PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol x(x): xx–xx
https://doi.org/10.14719/pst.4974

HORIZON
e-Publishing Group

RESEARCH ARTICLE

# Forecasting rice blast disease severity using weather-dependent regression and time series models

**Meena A G[1], Gopalakrishnan C[1]*, Santosh Ganapati Patil[2], Kamalakannan A[1], Jagadeeswaran R[3], Sathyamoorthy N K[4], S Manonmani[5]**

[1] Department of Plant Pathology, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

[2] Agricultural Statistics, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

[3] Department of Soil Science, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

[4] Department of Agronomy, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

[5] Department of Rice, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

*Email: pcgopalagri@gmail.com

OPEN ACCESS

ARTICLE HISTORY
Received: 06 September 2024
Accepted: 14 October 2024
Available online
Version 1.0 : 25 November 2024

Check for updates

**Additional information**

**Peer review**: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

**Reprints & permissions information** is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

**Publisher's Note**: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Indexing**: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting



**CITE THIS ARTICLE**
Meena AG, Gopalakrishnan C, Patil SG, Kamalakannan A, Jagadeeswaran R, Sathyamoorthy NK, Manonmani S. Forecasting rice blast disease severity using weather-dependent regression and time series models. Plant Science Today (Early Access). https:/doi.org/10.14719/pst.4974

## Abstract

Rice (*Oryza sativa*) is a staple food crucial for food security and economic stability, especially in developing countries. However, rice cultivation faces significant challenges, with rice blast disease, caused by the fungal pathogen *Magnaporthe oryzae*, being one of the most severe threats, potentially leading to yield losses of up to 30 %. This study aims to develop and apply regression models, including stepwise regression, multiple linear regression (MLR) and ARIMA (Autoregressive Integrated Moving Average), to predict the severity of rice blast disease based on weather parameters. Weekly data over 7 years (2017-2023) were collected from the Paddy Breeding Station at Tamil Nadu Agricultural University, Coimbatore, encompassing various weather factors such as temperature, relative humidity, rainfall and solar radiation. Data pre-processing included handling missing values, detecting outliers and creating time-lagged variables. The study revealed distinct seasonal patterns in rice blast incidence, with peak occurrences observed from mid-November to late January. Among the regression models, the ARIMA model incorporating weather variables as external regressors demonstrated superior performance with an R-squared value of 0.92, compared to 0.55 for stepwise regression and 0.57 for MLR. Accurate predictions of rice blast outbreaks could enable farmers and agricultural managers to implement timely and targeted disease management strategies, reducing dependence on broad-spectrum fungicides and minimizing crop losses. This study contributes to data-driven agriculture and disease management, potentially leading to more effective, economically viable and environmentally sustainable rice cultivation practices.

## Keywords

Rice blast; *Magnaporthe oryzae;* Prediction; Stepwise regression; Multiple Linear Regression; ARIMA

## Introduction

Rice (*Oryza sativa*) is a fundamental staple food for millions, playing a critical role in global food security by providing sustenance to over half of the world's population and serving as a primary calorie source for billions (1). Its cultivation is crucial not only for food security but also for the economic stability of many developing countries (2). Despite its importance, rice cultivation faces significant challenges, with diseases being a major threat. Among these, rice blast, caused by the fungal pathogen *Magnaporthe oryzae*, is particularly, devastating; potentially resulting in yield losses of up to 30 % in susceptible varieties under optimal conditions (3). The severity and spread of rice blast are significantly influenced by environmental factors, particularly weather conditions such as temperature, humidity, rainfall and sunlight exposure (4).

In India, the application of fungicides after disease onset is common but leads to economic losses and poor disease control. Efficient, economical and environmentally friendly management of rice blast can be achieved through prior knowledge of the infection process in relation to weather factors. To address these challenges, meteorological-based modeling for early forewarning of the disease, provides tools for predicting disease status, guiding farmers to take timely protection measures, which results in better disease management, resource savings and reduced fungicide use (5). Given the disease's seasonal nature and its strong correlation with weather patterns, utilizing meteorological data for predictive modeling of disease outbreaks holds considerable potential. Conventional disease management strategies, such as fungicides and resistant varieties, are often reactive and resource-intensive. Recently, statistical and machine learning techniques have gained prominence in agricultural challenges, offering novel approaches for disease management and crop protection (6).

Among these methodologies, regression analysis emerges as a powerful tool for understanding and forecasting complex biological phenomena based on multiple variables. This study focuses on developing and applying regression models, specifically stepwise regression, multiple linear regression (MLR) and ARIMA (Autoregressive Integrated Moving Average), to predict the severity of rice blast disease. Stepwise regression was selected for its ability to automatically identify the most significant weather variables affecting disease severity, simplifying the model while maintaining key predictors. MLR was chosen for its capacity to simultaneously assess the impact of multiple weather parameters on disease severity, providing a comprehensive view of the relationships between the independent variables. ARIMA was incorporated due to its strength in handling time series data, making it ideal for forecasting the disease based on historical patterns while incorporating weather variables as external regressors. By incorporating key weather parameters-such as minimum and maximum temperatures, relative humidity (morning and evening), rainfall, sunshine hours, wind speed, evaporation, solar radiation and leaf wetness-as independent variables, this research aims to construct a robust predictive model with disease severity as the dependent variable. The study utilizes weekly data from a single location, acknowledging the disease's seasonal occurrence. This approach aligns with previous studies that have successfully predicted plant diseases based on environmental factors (7, 8).

The implications of this research are substantial. Accurate predictions of rice blast outbreaks could enable farmers and agricultural managers to implement timely and targeted disease management strategies, reducing dependence on broad-spectrum fungicides and minimizing crop losses. Additionally, in the context of climate change and increasing weather variability, such predictive tools are crucial for ensuring food security and agricultural sustainability (9). By integrating statistical rigor with practical application, this study aims to contribute to the expanding body of knowledge on data-driven agriculture and disease management, potentially leading to more effective, economically viable and environmentally sustainable rice cultivation practices, ultimately benefiting millions of farmers and consumers globally.

## Materials and Methods

### Study Area and Data Collection

The study was conducted at Paddy Breeding Station (PBS), Tamil Nadu Agricultural University, Coimbatore, a region known for rice cultivation and periodic outbreaks of rice blast disease. Data were collected weekly over a period of 7 years (from the year of 2017 to 2023), encompassing all three growing seasons and the weekly data comprised of per cent (%) disease incidence (PDI) due to blast attack.

### *Weather Data Collection*

Weather parameters were recorded daily using an automated weather station (Yuktix Technologies, Bangalore, Karnataka) located within the study site. The data collected, included the following weather variables - Maximum temperature ($X_1$), Minimum temperature ($X_2$), Relative Humidity (morning) ($X_3$), Relative Humidity (evening) ($X_4$), Wind Speed ($X_5$), Evaporation ($X_6$), Sunshine ($X_7$), Solar radiation ($X_8$), Rainfall ($X_9$) and Leaf wetness ($X_{10}$). Data were aggregated into weekly averages for analysis to align, with the disease assessment schedule.

### Assignment of score

The standard scoring system provided by the International Rice Research Institute (IRRI), Philippines (Anonymous 2002), was used to assess rice blast incidence. The scores range from 0 to 9 (Table 1, Fig. 1). The disease severity was calculated using % Disease Index (10).

PDI = Sum of all individual rating/Total number of leaf observed x maximum rating x 100

### Data Pre-processing

Raw data were pre-processed using R software (version 4.1.0, R Core Team, 2021) to handle missing values, detect outliers and create time-lagged variables. Missing values (<5 % of the dataset) were imputed using the Miss Forest algorithm due to its ability to handle both continuous and categorical data while preserving complex relationships between variables (11). This non-parametric method is robust and has been shown to perform well in scenarios with small amounts of missing data. Outliers were identified using the Interquartile Range (IQR) method and were either corrected or removed based on field notes and expert consultation. Time-lagged variables were created for each weather parameter, considering lags of 1 to 4 weeks to account for the incubation period of rice blast disease.

**Table 1.** Rice blast disease scale: 0-9.

| Scale | Infected leaf area |
|---|---|
| 0 | There were no lesions seen |
| 1 | Pinpoint sized small brown spots |
| 2 | little roundish to slightly elongated 1-2 mm in diameter |
| 3 | Similar to scale 2, but on upper leaves as well |
| 4 | under 4 % |
| 5 | 4-10 % |
| 6 | 11-25 % |
| 7 | 26-50 % |
| 8 | 51-75 % |
| 9 | above 75 % |

**Fig. 1.** Disease scale 0-9 for rice blast.

0= There were no lesions seen; 1= Pin point sized small brown spots; 2= little roundish to slightly elongated 1 -2 mm in diameter; 3= similar to scale 2; but on upper leaves as well; 4= under 4 %; 5=4-10 %; 6=11-25 %; 7=26- 50 %; 8=51 -75 %; 9= above 75 %

## Severity incidence and seasonal distribution pattern of rice blast

Rice blast severity over the years was represented using a radial plot in R (version 4.3.1) packages**:** ggplot2, dplyr, scales, ggthemes. The severity grade was sorted into severe, high, moderate, less and no incidence categories. The ggplot2 package was used to create the radial bar plot, the data was mapped to aesthetic components using the aes function. Bars were created using geom_bar function and the plot was converted to a polar coordinate system using the coord_polar function.

The seasonal plot was constructed to visualize the seasonal pattern of rice blast using function Matplotlib, Seaborn and Stat models in Google colab for analysis and visualization. The time series plot was constructed with R software using ggplot2 and dplyr in R. The data, comprising weekly rice blast incidence, were pre-processed to calculate weekly means and visualize seasonal trends using line plots. A LOESS regression line was added for smoothing and the overall mean incidence was highlighted with a horizontal line.

## Model Development

Three types of regression models were developed and compared:

### Stepwise Regression

Stepwise regression was performed using the 'step' function in R, with both forward and backward selection. The Akaike Information Criterion (AIC) was used to determine the optimal model (12). Stepwise regression is an automated method that adds or removes predictor variables based on their statistical significance in explaining the dependent variable (13). The significance of each predictor in the final model was assessed using p-values. The model fit was evaluated using R-squared and adjusted R-squared values. Finally, regression coefficients (βi's) were interpreted to understand the impact of each weather parameter on rice blast incidence.

### Multiple Linear Regression (MLR)

MLR models were developed using the 'lm' function in R. The initial model included all weather variables and their time-lagged versions. Model assumptions (linearity, homoscedasticity, normality of residuals and absence of multicollinearity) were checked and addressed as necessary (14).

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_nX_2 + \varepsilon$$

Where: Y = Disease severity (dependent variable), $X_1$, $X_2$, ..., $X_n$ = Weather parameters (independent variables), $\beta_0$, $\beta_1$, $\beta_2$, ..., $\beta_n$ = Regression coefficients, ε = Error term

### ARIMA Model

ARIMA models were developed using the 'auto.arima' function from the 'forecast' package in R (15). The function automatically selects the optimal ARIMA parameters (p, d, q) based on the AIC (16). Weather variables were incorporated as external regressors.

$$\text{ARIMA(p,d,q)}: \Phi(B)(1-B)^dY_t = \theta(B)\varepsilon_t$$

Where: B = Backshift operator, Φ(B) = Autoregressive operator of order p, $(1-B)^d$ = Differencing operator of order d, θ(B) = Moving average operator of order q, $Y_t$ = Observed value at time t, $\varepsilon_t$ = Error term at time t.

### Model Validation and Comparison

Models were validated using a k-fold cross-validation approach (k=5). The dataset was split into training (80 %) and testing (20 %) sets. Model performance was evaluated using the following metrics: Root Mean Square Error (RMSE), which measures the average magnitude of error and gives higher weight to larger errors; Mean Absolute Error (MAE), which calculates the average absolute differences between predicted and actual values, treating all errors equally; R-squared ($R^2$), which indicates the proportion of variance explained by the model and the Akaike Information Criterion (AIC), which assesses the model's fit while penalizing for complexity to avoid overfitting. Additionally, residual diagnostics were performed to ensure model assumptions were met.

### Sensitivity Analysis

Sensitivity analysis was conducted to assess the relative importance of each weather parameter in predicting disease severity. This was done using the Sobol method with the 'sensitivity' package in R (17), employing both local and global sensitivity analysis methods.

### Statistical Analysis

All statistical analyses were performed using R software (version 4.1.0). Significance levels were set at α=0.05 for all tests. Graphical representations of results were created using the 'ggplot2' package (18). The flow chart was made to represent the stepwise process taken during the data analysis (Fig. 2).

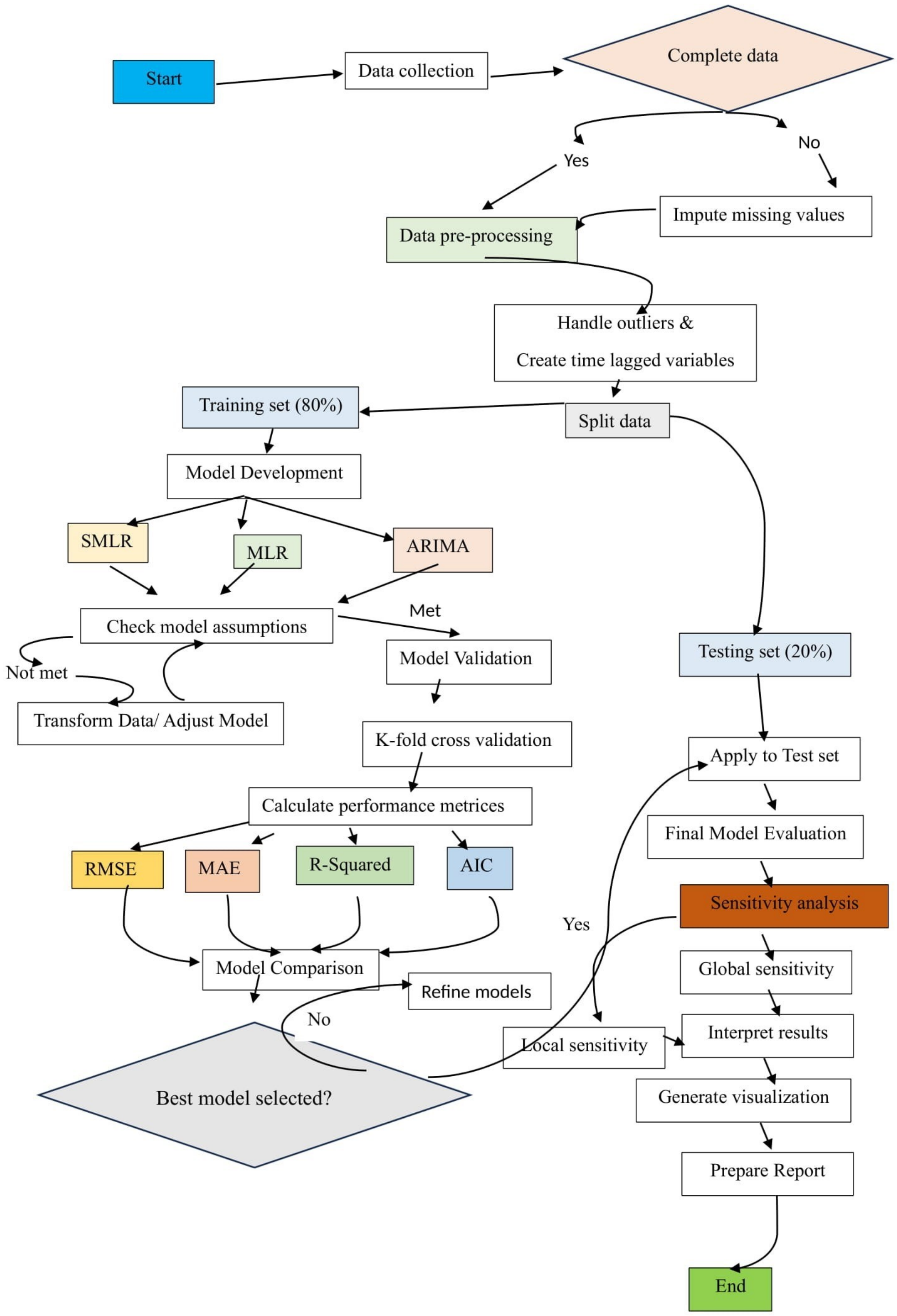


**Fig. 2.** The flow chart representing the step in data analysis.

## Results

### Blast severity occurrence over years

The radial bar plot was generated to visualize the weekly occurrences of blast. Data was collected, processed and visualized using R and the ggplot2 package. The resulting plot clearly indicates the frequency of blasts per week, with colour coding representing different levels of occurrences-darker shades indicated severe incidences, while lighter shades represent mild incidences. This method provides an effective way to identify patterns and trends in the data. The radial plot with weekly incidence of rice blast from 2017 to 2021 revealed a distinct seasonal pattern. Peak incidence was observed during mid-November to late-January (SMW 46 to 3), with severe levels of the disease. This period corresponds to the cooler months, when lower temperatures combined with higher relative humidity create favourable conditions for the fungal pathogen to thrive. High incidence periods were noted in late-January to mid-February (SMW 4 to 8) and late-September to mid-November (SMW 38 to 45), likely due to similar weather patterns with optimal moisture levels for disease development. Moderate incidence occurred during late-June to mid-September (SMW 26 to 38) and late-February to early-March (SMW 8 to 9), when fluctuating weather conditions likely moderated disease spread. Very low incidence was recorded from mid-April to mid-June (SMW 15 to 25) and there was no incidence from early-March to early-April (SMW 10 to 14), periods typically characterized by higher temperatures and lower humidity, which are less conducive to the growth of the pathogen. These findings suggest a strong seasonal influence on the occurrence of rice blast in Coimbatore (Fig. 3).

## Seasonal occurrence of rice blast over years

A seasonal and time series plot created using data from January 2017 to December 2023 showed an overall increasing trend in weekly rice blast incidence. Seasonal fluctuations were evident, with peaks typically increasing at the beginning and end of each year (Fig. 4A) and a typical peak occurrence was found at end of each year as shown in time series plot (Fig. 4B). A LOESS regression smoothing line highlighted this positive trend in both plots, aiding in the interpretation of seasonal patterns (Fig. 4).

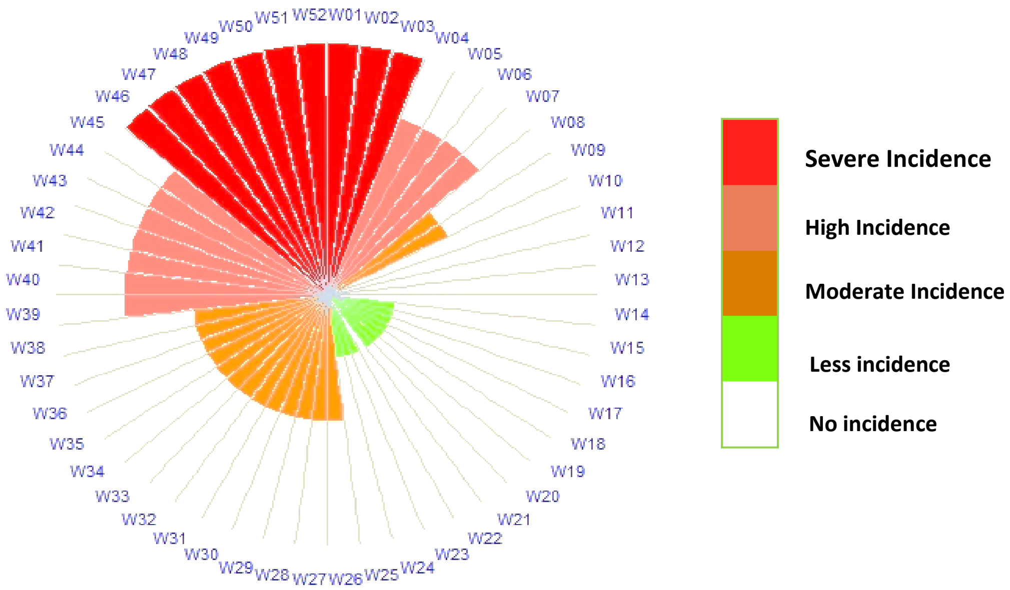


**Fig. 3.** Weekly occurrence of rice blast over years (2017-2021).

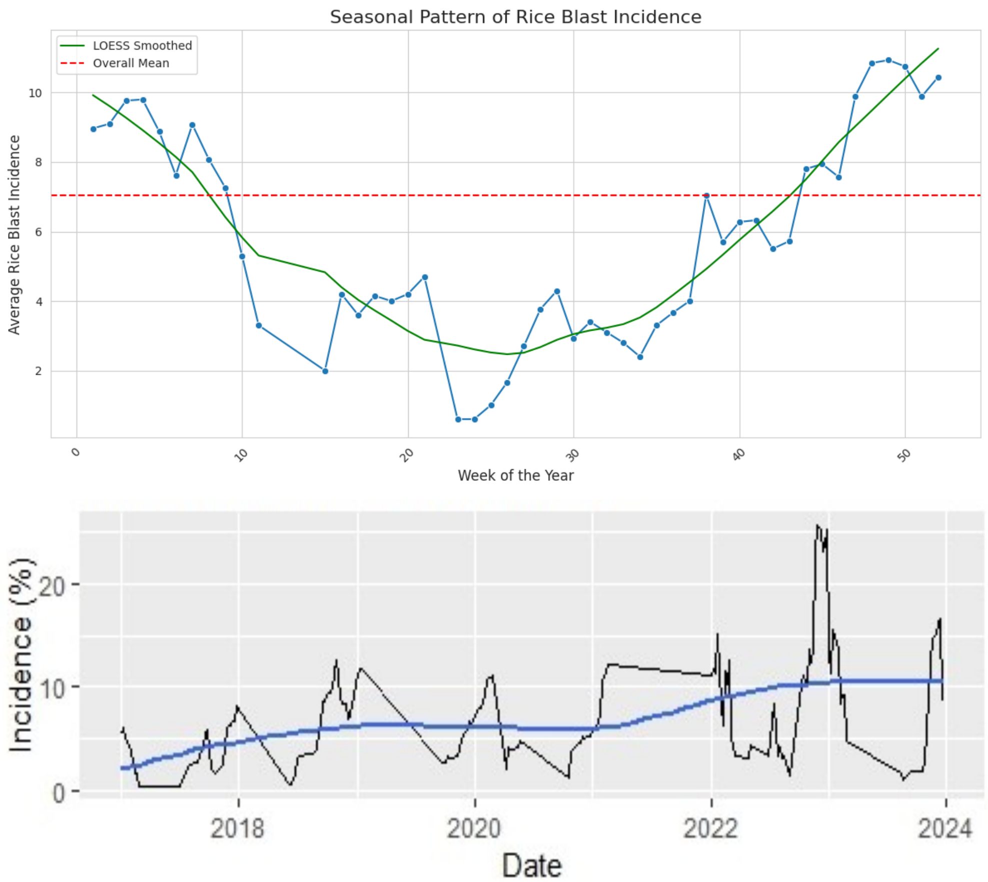


**Fig. 4.** Seasonal pattern and time series plot of rice blast incidence.

## Descriptive statistics

The descriptive statistics of the dependent variable (rice blast incidence) and exogenous variable (weather parameters) were calculated and presented in Table 2 were self-explanatory. The data under consideration were highly heterogeneous in nature, with rice blast incidence varying from 0.3 to 31.3, leading to a high coefficient of variation (CV) and abnormality in data distribution as indicated by skewness and kurtosis values out of the normal range.

## Model Development and Performance

Three regression models-Stepwise Regression, Multiple Linear Regression (MLR) and ARIMA-were developed to predict the severity of rice blast disease using the weather parameters as predictors.

## Stepwise Regression

The stepwise regression model selected the following significant predictors: Minimum Temperature (X2), Relative Humidity (morning) (X3), Evaporation (X7) and Leaf Wetness (X10) given in Table 3. The model achieved an R-squared value of 0.55 and an adjusted R-squared value of 0.54 (Table 6, Fig. 5).

## Multiple Linear Regression (MLR)

The MLR model included all weather variables and their time-lagged versions. The highly significant and influencing predictors were Minimum Temperature (X2), Relative Humidity (morning) (X3), Evaporation (X7) and Leaf Wetness (X10) given in Table 4. The model assumptions were checked and addressed. The MLR model achieved an R-squared value of 0.56 (Table 6, Fig. 6).

## ARIMA Model

The ARIMA model with parameters (p=3, d=0, q=0) was selected based on the AIC. The model incorporated weather variables as external regressors. The performance metrics for the ARIMA model were given in Table 5, Fig. 7.

## Model Validation and Comparison

The models were validated using a k-fold cross-validation approach (k=5). The performance metrics for each model were summarized below:

## Sensitivity Analysis of Weather Parameters on Rice Blast Severity

Sensitivity analysis was conducted using the Sobol method to determine the relative importance of various weather parameters in predicting rice blast disease severity. The analysis considered ten weather parameters: maximum temperature (Max), minimum temperature (Min), relative humidity 1 (RH1), relative humidity 2 (RH2), rainfall (RF), wind speed (Wind), evaporation (EVP), sunshine hours (SS), solar radiation (SR) and leaf wetness (LW). Table 7 presents the first-order (S) and total (T) sensitivity indices for each parameter. The first-order indices represent the direct effect of each parameter on rice blast severity, while the total indices account for both direct effects and interactions with other parameters. Solar radiation (SR) emerged as the most influential parameter, with a first-order index of 0.7803 and a total index of 1.0168, indicating that it accounts for approximately 78 % of the variance in rice blast severity directly and over 100 % when including interactions with other parameters. The second most important factor was

**Table 2.** Summary statistics of weather parameters on rice blast.

| | Mean | Std. Error | Co-var | Skewness | Kurtosis | Min | Max |
|---|---|---|---|---|---|---|---|
| **Rice blast** | 7.343 | 0.437 | 0.779 | 1.773 | 3.993 | 0.3 | 31.3 |
| **Max** | 30.691 | 0.159 | 0.068 | -0.547 | 4.305 | 19.429 | 35.6 |
| **Min** | 21.946 | 0.171 | 0.102 | -0.632 | -0.143 | 15.329 | 25.957 |
| **RH1** | 86.211 | 0.305 | 0.046 | -0.586 | 0.412 | 72.714 | 95.143 |
| **RH2** | 64.631 | 0.589 | 0.119 | 0.371 | 0.071 | 50 | 89.714 |
| **RF** | 2.438 | 0.341 | 1.831 | 2.962 | 10.069 | 0 | 26.714 |
| **Wind** | 6.004 | 0.191 | 0.417 | 2.097 | 5.434 | 2.357 | 17.686 |
| **EVP** | 4.593 | 0.099 | 0.281 | 0.171 | -0.634 | 2.015 | 7.486 |
| **SS** | 5.412 | 0.168 | 0.406 | 0.23 | -0.999 | 2.043 | 10.171 |
| **SR** | 333.306 | 4.729 | 0.186 | 0.598 | -0.361 | 250.029 | 543.314 |
| **LW** | 5.989 | 0.259 | 0.565 | 0.235 | -1.303 | 0.728 | 11.942 |

**Table 3.** Stepwise regression for weather parameters influencing rice blast severity.

| Predictor | Coefficient (β) | Std. Error | t-value | p-value |
|---|---|---|---|---|
| **Intercept** | -17.36127 | 10.39884 | -1.670 | 0.09690 |
| **Min** | -0.56361 | 0.17882 | -3.152 | 0.00193 |
| **RH1** | 0.45830 | 0.09581 | 4.784 | $3.78e^{-06}$ |
| **EVP** | -0.99361 | 0.25072 | -3.963 | 0.00011 |
| **LW** | 0.35504 | 0.12141 | 2.924 | 0.00394 |

**Table 4.** Multiple linear regression (MLR) for weather parameters influencing rice blast severity.

| Predictor | Coefficient (β) | Standard Error | t-value | p-value |
|---|---|---|---|---|
| **Intercept** | -38.619330 | 14.818506 | -2.606 | 0.01002 |
| **Max** | 0.049216 | 0.189357 | 0.260 | 0.79527 |
| **Min** | -0.516402 | 0.194661 | -2.653 | 0.00879** |
| **RH1** | 0.608490 | 0.121703 | 5.000 | 1.49e-06*** |
| **RH2** | 0.047074 | 0.042518 | 1.107 | 0.26989 |
| **RF** | -0.036754 | 0.072839 | -0.505 | 0.61454 |
| **Wind** | 0.223811 | 0.170880 | 1.310 | 0.19216 |
| **EVP** | -1.245951 | 0.319323 | -3.902 | 0.00014*** |
| **SS** | 0.084454 | 0.190065 | 0.444 | 0.65739 |
| **SR** | 0.006209 | 0.007081 | 0.877 | 0.38191 |
| **LW** | 0.371669 | 0.126661 | 2.934 | 0.00383** |

**Table 5.** ARIMA model for weather parameters influencing rice blast severity.

| Model | Parameters | Estimation | S.E. | Z Value | Probability | Model fitting | | Box-Pierce Non-correlation Test | |
|---|---|---|---|---|---|---|---|---|---|
| | ar1 | 0.380489 | 0.090492 | 4.2047 | <0.01 | | | | |
| | ar2 | -0.019755 | 0.096427 | -0.2049 | 0.837 | | | | |
| | ar3 | 0.299391 | 0.084551 | 3.5410 | <0.01 | Log likelihood | -164.26 | Original | Residual |
| | Intercept | 7.094515 | 0.204825 | 34.6369 | <0.01 | | | | |
| | Max | 1.966487 | 0.077294 | 25.4418 | <0.01 | | | | |
| | Min | -0.565207 | 0.085101 | -6.6416 | <0.01 | | | | |
| ARIMA (3,0,0) | RH1 | 0.886704 | 0.084097 | 10.5438 | <0.01 | | | | |
| | RH2 | -0.233276 | 0.089903 | -2.5947 | 0.009 | | | | |
| | RF | -0.262035 | 0.119394 | -2.1947 | 0.02 | | | | |
| | Wind | 1.036248 | 0.125351 | 8.2668 | <0.01 | AIC | 358.52 | $X^2$=118.17 ($p$<0.01) | $X^2$= 0.015372 ($p$=0.9013) |
| | EVP | -3.004875 | 0.163512 | -18.3770 | <0.01 | | | | |
| | SS | -0.628195 | 0.136667 | -4.5965 | <0.01 | | | | |
| | SR | -1.307819 | 0.152168 | -8.5946 | <0.01 | | | | |
| | LW | -1.169051 | 0.149401 | -7.8249 | <0.01 | | | | |

**Table 6.** Model validation.

| Model | Equation | RMSE | MAE | $R^2$ | AIC |
|---|---|---|---|---|---|
| SMLR | $Y=-17.4-0.5X_2+0.5X_3-0.9X_7+0.3X_{10}$ | 3.81 | 2.47 | 0.55 | 954.71 |
| MLR | $Y=-38.6 + 0.04X_1 - 0.5X_2 + 0.6X_3 + 0.04X_4 - 0.03X_9 + 0.2X_5 -1.2X_6 + 0.08X_7 + 0.01X_8 + 0.37X_{10}$ | 3.75 | 2.45 | 0.57 | 961.66 |
| ARIMA | $Yt =7.09+0.38Yt\text{-}1-0.01\ Yt\text{-}2+0.29\ Yt\text{-}3+\beta 1X1+\beta 2X2+..+\beta 10X10+\epsilon t$ | 2.39 | 1.74 | 0.92 | 364.12 |

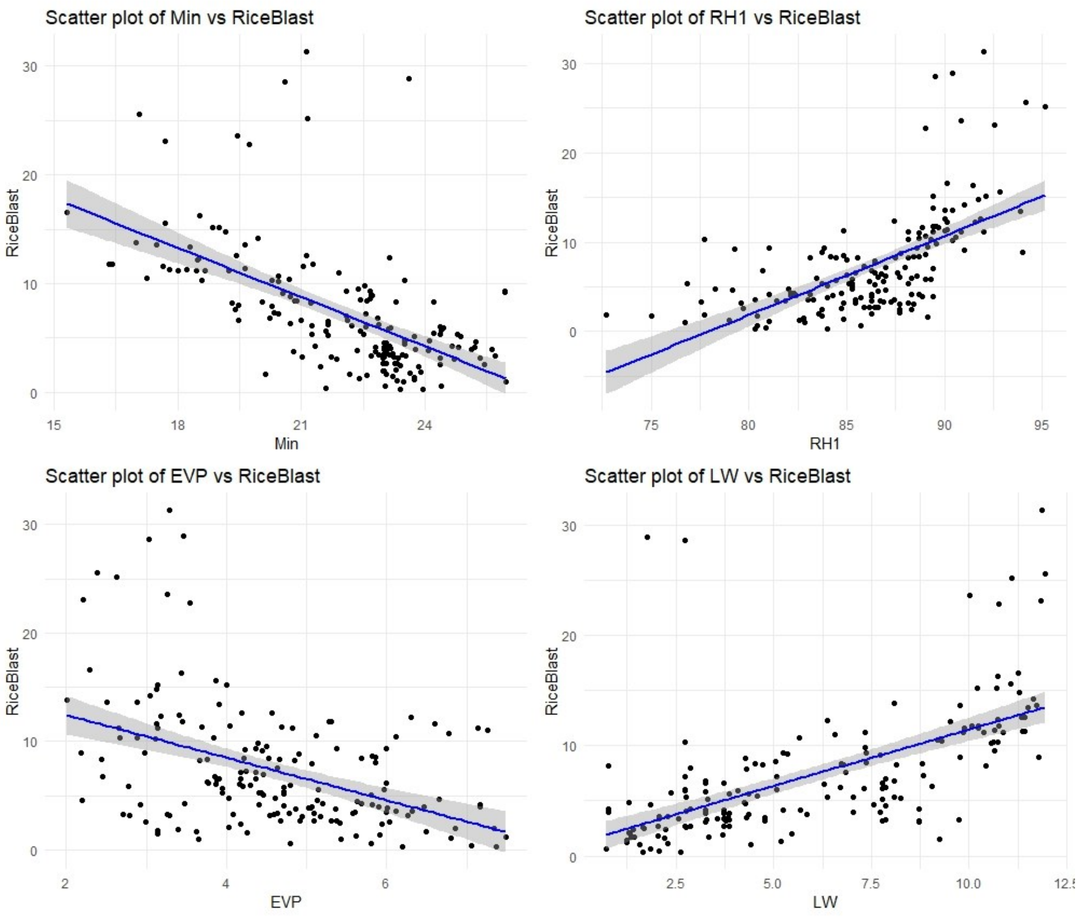


**Fig. 5.** Scatter plot showing weekly incidence of rice blast through correlation with the significant predictors obtained from stepwise regression.

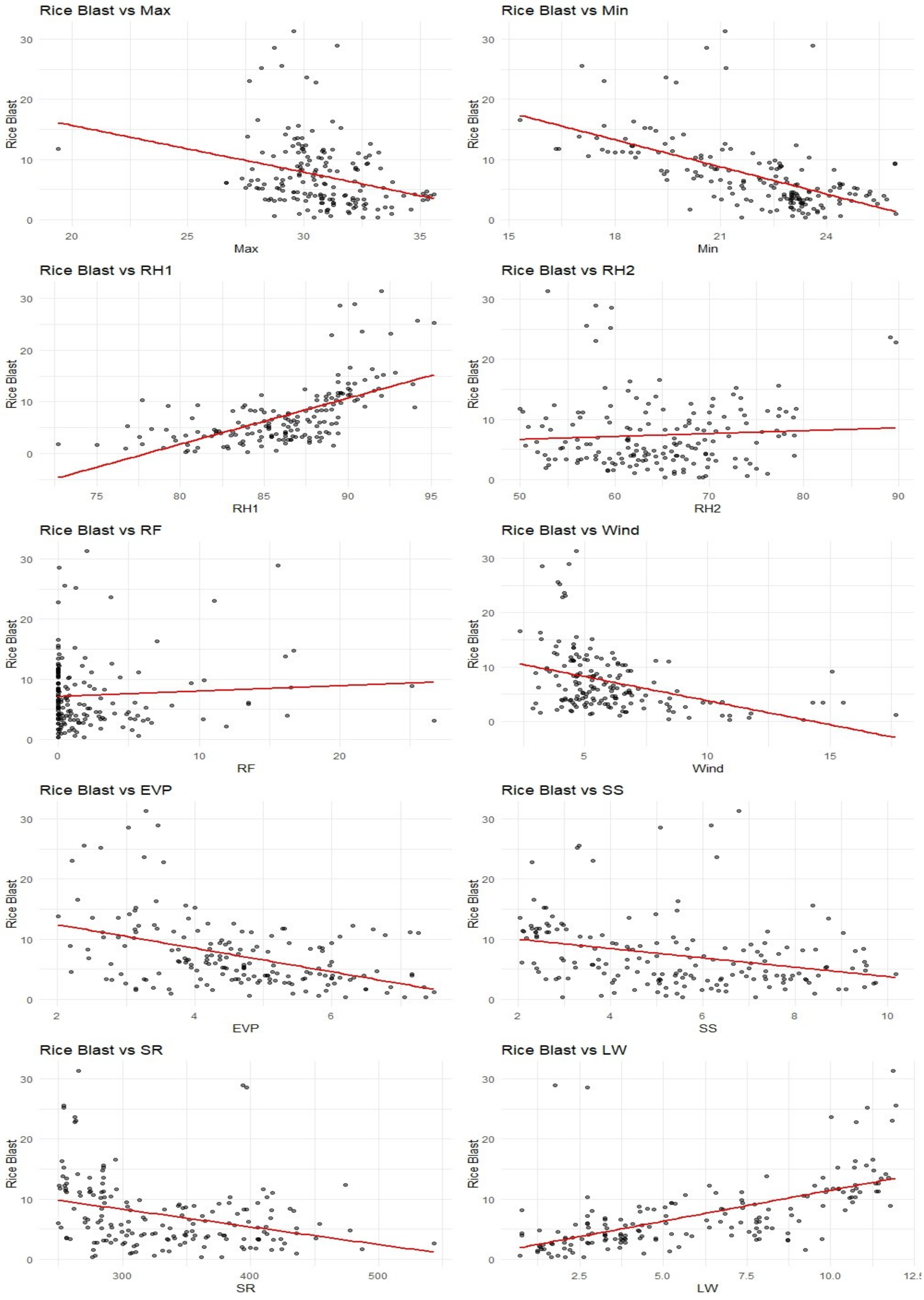


**Fig. 6.** Scatter plot showing weekly incidence of rice blast through correlation with the significant predictors obtained from multiple linear regression.

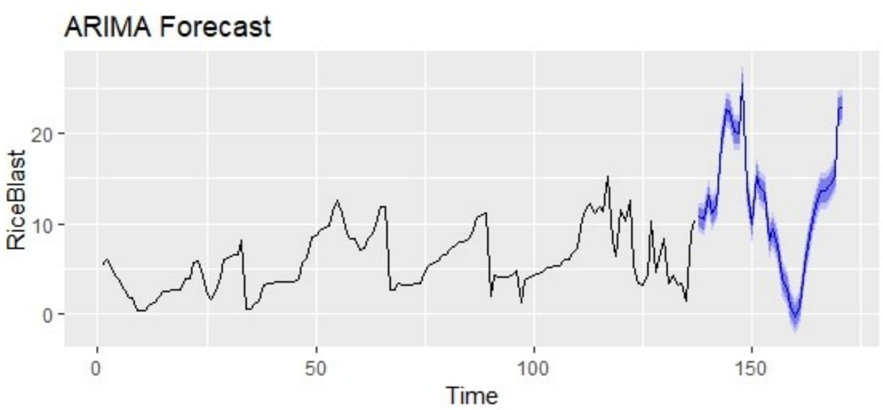


**Fig. 7.** ARIMA Forecast (1,1,1) of rice blast over years. The black line indicates the training set and blue line indicated the future forecast.

**Table 7.** Sobol sensitivity indices for weather parameters influencing rice blast severity.

| Parameter | First-order Index (S) | Total Index (T) |
|---|---|---|
| SR | 0.7803 | 1.0168 |
| RH2 | 0.0312 | 0.0312 |
| RH1 | 0.0534 | -0.0092 |
| Min | 0.0074 | -0.0046 |
| Max | 0.0055 | 0.0009 |
| LW | 0.0030 | -0.0017 |
| EVP | 0.0015 | -0.0009 |
| SS | 0.0008 | -0.0000 |
| RF | -0.0032 | 0.0070 |
| Wind | -0.0009 | 0.0025 |

relative humidity (RH2), with a first-order index of 0.1024 and a total index of 0.0312. Other parameters, including rainfall, wind speed and temperature, showed relatively small effects. Evaporation, sunshine hours and leaf wetness demonstrated minimal influence on rice blast severity prediction in this model. Interestingly, some parameters (RF, Wind) showed higher importance in their total indices compared to their first-order indices, suggesting that these factors might play a role through interactions with other parameters rather than direct effects. These results suggest that monitoring and managing solar radiation and relative humidity could be key factors in predicting and controlling rice blast disease. However, it's important to note that these findings are based on the specific predictive model used in this study and further research may be needed to validate these results across different environmental conditions and rice varieties.

## Discussion

This study investigated the relationship between weather parameters and Rice Blast disease severity in Coimbatore, India, over a 7 years period (2017-2023). The findings reveal significant seasonal patterns and highlight the complex interactions between climatic factors and disease incidence.

Radial plot analysis showed a peak incidence of rice blast from mid-November to late January (SMW 46 to 3), consistent with the findings of other studies (8). Their study utilized an artificial intelligence-based model to predict rice blast disease in South Korea, incorporating climatic data such as air temperature, relative humidity and sunshine hours, known to influence plant disease seasonal patterns. This period corresponds to the winter months in South Korea, suggesting that climatic conditions during this time may be conducive to the proliferation of the rice blast fungus. Specifically, *Magnaporthe oryzae* thrives in moderate temperatures, typically between 20 °C and 28 °C, combined with high humidity levels, which are characteristic of this season. The cooler temperatures can slow the plant's metabolic processes, weakening its natural defences against pathogens, while high relative humidity provides the moisture necessary for spore germination and subsequent infection. Furthermore, leaf wetness, caused by dew formation during these months, facilitates fungal colonization, as it allows the spores to adhere to the leaf surface more easily. Reduced solar radiation during the winter months decreases the drying effect on plant surfaces, maintaining moisture levels that promote fungal growth. These environmental conditions create an optimal environment for the pathogen's development, leading to increased infection rates. Seasonal and time series plots indicate an overall increasing trend in weekly rice blast incidence over the study period, emphasizing the importance of considering seasonal variations in disease management strategies. Similarly, other studies highlight the complexity of seasonal patterns and their varying impacts on different agricultural diseases, underscoring the need for targeted interventions during peak risk periods (19).

Regression models and sensitivity analysis identified several key weather parameters influencing rice blast severity. Solar radiation plays a critical role in influencing fungal growth and disease progression due to its direct effect on the biological and environmental conditions necessary for fungal survival. Solar radiation, particularly ultraviolet (UV) light, damages fungal cells by causing DNA mutations, impairing spore viability and hindering growth. These radiations also impact humidity and temperature, the two key factors that create an unfavourable environment for fungal pathogens. By reducing the relative humidity and increasing temperatures, solar radiation limits the moist conditions fungi need to thrive, thus significantly lowering disease severity. This aligns with the finding that solar radiation explains about 78 % of the variance in disease severity, as it directly disrupts fungal development. Solar radiation emerged as the most influential factor, accounting for approximately 78 % of the variance in disease severity (20). Other weather features such as average visibility, rainfall, sunshine hours, maximum wind speed and days of rain were identified as effective predictors for rice blast disease forecasting (21). Relative humidity, particularly evening relative humidity (RH2), was also identified as a significant predictor, supporting contemporary research on the importance of moisture in blast disease progression. Interestingly, the study found that minimum temperature had a negative relationship with disease severity in both stepwise and multiple linear regression models, indicating that lower temperatures generally favour blast disease development. This is consistent with the biological behaviour of *Magnaporthe oryzae*, the rice blast fungus, which thrives in cooler temperatures, typically between 20 °C and 28 °C. At lower temperatures, the plant's physiological processes slow down, reducing its natural defence mechanisms, such as the production of reactive oxygen species (ROS) and the activation of defence-related genes. Furthermore, cooler temperatures extend the leaf wetness duration, which is crucial for spore germination and successful fungal infection. Under these conditions, the fungus was able to proliferate more efficiently, producing more spores and infecting new plant tissue. Higher temperatures, by contrast, can inhibit fungal growth by accelerating evaporation, drying plant surfaces and reducing the humidity levels required for the pathogen to thrive. This temperature-dependent behaviour of the rice blast fungus has been widely observed in other studies, underscoring the importance of cooler, moist conditions in promoting the development of rice blast disease (22). Similarly, it was found that temperature and humidity significantly affect rice blast disease severity, with high severity associated with low temperature and high humidity (23). The scientist Ali developed a predictive model for lentil wilt severity based on meteorological variables, finding significant correlations between these variables and disease severity (24). Maximum temperature showed a negative correlation, while minimum temperature, rainfall and relative humidity exhibited positive correlations with lentil wilt severity.

The ARIMA model outperformed the other models in capturing the relationship between weather parameters and disease severity due to its ability to account for both the temporal trends and seasonality in the data. By modelling these time-dependent patterns, ARIMA captured the complex dynamics between weather variables and disease progression more effectively, resulting in the highest R-squared value (0.92) and lowest RMSE (2.39). This indicates that ARIMA was better at explaining variance and minimizing prediction errors compared to the other models. This suggests that time series approaches incorporating weather parameters as external regressors may be particularly effective for predicting rice blast severity. The strong performance of the ARIMA model indicates its potential utility in developing early warning systems for rice blast outbreaks, aligning with recent trends in agricultural forecasting. The forecasted *Botrytis cinerea* spores in Galicia and Northern Portugal using ARIMA models, combining meteorological and aerobiological parameters to provide useful tools for forecasting spore concentrations and reducing infection risks, forming the basis for a modern integrated grapevine pest-management strategy (25).

The sensitivity analysis provided valuable insights into the relative importance of different weather parameters. The high influence of solar radiation and relative humidity suggests that these factors should be prioritized in monitoring and predictive efforts. However, the complex interactions between parameters, as indicated by the differences between first-order and total sensitivity indices for some variables (e.g., rainfall and wind speed), underscore the need for comprehensive models that account for these interactions, as emphasized in recent literature.

While this study provides valuable insights, it has some limitations. The data were collected from a single location, which may limit the generalizability of the findings to other rice-growing regions. Future research should consider multi-location studies to validate these results across different agro-climatic zones. Additionally, the study focused on weather parameters, but other factors such as rice variety susceptibility, soil conditions and management practices also influence blast disease development (20, 26, 27). Incorporating these factors into future models could further improve predictive accuracy and provide more comprehensive disease management recommendations.

## Conclusion

This study demonstrates the significant influence of weather parameters, particularly solar radiation and relative humidity, in driving rice blast disease severity in Coimbatore. The ARIMA model's superior performance in predicting disease outbreaks provides a valuable tool for timely and targeted disease management strategies, which could help minimize crop losses and improve resource efficiency. Future research should aim to incorporate additional factors like soil conditions and rice variety susceptibility and extend to multi-location studies. Integrating machine learning techniques could further enhance predictive accuracy and bolster early warning systems for broader agricultural application.

## Acknowledgements

The authors express their heartfelt gratitude to the Department of Plant Pathology, Centre for Plant Protection studies and the Department of Rice at TNAU, Coimbatore, Tamil Nadu, India, for providing laboratory and free facilities essential for this research.

## Authors' contributions

AGM: Conceptualization, data curation, formal analysis, methodology, software, writing original draft, investigation, writing - original draft. CG: Conceptualization, methodology, supervision, writing- and editing. AK: Conceptualization, formal analysis, supervision, writing - review and editing. SGP: Conceptualization, software, supervision, writing - review and editing. JR: Supervision, visualization. NKS: Supervision, visualization. SM: Supervision, visualization.

## Compliance with ethical standards

**Conflict of Interests:** The authors declare that there are no conflicts of interest related to this article.

**Ethical issues:** None.

**Declaration of generative AI and AI-assisted technologies in the writing process**

During the preparation of this work the author(s) not used AI tools and the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## References

1. Muthayya S, Sugimoto JD, Montgomery S, Maberly GF. An overview of global rice production, supply, trade and consumption. Annals of the New York Academy of Sciences. 2014;1324:7-14. https://doi.org/10.1111/nyas.12540
2. Mohanty S. Trends in global rice consumption. Rice Today. 2013;12:44-45.
3. Wilson RA, Talbot NJ. Under pressure: Investigating the biology of plant infection by *Magnaporthe oryzae*. Nature Reviews Microbiology. 2009;7:185-95. https://doi.org/10.1038/nrmicro2032
4. Groth DE. Azoxystrobin rate and timing effects on rice yield and rice disease control. Agronomy Journal. 2006;98:389-94. https://doi.org/10.1094/PD-90-1055

5. Kumar A. Forewarning models for *Alternaria* blight in mustard (*Brassica juncea*) crop. Indian Journal of Agricultural Sciences. 2013;81:116-18. https://doi.org/10.54386/jam.v24i1.782

6. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: A review. Sensors. 2018;18:2674. https://doi.org/10.3390/s18082674 PubMed: 30110960

7. Nettleton DF, Katsantonis D, Kalaitzidis A, Sarafijanovic-Djukic N, et al. Predicting rice blast disease: Machine learning versus process-based models. BMC Bioinformatics. 2019;20:514. https://doi.org/10.1186/s12859-019-3065-1 PubMed: 31640541

8. Kim Y, Roh JH, Kim HY. Early forecasting of rice blast disease using long short-term memory recurrent neural networks. Sustainability. 2017;10:34. https://doi.org/10.3390/su10010034

9. Bebber DP, Ramotowski MAT, Gurr SJ. Crop pests and pathogens move polewards in a warming world. Nature Climate Change. 2013;3:985-88. https://doi.org/10.1038/nclimate1990.

10. Hossain M, Ali MD, Hossain D. Occurrence of blast disease in rice in Bangladesh. American Journal of Agricultural Science. 2017;4:74-80. http://www.aascit.org/journal/ajas

11. Stekhoven DJ, Buhlmann P. Miss_Forest-non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28 (1):112-18. https://doi.org/10.1093/bioinformatics/btr597.

12. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19(6):716-23. https://doi.org/10.1109/TAC.1974.1100705.

13. Liu ZY, Huang JF, Shi JJ, Tao RX, et al. Characterizing and estimating rice brown spot disease severity using stepwise regression, principal component regression and partial least-square regression. Journal of Zhejiang University Science B. 2007:(10):738-44. https://doi.org/10.1631/jzus.2007.B0738.

14. Neetoo H, Chuttur Y, Nazurally A, Takooree S, Mamode Ally N. Crop disease prediction using multiple linear regression modelling. International Conference on Soft Computing and its Engineering Applications. 2021;312-26. https://doi.org/10.1007/978-3-031-05767-0_25

15. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software. 2008;27(1):1-22. https://doi.org/10.18637/jss.v027.i03

16. Ling AS, Darmesah G, Chong KP, Ho CM. Application of ARIMAX model to forecast weekly cocoa black pod disease incidence. Mathematics and Statistics. 2019;7(4):29-40. https://doi.org/10.13189/ms.2019.070705

17. Pujol G, Iooss B, Janon A. With contributions from Laurent T. Package ‘sensitivity’. R Package Version 1.61. 2015. http://cran.nexr.com/web/packages/sensitivity/sensitivity.pdf

18. Wickham H, Chang W. ggplot2: Create elegant data visualisations using the grammar of graphics. Version 2.1. Journal of Statistical Software. 2016;189:1-89. https://doi.org/10.1007/978-3-319-24277-4.

19. Yohana JM, Yunxian Y. Analysis of seasonal pattern and variation of rice prices in Tanzania. Asian Journal of Agricultural Extension Economics and Sociology. 2019;37(3):1-10. https://doi.org/10.9734/ajaees/2019/v37i330267

20. Rotem J, Aust HJ. The effect of ultraviolet and solar radiation and temperature on survival of fungal propagules. Journal of Phytopathology. 1991;133(1):76-84. https://doi.org/10.1111/j.1439-0434.1991.tb00139.x

21. Sriwanna K. Weather-based rice blast disease forecasting. Computers and Electronics in Agriculture. 2022;193. https://doi.org/10.1016/j.compag.2022.106685

22. Greer CA, Webster RK. Occurrence, distribution, epidemiology, cultivar reaction and management of rice blast disease in California. Plant Disease. 2001;85(10):1096-102. https://doi.org/10.1094/PDIS.2001.85.10.1096

23. Kirtphaiboon S, Humphries U, Khan A, Yusuf A. Model of rice blast disease under tropical climate conditions. Chaos Solitons and Fractals. 2021;143. https://doi.org/10.1016/j.chaos.2020.110530

24. Ali Y, Qin A, Aatif HM, Ijaz M, Khan AA, Ahmad S, et al. A stepwise multiple regression model to predict *Fusarium* wilt in lentil. Meteorological Applications. 2022;29. https://doi.org/10.1002/met.2088

25. Fernandez-Gonzalez M, Rodríguez-Rajo FJ, Jato V, Aira MJ, et al. Forecasting ARIMA models for atmospheric vineyard pathogens in Galicia and Northern Portugal: *Botrytis cinerea* spores. Annals of Agricultural and Environmental Medicine. 2012;19:255-62.

26. Atugala DM, Deshappriya N. Effect of endophytic fungi on plant growth and blast disease incidence of two traditional rice varieties. Journal of the National Science Foundation of Sri Lanka. 2015;43(2). https://doi.org/10.4038/JNSFSR.V43I2.7945

27. Pajor R, Falconer R, Hapca S, Otten W. Modelling and quantifying the effect of heterogeneity in soil physical conditions on fungal growth. Bio-geosciences. 2010;7(11):3731-40. https://doi.org/10.5194/bg-7-3731-2010