–2099 under SSP2-4.5 and SSP5-8.5, respectively. The SSP2-4.5 and SSP5-8.5 climatic scenarios for 2030 were presented in our study. The study by Lakhyar et al. [39] focuses on an epidemic of black rust (*Puccinia graminis*) that affects wheat. But according to the current study, there are four types of wheat crop diseases: rust, fusarium wilt, powdery mildew, and blight [40]. focuses on important wheat diseases that are present in the United States. Research identifies a number of illnesses that impact wheat harvests, particularly in Morocco. Nigam et al. [41] concentrated on a model for the automatic diagnosis of major wheat rust diseases that was based on the EfficientNet architecture. With the introduction of the CropNet technique, wheat diseases were identified with an astonishing 99.80% accuracy rate [42]. To identify different wheat diseases, we used the NIRFO framework, such as RF-DE, RF-GA, RF-GWO, RF-HBA, and RF-PSO. The RF-GA model achieved the highest performance among all evaluated RF models for wheat disease prediction in the NIRFO system, with an AUC exceeding 0.83. Another study by Jain et al. [43], GA and PSO are commonly used in water-resource optimization. A multi-objective model was developed to evaluate groundwater–surface water interactions in low-flow regions. The study shows that modifying paddy, cotton, and pulse cultivation areas can improve net farm returns while maintaining groundwater sustainability. Model scenarios

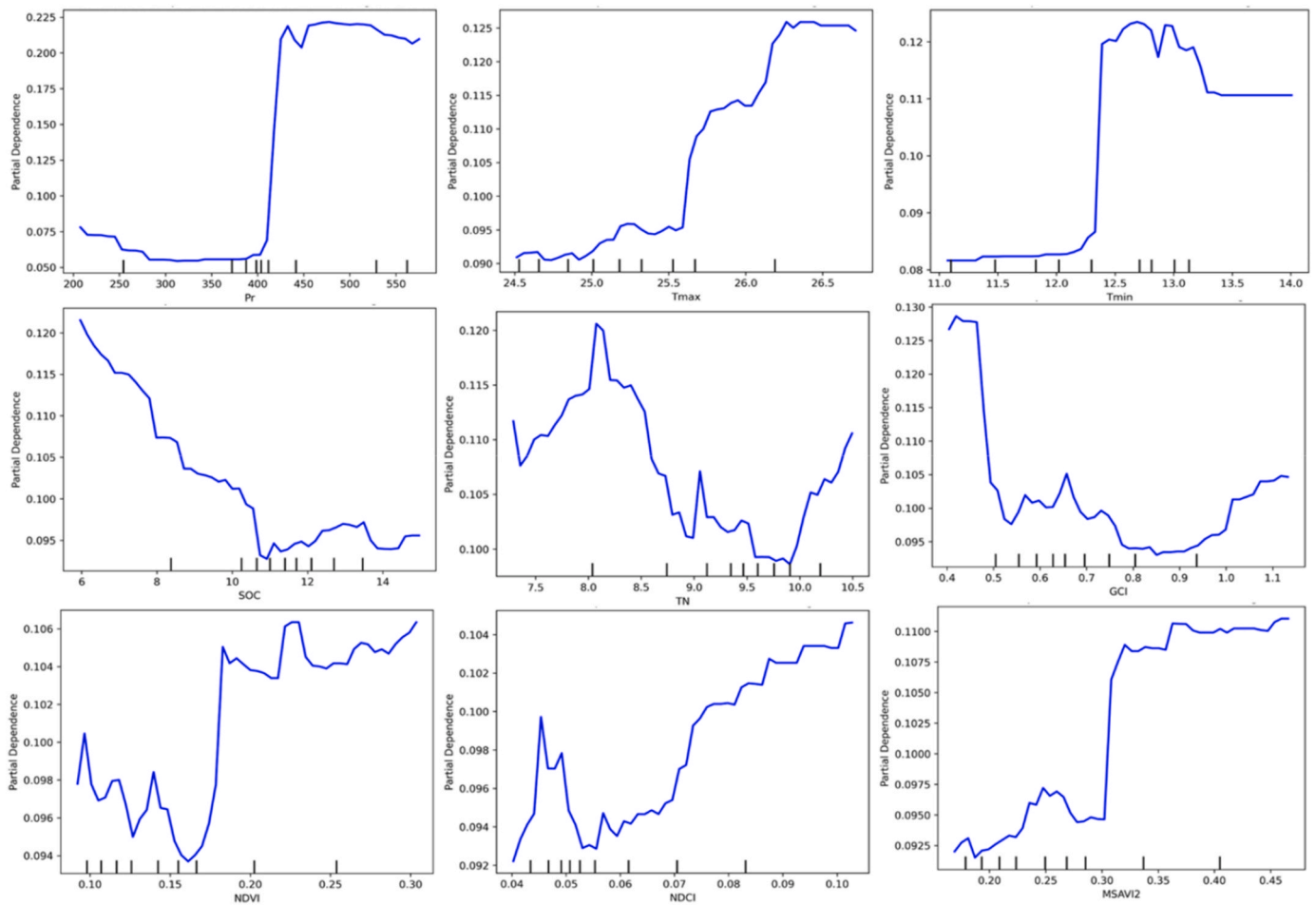


Fig. 16. Partial dependence plots for influencing parameters of *Blight* wheat disease.

considered 5–25% changes in cropping areas and 0–70% groundwater extraction. A comparable study by Sahoo et al. [9] reported that the RF model demonstrated the highest performance, attaining a test accuracy of 0.70 in predicting rice crop diseases across eastern India by utilizing climatic variables and RS-derived vegetation indices. Bijlwan et al. [12] demonstrated that integrating real-time meteorological data with machine learning, especially artificial neural networks (ANN), enables accurate prediction of wheat disease severity. Elfouly et al. [44] recently introduced an advanced transfer learning model within a data augmentation framework for assessing corn disease severity, achieving a classification accuracy of 97.33%.

In their study, Hessane et al. [45] used a publicly accessible dataset. Our study used field data to conduct a field research project. Using data from the NASA POWER database, Cól et al. [46] conducted a study on daily climatic variables such as Tmin, Tmax, mean temperature (Tmean), relative humidity (RH), and Pr. Pr, Tmin, Tmax, SOC, TN, GCI, NDVI, NDCI, and MSAVI2 were among the variables that were included in our investigation. These are significant predictor factors that can influence model predictions in either a good or a negative way. Notably, temperature and Pr have a significant impact on how severe diseases are in wheat crops. Ganeva et al. [47] analyzed the relationship between chlorophyll content (CC) and infection response in wheat using nine vegetation indices (VIs) sensitive to rust and leaf spot. Regression results indicated that leaf CC could be accurately estimated only in the late growth group, with R^2 values of 0.63 (cross-validation). Our SHAP-based explainable AI analysis revealed that Pr and temperature are the most influential factors determining the severity of wheat diseases. Lee et al. [48] developed ML models using Accumulated Local Effects (ALE) to explain how meteorological variables influence

Fusarium Head Blight occurrence, identifying precipitation around 60 mm during the flowering stage as a critical factor for disease incidence. According to Pequeno et al. [49], the current climate makes 6.4 million hectares of cropland susceptible to wheat blast; nevertheless, future temperature and humidity increases are anticipated to increase the area susceptible to infection, especially in the Southern Hemisphere. According to our research, Blight, Fusarium Wilt, Powdery Mildew, and Rust are mainly found in the northern, central, and southern parts of our selected study area now. However, by 2030, all these diseases will decline because of the adoption of climate change adaptation strategies, timing of fertilizer applications, and irrigation management techniques.

6. Adaptation strategies and policy recommendations

6.1. Adaptation strategies

6.1.1. Develop disease-resistant wheat varieties

Accelerated Breeding Programs: Utilize genomic selection, speed breeding, and doubled haploid technologies to rapidly develop wheat varieties resistant to Fusarium, yellow rust, and Hessian fly. For example, short-duration varieties have shown higher resilience under Moroccan rainfed conditions.

Hybrid Wheat: Invest in hybrid wheat research to enhance genetic diversity and disease resistance while maintaining agronomic performance.

Collaborative Research: Strengthen partnerships with international institutions (e.g., University of Nebraska–Lincoln) to identify resistant genotypes and share knowledge on disease management.

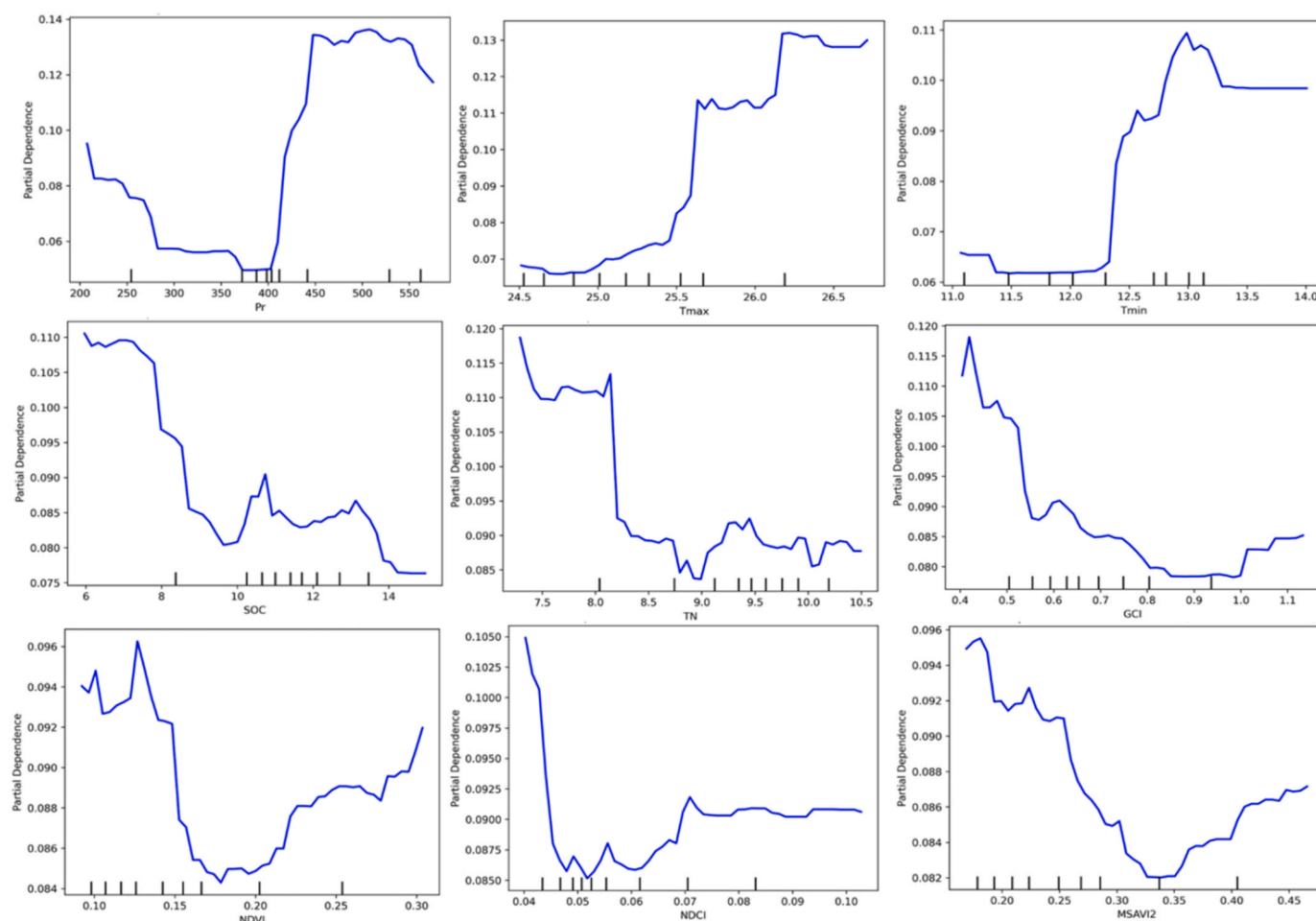


Fig. 17. Partial dependence plots for influencing parameters of *Fusarium Wilt* wheat disease.

6.1.2. Strengthen seed systems and distribution

Early Generation Seed (EGS) Production: Address bottlenecks in breeder seed multiplication to ensure farmers have access to high-quality, disease-resistant varieties. Current inefficiencies in Morocco's seed system have blocked genetic gains for over a decade.

Farmer Education: Promote certified seed adoption through extension services, emphasizing the economic benefits of resistant varieties over traditional seeds.

6.1.3. Integrated disease management

Agronomic Practices: Optimize sowing dates (November–early December) to reduce exposure to disease-favorable conditions. Delayed planting increases vulnerability to *Fusarium* and root rot.

Chemical and Biological Controls: Combine fungicides with biocontrol agents to manage mycotoxin contamination from *Fusarium*. Monitor pesticide use to avoid resistance buildup.

Crop Rotation and Soil Health: Implement conservation agriculture to break disease cycles and improve soil resilience against pathogens.

6.1.4. Climate-smart monitoring and early warning systems

Crop Modeling: Use tools like APSIM to predict disease outbreaks under climate scenarios and recommend adaptive measures (e.g., irrigation adjustments, cultivar selection).

Remote Sensing: Deploy satellite-based monitoring to detect disease hotspots and guide targeted interventions.

6.1.5. Enhance farmer capacity and collaboration

Extension Services: Train farmers on disease identification, resistant varieties, and adaptive practices through Morocco's Green Morocco Plan (GMP) networks.

International Exchanges: Expand programs like the USDA Scientific Exchanges to facilitate knowledge transfer on disease epidemiology and management.

6.2. Policy recommendations

- Allocate funding for national breeding programs (e.g., INRA) and incentivize private-sector involvement in developing climate-smart seeds. Revise seed royalty agreements to ensure the timely distribution of certified seeds to farmers.
- Update seed policies to streamline EGS production and distribution. Establish public-private partnerships to increase certified seed availability, particularly in rainfed regions.
- Subsidize eco-friendly pesticides and provide integrated pest management (IPM) training through Morocco's Green Generation Strategy.
- Invest in national agricultural research stations (e.g., Marchouch, Sidi El Aidi) to build capacity in predictive analytics and real-time data dissemination.
- Integrate disease resilience into national strategies (e.g., Green Generation 2020–2030) and prioritize funding for farmer-led innovation hubs.

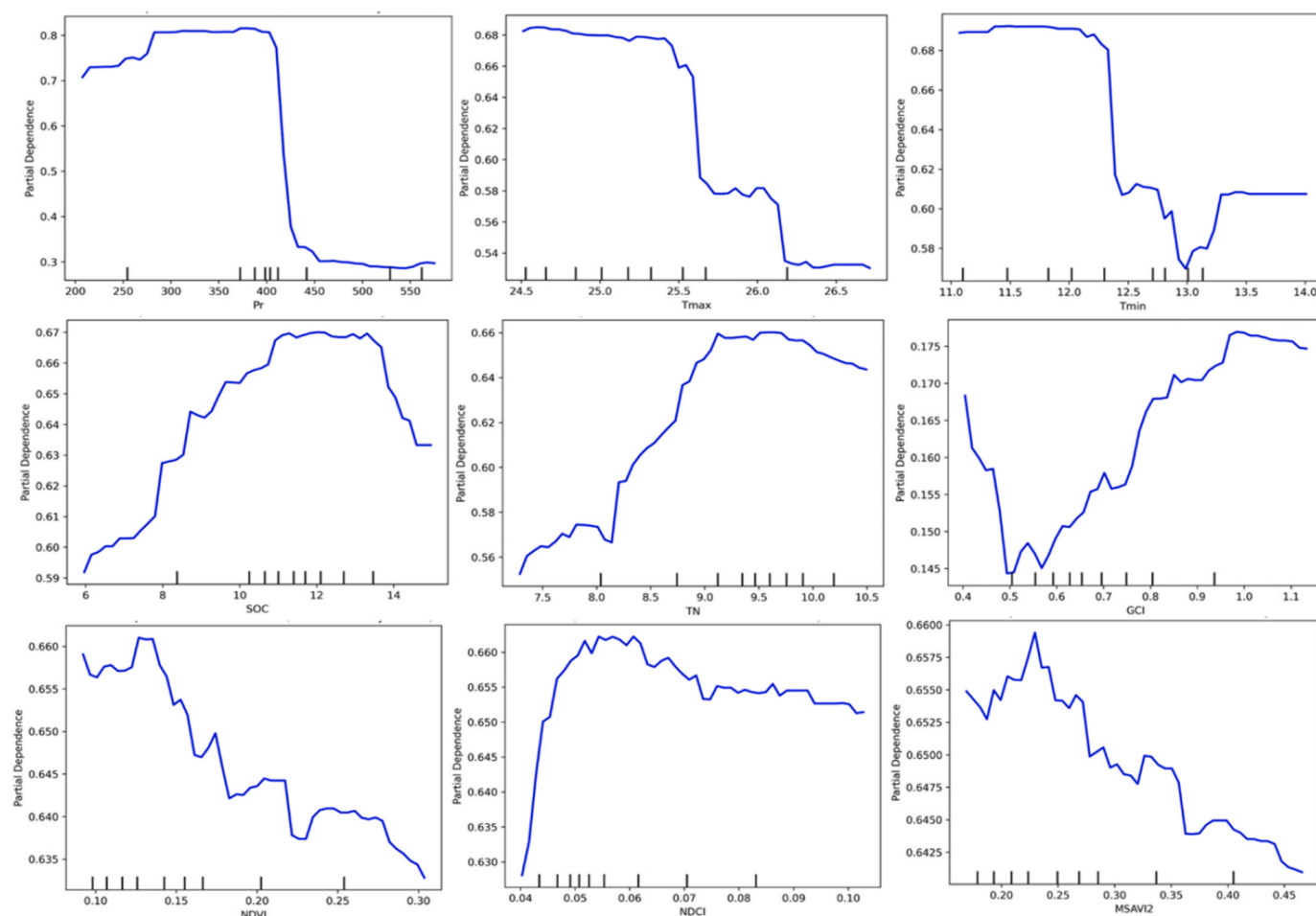


Fig. 18. Partial dependence plots for influencing parameters of *Powdery Mildew* wheat disease.

7. Disease decision-support system

Fig. 21 constitutes a fully functional decision-support system (DSS) remains unsubstantiated, as it lacks an operational design encompassing user interfaces, real-time data acquisition, and feedback mechanisms necessary for practical implementation. A robust agricultural DSS requires a well-defined pipeline integrating data collection, modeling, and actionable output delivery rather than stand-alone predictive outputs [50]. Existing systems such as CGMS-MAROC (Crop Growth Monitoring System) [51], the European MARS system (URL: <https://agri4cast.jrc.ec.europa.eu/>), CropWatch (URL: <http://cloud.cropwatch.com.cn/>), the American USDA-FAS (URL: <https://www.fas.usda.gov/>), CRAFT (URL: <https://ccafs.cgiar.org/tags/craft>), SEPTTR (Septoria Leaf Blotch Model) [52], the Integrated Canadian Crop YieldForecaster28 (ICCYF) (URL: https://www.statcan.gc.ca/en/statistical-programs/document/5225_D1_T9_V1); eDWIN and Online Pest Warning System (OPWS)

[50] for instance, provide early warning for drought and yield forecasting but do not yet incorporate real-time crop disease monitoring or farmer-oriented decision frameworks [51]. Furthermore, successful web-based DSS platforms illustrate that user-centered design, participatory validation, and field-level advisory interfaces are critical for usability and impact [53]. In contrast, an effective climate-integrated DSS for Morocco's wheat sector should combine real-time or near-real-time climatic data (temperature, precipitation, humidity) with biophysical indicators such as SOC, TN, and VIs (NDVI, MSAVI2, NDCI, CGI) to predict disease outbreaks of blight, powdery mildew, rust, and fusarium wilt. The prediction engine should employ epidemiological or ML models coupled with future-climate scenarios (SSP2-4.5, SSP5-8.5) to

generate early warnings when environmental thresholds conducive to infection are exceeded [54]. Decision rules should translate model outputs into actionable agronomic recommendations such as optimized fungicide application timing, resistant cultivar selection, and adaptive management aligned with sustainability goals and the SDGs [55]. To ensure long-term resilience, such a DSS must include a feedback loop for continuous validation, model recalibration, and integration with national early-warning systems and adaptation policies. Overall, this holistic integration of explainable AI, climate projections, and participatory design is essential to transform predictive modeling into a truly operational DSS for wheat disease management under changing climate conditions. According Patel et al. [56] contemporary pest surveillance systems increasingly employ Internet of Things (IoT) technologies, including interconnected sensors and smart imaging devices, to monitor insect populations in real time. The collected data are transmitted wirelessly to cloud-based platforms, where analytical algorithms process and interpret the information to produce automated alerts and management suggestions. To support accurate crop disease modelling and classification, stratified sampling was implemented to maintain proportional representation of disease classes across varying field conditions, thereby reducing sampling bias and improving model performance [57]. These outputs are then delivered to farmers and agricultural advisors, supporting timely, evidence-based decisions for improved crop protection and pest control efficiency.

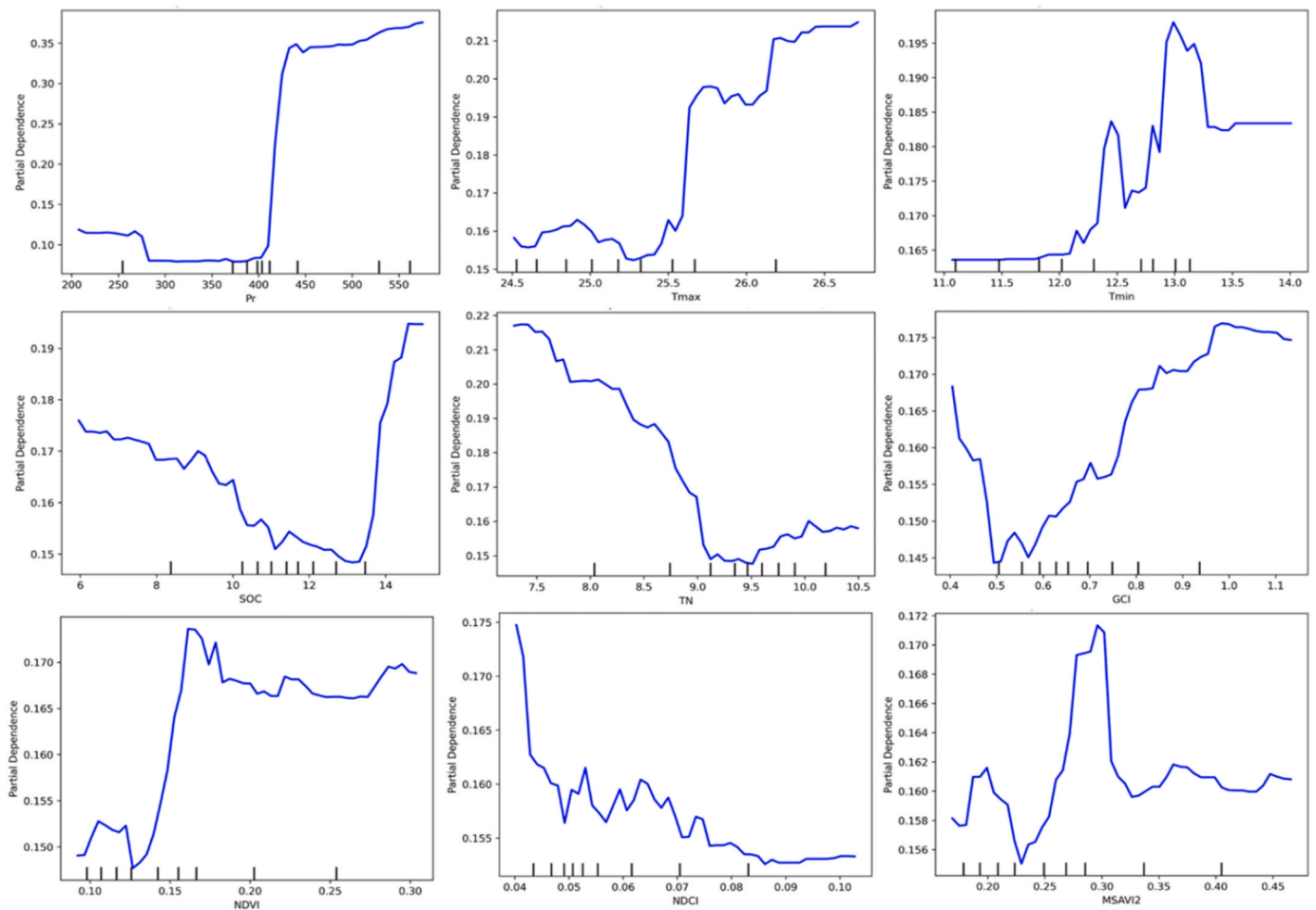


Fig. 19. Partial dependence plots for influencing parameters of *Rust* wheat disease.

8. Limitations and future scope

8.1. Variable dataset sizes for diseases

Different numbers of cases of each detected WCD were included in the dataset. For instance, rust had 111, blight had 80, powdery mildew had 545, and fusarium wilt had the fewest, with only 45 occurrences. Significant variations in the number of instances per class can cause class imbalance problems while training ML models, even if this isn't specifically mentioned as a disadvantage. The PSO method demonstrated a notable improvement in managing class imbalance, according to the study, indicating that this was a significant component.

8.2. Complexity of the mapping process

According to the sources, the "intricate interplay" of numerous environmental and agricultural factors, such as climatic conditions, soil nutrient levels, and vegetation health, drives the "complex process" of creating thematic maps to show WCD occurrences. It takes "careful consideration of these multifaceted influences" to map and classify diseases accurately. This demonstrates how challenging and intricate the issue the methodology is attempting to solve is by nature.

8.3. Uncertainty in future climate projections

To forecast future WCD trends under various climatic scenarios, the study makes use of CMIP6 datasets. According to the sources, in order to "minimize the uncertainty associated with projected climate scenarios,"

ensemble averages were computed using several GCMs. Although a technique to lessen uncertainty was utilized, this suggests that the future climate data used for forecasts has intrinsic uncertainty.

8.4. Model sensitivity to input variations

Through the measurement of performance changes under input perturbations, the sensitivity analysis method assesses the resilience of the model. High sensitivity denotes susceptibility to noise in the input data or adversarial attacks. Although this is a measure of the model's characteristics rather than a flaw in the analysis process itself, if left unchecked, the possibility that model predictions could diverge greatly from tiny input differences could pose a problem for real-world application.

8.5. Integration of multi-resolution data

With original spatial resolutions varying from 10 m (Sentinel-2 vegetation indices) to around 4.6 km (TerraClimate) and even 27.8 km (CMIP6), the study combined CDIF layers from many sources. A consistent 30-m resolution was used to resample each of these layers. Although it isn't stated in the sources specifically, resampling data from coarser to finer resolutions may result in the loss or alteration of information fidelity.

8.5.1. Future scope

Future wheat disease mapping in Morocco holds significant potential through the integration of high-resolution satellite imagery i.e

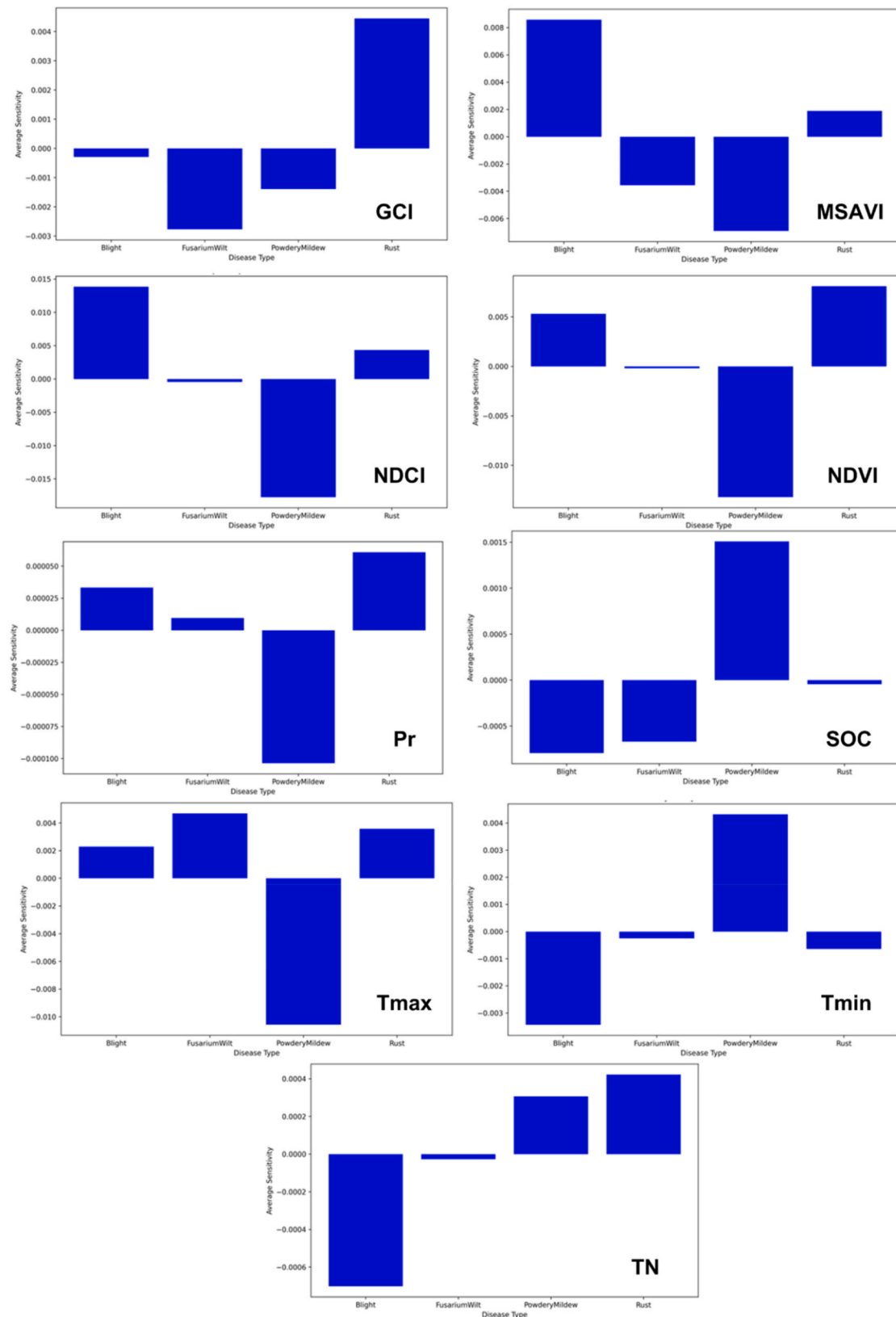


Fig. 20. Perturbation sensitivity analysis for the importance of the influencing parameters for the prediction of wheat crop diseases.

PlanetScope, EcoStress, WorldView-3, SkySat and vegetation indices to detect early signs of outbreaks, alongside the use of UAVs equipped with multispectral or thermal sensors for precise field-level monitoring. Developing predictive models using advanced hybrid transfer DL and

artificial intelligence can improve the accuracy and efficiency of disease forecasting. Leveraging large datasets encompassing climate, soil, and crop health information can enhance predictive models for timely disease risk assessment. Climate-disease interaction studies are vital for

Wheat Disease Management System

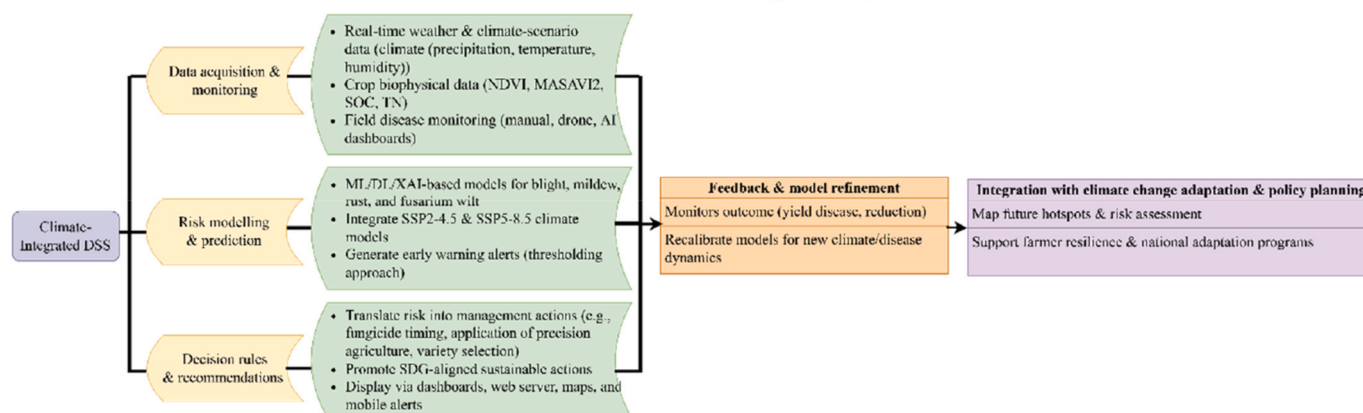


Fig. 21. Climate-integrated decision support system (DSS) for wheat disease management.

understanding how changing environmental conditions influence disease dynamics, enabling adaptive management strategies. Mobile-based early warning systems can empower farmers with real-time alerts and recommendations, while big data, IoT, and cloud computing platforms like GEE facilitate large-scale, data-driven analysis. Participatory approaches involving farmers can improve local surveillance accuracy, and disease-specific spatial risk maps for threats like stripe rust, septoria, and Fusarium will aid in targeted intervention. Integrating these efforts with sustainable agricultural practices, including crop rotation and resistant varieties, will support long-term disease control and improved wheat production resilience.

9. Conclusions

The current study uses nature-inspired random forest optimization (i.e., GA, PSO, GWO, DE, and HBA) and SHAP XAI in Morocco to predict wheat crop disease based on biophysical characteristics. For the purpose of climate forecasts, the current study also incorporated the SSP2-4.5 and SSP5-8.5 scenarios from three GCMs (EC-Earth3, NorESM2-LM, and MIROC6) under CMIP6. Using Boruta feature selection approaches, the results demonstrate that precipitation has the most influence over the other parameters. It was noted that blight disease, which is brought on by the fungi *Fusarium culmorum* and *Fusarium graminearum* under warm, humid conditions, is mostly found in the north, with sporadic incidences in the south. Additionally, it was noted that rust is more common in the northern, central, and southern regions, whereas powdery mildew wheat crop disease is very uncommon in the southern and eastern regions. It was discovered that when compared to another model, RF-GA predicts wheat crop diseases remarkably well. Future findings demonstrate that the RF-HBA, RF-PSO, and RF-GA models are capable of effectively identifying the intricate and dense blight disease in the northern regions. This disease is brought on by a number of diseases, such as bacteria and fungi, that affect the leaves. According to SHAP dependence results, temperature and precipitation have a significant impact on wheat crop health and raise the likelihood of developing various diseases, in addition to powdery mildew.

It was informed that MSAVI exhibits a strong positive sensitivity to blight and a moderate negative sensitivity to powdery mildew disease, whereas GCI exhibits a significant positive sensitivity to rust, indicating that GCI is especially adept at identifying or simulating rust. Morocco's wheat crops are prone to a number of diseases that can seriously affect their quality and productivity. Morocco's wheat crops are susceptible to diseases that lower grain quality and yield, such as Fusarium head blight, Septoria leaf blotch, tan spot, and stem, leaf, and stripe rusts. The country's varied temperatures, especially the wetter northern regions, are ideal for the growth of these fungal infections. To reduce losses, farmers control them with fungicides, crop rotation, resistant cultivars,

and better irrigation techniques. Controlling outbreaks and guaranteeing steady wheat output depend heavily on monitoring and early identification.

CRedit authorship contribution statement

Satiprasad Sahoo: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Chiranjit Singha:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Ajit Govind:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

CGIAR Pooled Funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We acknowledge the support from the CGIAR Digital Transformation Accelerator (DTA) AoW-2 Actions Lab. We also thank the CGIAR ClimBer (Climate Resilience) Initiative through which the trans Morocco survey was undertaken. A subset of this survey data was used in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jafr.2026.102924>.

Data availability

Data will be made available on request.

References

- [1] J. Benjamin, D.O. Oyedokun, E.V. Oziegbe, J. Oni, E.B. Ogundare, G.O. Ujah, A. Adebayo, Cereal production in Africa: the threat of current plant pathogens in changing climate—a review, *Discover Agriculture* 2 (1) (2024) 33.
- [2] A. Ramdani, H. Ibriz, C. Royo, M. Nachit, N. Fonzo, J. Araus, Survey of wheat diseases in Morocco during three consecutive seasons 1997, 1998 and 1999, *Proc. Durum wheat improvement in the mediterranean region: New challenges*, Zaragoza (C. Royo, M. Nachit, N. Di Fonzo, J. Araus, eds). CHEAM, IRTA, CIMMYT e ICARDA (2000) 407–411.
- [3] M. El Jarroudi, R. Lahlali, L. Kouadio, A. Denis, A. Belleflamme, M. El Jarroudi, B. Tychon, Weather-based predictive modeling of wheat stripe rust infection in Morocco, *Agronomy* 10 (2) (2020) 280.
- [4] W. Bozalmat, S.B. Alaoui, A.A.H. Sidikou, A. Abouabdillah, The management of fungal diseases in organic production systems through a mixture of durum wheat varieties, *Sustainability* 16 (21) (2024) 9304.
- [5] M. El Jarroudi, R. Lahlali, H. El Jarroudi, B. Tychon, A. Belleflamme, J. Junk, L. Kouadio, Employing weather-based disease and machine learning techniques for optimal control of septoria leaf blotch and stripe rust in wheat, in: *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019) Volume 2-Advanced Intelligent Systems for Sustainable Development Applied to Agriculture and Health*, Springer International Publishing, 2020, pp. 157–165.
- [6] J. Bouarda, F.M. Bassi, H. Wallwork, M. Benchacho, M. Labhili, I. Maafa, F. Bentata, Evaluation of durum wheat genotypes for resistance against root rot disease caused by Moroccan *Fusarium culmorum* isolates, *Plant Pathol. J.* 38 (1) (2022) 1.
- [7] E.M. Zahraoui, A. Ramzi, W. Habybellah, B. El Yousfi, K. Boukachabine, Alternative control of helminthosporium leaf spot on wheat using essential oils of *Origanum compactum* and *Thymus satureioides*, *Moroccan Journal of Agricultural Sciences* 5 (1) (2024) 63–69.
- [8] E.M. Zahraoui, B. El Yousfi, M. Chahbouni, F.E. Chellakhi, S. Qarchaoui, K. Boukachabine, Effect of essential oils from *Thymus satureioides* and *Origanum compactum* on wheat root rot induced by *Fusarium culmorum* and *Bipolaris sorokiniana*, *Moroccan Journal of Agricultural Sciences* 4 (2) (2023) 86–92.
- [9] S. Sahoo, C. Singha, A. Govind, M. Sharma, Leveraging ML to predict climate change impact on rice crop disease in Eastern India, *Environmental Monitoring and Assessment* 197 (4) (2025) 366. <https://doi.org/10.1007/s10661-025-13744-wS>.
- [10] J.T. Abatzoglou, S.Z. Dobrowski, S.A. Parks, K.C. Hegewisch, Terracimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015, *Scientific Data* 5 (2018) 170191, <https://doi.org/10.1038/sdata.2017.191>.
- [11] C.F. Ajillogba, S. Walker, Using crop modeling to find solutions for wheat diseases: a review, *Front. Environ. Sci.* 10 (2023), <https://doi.org/10.3389/fenvs.2022.987765>.
- [12] A. Bijlwan, R. Ranjan, M. Singh, R. Purohit, A.J. Nath, S. Chakravarty, Predicting crop disease severity using real time weather variability through machine learning algorithms, *Sci. Rep.* 15 (1) (2025) 34767, <https://doi.org/10.1038/s41598-025>.
- [13] A. Sowmiya, M. Suganthi, R. Saranya, K. Ganesan, P. Janaki, Harnessing the power of microbes for plant protection in regenerative agriculture: a comprehensive review, *Journal of Natural Pesticide Research* 14 (2025) 100170, <https://doi.org/10.1016/j.napere.2025.100170>.
- [14] N.J. Maywald, D. Francioni, M. Mang, U. Ludewig, Role of mineral nitrogen nutrition in fungal plant diseases of cereal crops, *Crit. Rev. Plant Sci.* 42 (3) (2023) 93–123, <https://doi.org/10.1080/07352689.2023.2196100>.
- [15] S.R. Sultana, A. Ali, A. Ahmad, M. Mubeen, M. Zia-Ul-Haq, S. Ahmad, S. Ercisli, H. Z.E. Jaafar, Normalized difference vegetation index as a tool for wheat yield estimation: a case study from faisalabad, Pakistan, *Sci. World J.* 2014 (2014) 1–8, <https://doi.org/10.1155/2014/725326>.
- [16] R. Vidican, A. Mălinaș, O. Ranta, C. Moldovan, O. Marian, A. Ghețe, C.R. Ghișe, F. Popovici, G.M. Cătușescu, Using remote sensing vegetation indices for the discrimination and monitoring of agricultural crops: a critical review, *Agronomy* 13 (12) (2023) 3040, <https://doi.org/10.3390/agronomy13123040>.
- [17] A. Nega, Climate change impacts on agriculture: a review of plant diseases and insect pests in Ethiopia and East Africa, with adaptation and mitigation strategies, *Advances in Agriculture* 2025 (1) (2025), <https://doi.org/10.1155/aia/5606701>.
- [18] P. Prasad, S.C. Bhardwaj, R.K. Thakur, S. Adhikari, O.P. Gangwar, C. Lata, S. Kumar, Prospects of climate change effects on crop diseases with particular reference to wheat, *Journal of Cereal Research* 13 (2) (2021), <https://doi.org/10.25174/2582-2675/2021/112817>.
- [19] L. Breiman, *Machine Learning* 45 (1) (2001) 5–32.
- [20] H. Zhou, L. Ma, Y. Xiang, Y. Su, J. Li, J. Chen, S. Lu, C. Chen, Q. Wu, Estimating reference evapotranspiration using hybrid models optimized by bio-inspired algorithms combined with key meteorological factors, *Comput. Electron. Agric.* 230 (2024) 109862, <https://doi.org/10.1016/j.compag.2024.109862>.
- [21] A.M. Helmi, M.E. Lotfy, Recent advances of nature-inspired metaheuristic optimization, in: M. Khosravi, N. Gupta, N. Patel, T. Senjyu (Eds.), *Frontier Applications of Nature Inspired Computation*, Springer, Singapore, 2020.
- [22] L.J. Goodfellow, J. Shlens, C. Szegedy, Explaining and harnessing adversarial examples, *arXiv (Cornell University)* (2014). <https://doi.org/10.48550/arxiv.1412.6572>.
- [23] J. Kennedy, R. Eberhart, Particle swarm optimization, in: *Proceedings of ICNN'95-International Conference on Neural Networks*, IEEE, 1995, pp. 1942–1948.
- [24] J. Kennedy, R. Eberhart, Particle swarm optimization, in: *Proceedings of ICNN'95-International Conference on Neural Networks*, Perth, WA, Australia vol. 4, 1995, pp. 1942–1948. <https://doi.org/10.1109/ICNN.1995.488968>.
- [25] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey Wolf Optimizer, *Advances in Engineering Software* 69 (2014) 46–61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>.
- [26] R. Storn, K. Price, Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, *Journal of Global Optimization* 11 (1997) 341–359, <https://doi.org/10.1023/A:1008202821328>.
- [27] F.A. Hashim, E.H. Houssein, K. Hussain, M.S. Mabrouk, W. Al-Atabany, Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems, *Mathematics and Computers in Simulation* 192 (2021) 84–110. <https://doi.org/10.1016/j.matcom.2021.08.013>.
- [28] K. Lee, H. Jeon, S. Park, J. Cho, K. Kim, Comparison of projected rice blast epidemics in the Korean peninsula between the CMIP5 and CMIP6 scenarios, *Clim. Change* 173 (1–2) (2022) 12, <https://doi.org/10.1007/s10584-022-03410-2>.
- [29] B. Thrasher, E.P. Maurer, C. McKellar, P.B. Duffy, Technical note: bias correcting climate model simulated daily temperature extremes with quantile mapping, *Hydrol. Earth Syst. Sci.* 16 (9) (2012) 3309–3314, <https://doi.org/10.5194/hess-16-3309-2012>.
- [30] A. Talikan, R. Salapuddin, J. Aksan, R. Rahimulla, A. Ismael, R. Jimlah, N. Idris, R. Dammang, D. Jamar, E. Sarahadil, R. Ajan, On Paired Samples T-Test: Applications, Examples and Limitations, 2025, <https://doi.org/10.5281/zenodo.10987546>.
- [31] T. Ye, J. Nie, J. Wang, P. Shi, Z. Wang, Performance of detrending models of crop yield risk assessment: evaluation on real and hypothetical yield data, *Stoch. Environ. Res. Risk Assess.* 29 (1) (2014) 109–117, <https://doi.org/10.1007/s00477-014-0871-x>.
- [32] S. Franceschini, E. Gandola, M. Martinoli, L. Tancioni, M. Scardi, Cascaded neural networks improving fish species prediction accuracy: the role of the biotic information, *Scientific Reports* 8 (1) (2018). <https://doi.org/10.1038/s41598-018-22761-4>.
- [33] L.J. Goodfellow, J. Shlens, C. Szegedy, Explaining and harnessing adversarial examples, *arXiv (Cornell University)* (2014). <https://doi.org/10.48550/arxiv.1412.6572>.
- [34] S.M. Lundberg, S. Lee, A unified approach to interpreting model predictions, *arXiv (Cornell University)* (2017). <https://doi.org/10.48550/arxiv.1705.07874>.
- [35] B. Mieslerová, R.T. Cook, C.P. Wheeler, A. Lebeda, Ecology of powdery mildews—Influence of abiotic factors on their development and epidemiology, *Crit. Rev. Plant Sci.* 41 (6) (2022) 365–390.
- [36] Y. Ali, S. Iqbal, H.M. Aatif, K. Naveed, A.A. Khan, M. Ijaz, M.M. Magsi, S. Ahmad, A.U.A. Syed, M.A. Magsi, R.K. Iqbal, N.A. Bukhari, A.A. Hatamleh, A. Raza, Predicting stripe rust severity in wheat using meteorological data with environmental response modeling, *J. King Saud Univ. Sci.* 35 (4) (2023) 102591, <https://doi.org/10.1016/j.jksus.2023.102591>.
- [37] J.Y. Jung, J.H. Kim, M. Baek, C. Cho, J. Cho, J. Kim, K.H. Kim, Adapting to the projected epidemics of fusarium head blight of wheat in Korea under climate change scenarios, *Front. Plant Sci.* 13 (2022) 1040752.
- [38] M. Nouri, S. Veysi, CMIP6 multi-model ensemble projection of reference evapotranspiration using machine learning algorithms, *Agric. Water Manag.* 306 (2024) 109190, <https://doi.org/10.1016/j.agwat.2024.109190>.
- [39] Z. Lakhyar, S. Khoali, E.M. Farah, A. Laamari, Assessing Black rust spreading in wheat production using mathematical model, *Commun. Math. Biol. Neurosci.* 2024 (2024). Article-ID.
- [40] J. Singh, B. Chhabra, A. Raza, S.H. Yang, K.S. Sandhu, Important wheat diseases in the US and their management in the 21st century, *Front. Plant Sci.* 13 (2023) 1010191.
- [41] S. Nigam, R. Jain, S. Marwaha, A. Arora, M.A. Haque, A. Dheeraj, V.K. Singh, Deep transfer learning model for disease identification in wheat crop, *Ecol. Inform.* 75 (2023) 102068.
- [42] O. Jouini, M.O.E. Aoueilayine, K. Sethom, A. Yazidi, Wheat leaf disease detection: a lightweight approach with shallow CNN based feature refinement, *AgriEngineering* 6 (3) (2024) 2001–2022.
- [43] S. Jain, D. Ramesh, M.C. Trivedi, D.R. Edla, Evaluation of metaheuristic optimization algorithms for optimal allocation of surface water and groundwater resources for crop production, *Agric. Water Manag.* 279 (2023) 108181, <https://doi.org/10.1016/j.agwat.2023.108181>.
- [44] M.K. Elfouly, A.M. AbdelAziz, W.H. Gomaa, M. Abdalla, A deep learning-based framework for large-scale plant disease detection using big data analytics in precision agriculture, *J. Big Data* 12 (1) (2025), <https://doi.org/10.1186/s40537>.
- [45] A. Hessane, A. El Youssefi, Y. Farhaoui, B. Aghoutane, F. Amounas, A machine learning based framework for a stage-wise classification of date palm white scale disease, *Big Data Mining and Analytics* 6 (3) (2023) 263–272.
- [46] M.C.D. Cól, Weather-Based Logistic Regression Models for Predicting the Risk of Wheat Blast Epidemics, 2023.
- [47] D. Ganeva, L. Filchev, E. Roumenina, R. Dragov, S. Nedyalkova, V. Bozhanova, Winter Durum Wheat Disease Severity Detection with Field Spectroscopy in Phenotyping Experiment at Leaf and Canopy Level, *Remote Sensing* 16 (10) (2024) 1762, <https://doi.org/10.3390/rs16101762>.
- [48] N. Lee, J. Yang, J. Jung, Y. Kim, K. Kim, An interpretable machine learning technique to identify the key meteorological factors influencing the incidence of wheat fusarium head blight in Korea, *Comput. Electron. Agric.* 235 (2025) 110355, <https://doi.org/10.1016/j.compag.2025.110355>.
- [49] D.N. Pequeno, T.B. Ferreira, J.M. Fernandes, P.K. Singh, W. Pavan, K. Sonder, S. Asseng, Production vulnerability to wheat blast disease under climate change, *Nat. Clim. Change* 14 (2) (2024) 178–183.
- [50] A. Tratwal, M. Jakubowska, A. Pietrusińska-Radzio, Decision support systems in integrated Pest and disease management: innovative elements in sustainable agriculture, *Sustainability* 17 (18) (2025) 8111, <https://doi.org/10.3390/su17188111>.

- [51] Riad Balaghi, Mouanis Lahlou, Meriem Alaouri, CGMS-MAROC: the Drought Early Warning and Cereal Yield Forecasting System of Morocco, 2024.
- [52] P. Delfosse, Assessing the interplay between weather and septoria leaf blotch severity on lower leaves on the disease risk on upper leaves in winter wheat, *J. Fungi* 8 (2022) 1119.
- [53] P. Damos, Modular structure of web-based decision support systems for integrated Pest management. A review, *Agron. Sustain. Dev.* 35 (2015) 1347–1372, <https://doi.org/10.1007/s13593-015-0319-9>.
- [54] P. Juroszek, A. von Tiedemann, Climate change and potential future risks through wheat diseases: a review, *Eur. J. Plant Pathol.* 136 (2013) 21–33, <https://doi.org/10.1007/s10658-012-0144-9>.
- [55] K.C. Prah, H. Klink, M. Hasler, S. Hagen, J. Verreet, T. Birr, Can decision support systems help improve the sustainable use of fungicides in wheat? *Sustainability* 14 (23) (2022) 15599 <https://doi.org/10.3390/su142315599>.
- [56] A. Patel, B. Sahoo, P.R. Sahoo, Smart Pest Monitoring using Image Recognition and Web-Integrated IoT Devices. <https://doi.org/10.1109/elexcom67950.2025.11451404>.
- [57] S. Sahoo, C. Singha, A. Govind, M. Sharma, Leveraging ML to predict climate change impact on rice crop disease in Eastern India, *Environmental Monitoring and Assessment* 197 (4) (2025) 366. <https://doi.org/10.1007/s10661-025-13744-w>.