GMR
The Original

WEATHER-DRIVEN EPIDEMIOLOGICAL MODELLING FOR PREDICTION OF BACTERIAL LEAF BLIGHT OF RICE IN THE STATE OF TAMIL NADU

N. Revathy[1*], Aishwarya S[1] and R. Thilagavathi[2]

[1]Department of Plant Pathology, Agriculture College and Research Institute, TNAU, Madurai
[2]Department of Plant pathology, KVK, Needhamangalam
***Corresponding author:** N. Revathy, nrevathy@tnau.ac.in

ABSTRACT
Bacterial Leaf Blight (BLB) incidence in rice is a major factor, delimiting rice yield. The pathogen, Xanthomonas oryzae pv oryzae is known to cause huge losses both during the initial stages and later stages of the crop. In order to manage the disease, a timely detection is necessary. BLB in rice is highly influenced by the prevailing weather factors. Many studies have reported, temperature, relative humidity, rainfall and wind speed to affect the disease occurrence and severity. In this study, we have studied the effect of weather factors on the incidence of BLB in Madurai district of Tamil Nadu from 2015-2024. We observed maximum temperature, minimum temperature, morning and evening relative humidity and rainfall as the major weather factors affecting disease develop to formulate a multiple linear regression (MLR) model for forecasting the occurrence of the disease. While carrying out the statistical analysis we identified multicollinearity problems between two weather factors and based on the previous literatures, we decided to remove one of the correlated weather variable to develop the MLR model. The developed model exhibited a high R^2 of 0.917 and an adjusted R^2 of 0.851. This indicated that the model could explain 91.7% of the variation observed in BLB disease incidence in the region. The developed regression model provides a preliminary framework for weather-based prediction of rice BLB incidence in the study region. Including the major weather factors in the development of a disease prediction model will facilitate early assessment of disease risk and hence, facilitate timely disease management decisions.

KEYWORDS: Bacterial leaf blight (BLB), rice, prediction, Madurai, MLR

1. INTRODUCTION

Rice is a principal food crop in Tamil Nadu, that sustains rural livelihoods, households caloric consumption and the agricultural GDP (Arivelarasan et al., 2023). Seasonal rainfall, surface irrigation and monsoon timing across Tamil Nadu, shape its rice acreage (Sellaperumal et al., 2025). Cultivated area under rice ranges between 18-22 Lakh ha (Vasanthi et al., 2017). But rice is affected by a number of diseases like rice blast, brown spot, sheath blight and bacterial leaf blight amounting to huge yield losses (Shivappa et al., 2021). Amongst these, the BLB or rice has been reported to cause susbstantial yield losses of upto 74% in the tropical countries of Asia (Reddy et al., 1979). The disease is considered to be a vascular wilt disease and appears during the early stages of plant growth and also in the later stages (Reddy et al., 1979). The disease is caused by Xanthomonas oryzae pv oryzae that enters the leaf tissues through wounds or hydathodes (Wang et al., 2019). The pathogen forms network with xylem vessels of parenchyma cells and spreads within the plant (Yabe, 1987). The grain yield losses due to the infection have been reported to be upto 80% in India (Shekhar, Sinha, & Kumari, 2020).

The BLB disease incidence in rice is highly affected by climatic factors such as temperature, rainfall, humidity and wind movement (Shakoor, 2015). In a study conducted heavy rainfall, wind movement and a minimum temperature of 24^0C have been shown to favour the disease development (Naqvi et al., 2016). A number of other studies have reported the impact of weather factors in increasing the severity of the diseases (Rehman and Atiq, 2022 and Reddy et al., 2015). Considering the influence of weather factors on blb incidence in rice, a prediction model that can predict the occurrence of the disease before an epidemic development is essential.

Although several studies have developed weather-based prediction model for BLB of rice, the predictive relationship may vary with agroclimatic conditions, observation period and temporal resolution of data. The study hence aims to examine the suitability of the selected weather variables in predicting the disease incidence in the state of Tamil Nadu

2. MATERIALS AND METHODS

2.1. Location of the experiment

The experiment was conducted at the farm are in Agricultural college and Research Institute, Madurai, Tamil Nadu. Madurai region of Tamil Nadu, has a hot, semi-arid to seasonally dry tropical climate, with a pronounced hot period between March to May and rainfall that occurs between September to November (Dash et al., 2013; Surendran et al., 2017). The environmental variables, known to affect disease incidence of blb in rice were recorded for 10 years at Agricultural college and research institute, Madurai. The campus is located at 9° 58' 6.5496" N and 78° 12' 17.9521" E.

2.2. Field preparation

Rice plants were transplanted to the main fields after 30 days of sowing. The soil available for cultivation was clay loam black alluvium soil at Agricultural college and research institute, Madurai's central farm. The area of the field that was maintained for cultivation was ?. In order to reduce the effect of soil borne inoculum from the left-over stubbles of the previous season, the location of the field was changed every 2 seasons, in and around the central farm. The variety ADT-38, was used for cultivation all throughout the period of the experiment.

2.3. Disease assessment

Blb symptoms were noted as yellowish-white water-soaked streaks at the leaf margins (Rice, 1996). The percentage of disease incidence was calculated using the formula (Rice, 1996)

$$Percentage\ disease\ incidence(\%) = \frac{No.of\ blb\ plants\ infected}{Total\ no.of\ plants\ observed}\ x\ 100$$

2.4. Collection of weather data

The environmental variables, including maximum temperature, minimum temperature, rainfall, morning relative humidity and evening relative humidity were recorded for 10 years at Agricultural college and research institute, Madurai. Meteorological data of the period from November to March was considered for epidemiological studies. The selection of study period was based on the prevailing rice calendar for Tamil Nadu, in which the state is known to cultivate rice across 4-5 seasons in a year (Tamil Nadu Agricultural University, n.d.). The Navarai/Kodai season of rice cultivation in Tamil Nadu extends between December to March. In addition to this, the Samba/Thaladi season extends into February-March (Tamil Nadu Agricultural University, n.d.). Therefore, November-March period was selected to include the impact of weather from both the major seasons of rice cultivation in Madurai, Tamil Nadu. These variables were recorded from November to march each year at a regular interval of 15 days between each observation.

2.5. Statistical analysis and development of the disease prediction model

All the statistical analysis was carried out using the IBM SPSS statistics software version 27 and R software version 4.5.3.

Annual blb disease incidence was taken as the dependent variable and the weather factors, Maximum and minimum temperature, morning and evening relative humidity and rainfall were taken to be the independent variables during the statistical analysis. Karl Pearson's correlation analysis was conducted to determine the relation between disease incidence and each weather variable. Correlation amongst the weather variables were tested using the correlation analysis and multicollinearity statistics. Highly correlated weather variables were removed because it could affect the model's prediction capability. A multiple linear regression model was developed after retaining only the significant variables. The model validation and prediction capability was done using the already available weather data.

3. RESULTS

3.1. Descriptive statistics of bacterial leaf blight incidence and weather variables over the ten-year time period

3.2.

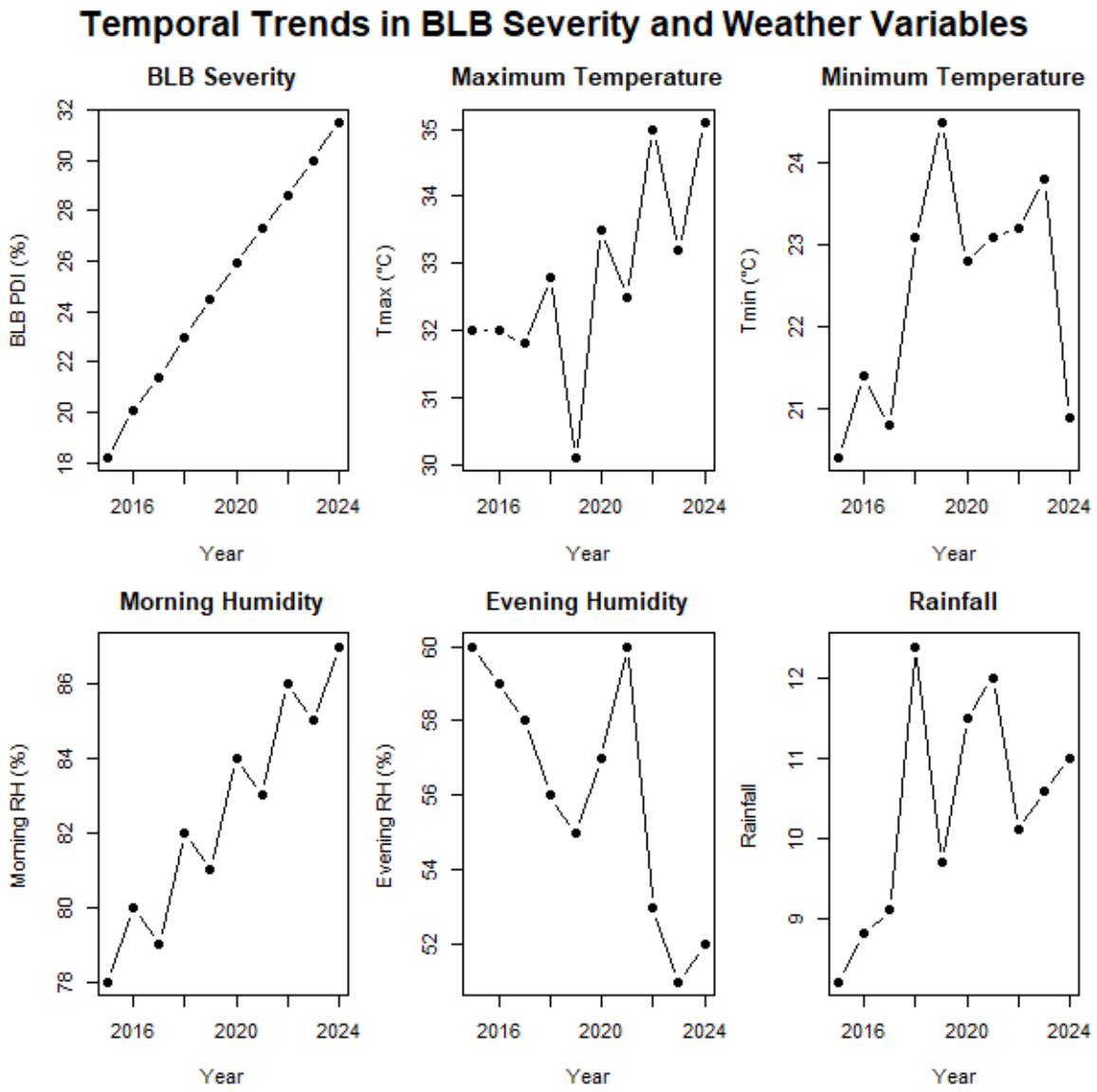
Table 1: Descriptive statistics of bacterial leaf blight incidence and weather variables

Weather factor	N	Minimum	Maximum	Mean	Std. Deviation
Annual disease incidence %	10	18.20	31.50	25.0500	4.39375
Maximum temperature	10	30.10	35.10	32.8000	1.50702
Minimum temperature	10	20.40	24.50	22.4000	1.41264
Morning Relative humidity	10	78.00	87.00	82.5000	3.02765
Evening Relative humidity	10	51.00	60.00	56.1000	3.28126
Rainfall (mm)	10	8.20	12.40	10.3400	1.40649

The annual blb incidence varied over the 10-year time period, with a mean disease incidence of 25% and a standard deviation of 4.39%, a maximum of 31.5% and a minimum of 18.20%. The mean maximum and minimum temperatures ranged between 32.800 ± 15^0C and 22.40 ± 1.41^0C respectively, while the mean morning and evening relative humidity ranged between 82.5 ± 3.02 % and 56.1 ± 3.28% and the mean rainfall was found to be 10.34 ± 1.40 mm during the 10-year time period.

3.3. Exploratory plots of BLB disease incidence percentage against time and weather variables

The plots generated exhibited the changes in disease incidence and weather variables over time.



The box plot analysis was performed to check for the presence of outliers in the data for disease incidence and weather variables.

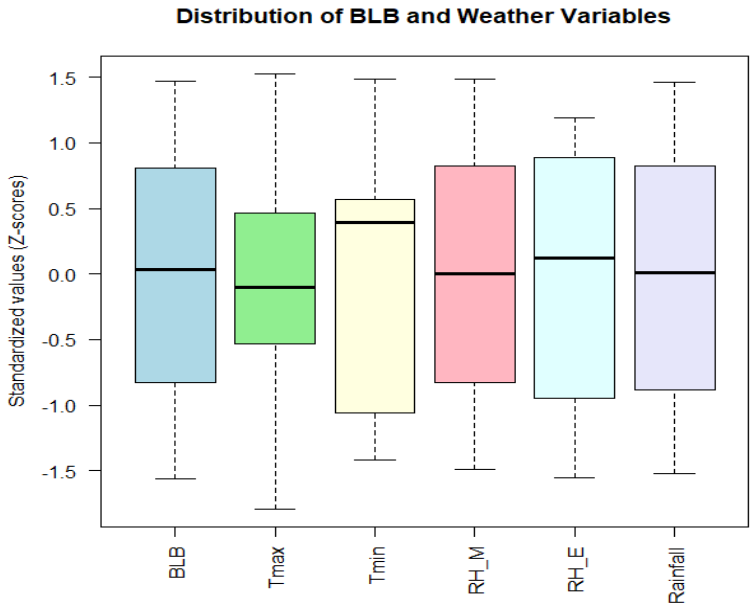


Figure 2: Box plot analysis of the statistical data. a. Annual disease incidence percentage (BLB) b. Maximum temperature (T_{max}) c. Minimum temperature (T_{mini}) d. Morning relative humidity (RH_M) e. Evening relative humidity (RH_E) f. Rainfall (mm)

The box plots didn't show any outliers indicating that no observation lied in the unusual extreme relative to the interquartile range of the respective variables.
The box plot analysis of minimum temperature revealed that the median was placed closer to the upper quartile, indicating an assymetric distribution of the central 50% observation. However, the shapiro-wilk test shoed no significant deviation from normality. Hence, the data for minimum temperature was considered to be in line with normal distribution despite the slight asymmetry observed in the box plot.
The values of p, obtained from the Shapiro-Wilk test was found to be >0.05 for all the factors, giving no evidence against normal distribution of data.

3.5. Correlation analysis between weather factors and disease severity percentage.

Pearson's correlation analysis was carried out to determine the associtation between bacterial leaf blight severity perncetage and the recorded weather factors.

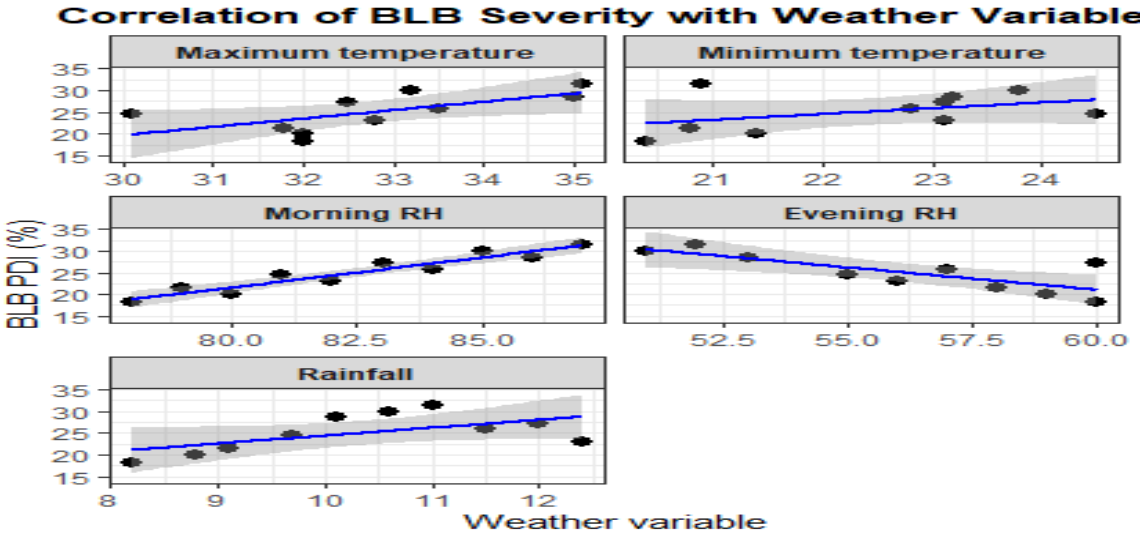


Figure3: Correlation between weather variables and blb disease incidence percentage over the 10 year time period

Correlation analysis		Annual disease incidence %	Maximum temperature	Minimum temperature	Morning Relative humidity	Evening Relative humidity	Rainfall (mm)
Annual disease severity %	Pearson Correlation	1	.656*	.428	.953**	-.773**	.582
Maximum temperature	Pearson Correlation	.656*	1	-.142	.806**	-.521	.384
Minimum temperature	Pearson Correlation	.428	-.142	1	.366	-.396	.487
Morning Relative humidity	Pearson Correlation	.953**	.806**	.366	1	-.766**	.603
Evening Relative humidity	Pearson Correlation	-.773**	-.521	-.396	-.766**	1	-.251
Rainfall (mm)	Pearson Correlation	.582	.384	.487	.603	-.251	1

Table 2: Correlation matrix of weather factors with disease incidence percentage

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
The annual disease incidence percentage can be seen rising linearly every year, during the 10-year time period. It can also be seen that the disease incidence percentage increases with maximum and minimum temperature, morning relative humidity and rainfall, while evening relative humidity had a negative influence on disease incidence percentage, suggesting a negative correlation between the two factors
Percent disease sevverity exhibited a strong positive correlation with maximum temperature ($r=0.656$; $p<0.01$), indicating that higher temperature was associated with increased disease severity. The morning relative humidity, also exhibited a significant positive correlation with disease severity percentage ($r= 0.953$, $p<0.01$). A significantly negative correlation was observed with evening relative humidity ($r = -0.773$, $p<0.01$), indicating that with the increase in evening relative humidity, disease severity decreases.

Amongst the weather varables significant positive correlation was seen between maximum temperature and morning relative humidity (r=0.806, p<0.01), and significantly negative correlation was observed between evening and morning relative humidity (r=0.076, p<0.01). These interrelationships indicate that the weather variables were not completely independent of each other and further tests of multicolinearity was needed before regression analysis.

3.6. Analysis of multicolinearity amongst the weather variables

Presence of multicolinearity amongst the weather factors was tested using the VIF (Variance inflation factor) and Tolerance analysis using the baseline rule (Marcoulides & Raykov, 2019)

Table 3: Standard baseline values for VIF and Tolerance to interpret multicollinearity

Multicolinearity studies with weather factors revealed that VIF and tolerance level were extremely high and low respctively for Maximum temperature (VIF = 7.993 and Tolerance = 0.12) and Morning relative humidity (VIF = 13.245 and Tolerance = 0.076).

Considering the correlation between maximum temperature and morning relative humidity (r = 0.806, p < 0.01) and the previous literatures on impact of weather factors on blb development, (Haque et al., 2022; Naqvi et al., 2016; Rehman & Atiq, 2022) we decided to remove maximum temperature from our regression analysis.

Table 5: Multicollinearity statistics amongst the weather variables after omitting maximum temperature

All the VIF values were <5 and tolerance values were above 0.20, suggesting a considerable decrease in multicollinearity amongst the weather variables.

Multicollinearity diagnostics		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	Minimum temperature	.635	1.576
	Morning Relative humidity	.217	4.611
	Evening Relative humidity	.289	3.461
	Rainfall (mm)	.403	2.482

VIF value	**Tolerance**	**Interpretation of multicolinearity**
<5	Above 0.20	Generally acceptable
5 to 10	0.10 to 0.20	Potential multicolinearity
>10	<0.10	Strong concern about multicolinearity

3.7. Development of an empirical regression model for the prediction of bacterial blight of rice

3.7.1. Model statistics

The new model statistics R square, adjusted R square, F value were generated using SPSS version 27 software.

Table 5: Model statistics of the revised regression model

R square	Adjusted R square	F value	Sig.
0.917	0.851	13.894	0.006

After the removal of the highly collinear predictor the revised multiple regression model was developed. The revised model expressed a slight decrease in R square and adjusted R square and overall F statistic compared with that of the initial model (R^2 = 0.955, Adj R^2 = 0.898, F (5,4) = 16.884 at p = 0.009). This suggests a slight decrease in explanatory power of the model. However, removal of the predictor decreased the VIF values considerably, with the VIF vlaues of all the predictor remaining below 5 and hence, reducing multicolinearity between the weather factors.

The model also showed a good fit to the obserrved data with an Rsquare of 0.917 and an adjusted R square of 0.851. It was able to explain 91.7% of the variation in disease incidence. The regression model was also found to be statistically significant (F(4,5) = 13.894, p= 0.006), indicating that the selected variable explained a significant proportion of variation in blb disease incidence. Hence, we suggest the use of the revised model to develop an empirical equation that can predict blb incidence in Tamil Nadu.

Regression model for the prediction of disease incidence

The regression model that we obtained after renoving the highly correlated weather variables is as follows:
$Y = -75.835 + 0.239\ (T_{min}) + 1.235(RH_m) + -0.119(RH_e) + 0.028\ (RF)$

Where,
Y= Percentage disease incidence
T_{min}= Minimum temperature
RH_m= Morning Relative humidity
RH_e= Evening Relative humidity
RF= Rainfall

3.8. Evaluation of Model Predictive performance

3.8.1. Calculation of predicted pdi and residual analysis from the developed regresion model

The predictive performance of the model developed was evaluated by comparing the observed disease incidence percentage with that predicted by the model during each year. Assessment of prediction error was done using the root mean square error (RMSE) value.

Table 6: Difference (Residual) between the predicted and the observed PDI

Year	Observed PDI	Predicted PDI	Residual
2015	18.20	18.49100	-.29100
2016	20.10	21.33582	-1.23582
2017	21.40	20.08431	1.31569
2018	23.00	24.66891	-1.66891
2019	24.50	23.81183	.68817
2020	25.90	26.92385	-1.02385
2021	27.30	25.41808	1.88192
2022	28.60	29.92577	-1.32577
2023	30.00	29.08533	.91467
2024	31.50	30.75511	.74489

The RMSE value was calculated using the R software version 4.5.3. It was found to be equal to 1.197534 PDI percentage points. This suggests that the difference between the observed pdi and the predicted value which was developed using the fitted regression model is very less indicating that the developed model could predict the occurrence of blb disease incidence with high accuracy.

3.8.2. Graphical validation of the developed regression model

The concurrence between the observed and predicted blb incidence was assessed using the scatter plot graph.

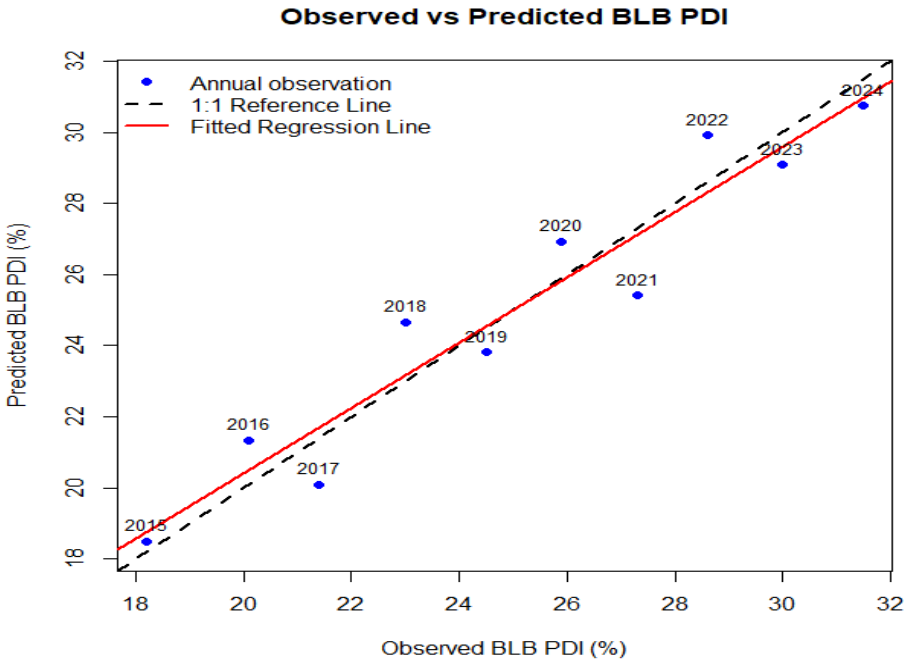


Figure 8: Scatter plot of observed and predicted blb incidence.

The proximity of the plotted points (blue colour) to the 1:1 reference line (dotted line) indicates the degree of agreement between the observed and the predicted values. Points close to the line indicate smaller prediction errors while those away from the line indicated the difference between the observed and the predicted values. Similarly, the values above and below the line indicate overprediction and underprediction respectively of disease incidence by the developed regression model. The distribution of the points in close proximity to the reference line provides a visual assessment of a good fit of the regression model to the observed data.

4. DISCUSSION

The developed regression model was identifed to predict the BLB disease occurrence with high accuracy. The weather variables included in the development of the model were minimum temperature, rainfall, morning and evening relative humidity which were found to be positively correlated with disease incidence except for evening relative humidity that showed a negative correlation. Though, maximum temperature was recorded during the 10 year time period but it had a significant correlation with morning relative humidity, thus presenting the problems with multicollinearity in development of a multiple linear regression model. Hence, based on the previous literature we decided to remove maximum temperature for the development of the multiple regression model. The developed model was able to predict the disease occurrence with high accuracy as was proved using within-data validation. In an earlier study it was found that the BLB incidence was positively correlated with relative humidity, rainfall and wind speed (Rehman & Atiq, 2022). Another study also found that the evening relative humidity has a negative correlation with blb disease incidence (Reddy et al., 2015) which was in concurrence with our findings. The study also developed a multiple linear regression model to check the influence of weather factors, minimum temperature, morning relative humidity, evening relative humidity, rainfall and maximum temperature. It found that these weather factors afftected 99% of disease development (Reddy et al., 2015). This justifies our use of the weather variables in the multiple linear regression model. A GIS-based Ordinary Least Square (OLS) spatial model has also been developed using the weather factors like relative humidity (RH), atmospheric pressure (A.P), minimum temperature alongwith soil factors to predict the occurrence of BLB incidence in rice. The developed model could only explain 65% of the variation in data (Ahmad et al., 2023). Weather based disease prediction models are based on the fact that the meteorological factors influence pathogen development, infection and disease progression. Temperature, relative humidity and rainfall are commonly incorporated as predictors because they influence key stages of plant disease development. Therefore, evaluating their relationships with BLB help in developing a weather based disease forecasting model (De Wolf & Isard, 2007). In this study, we have also attempted the same by developing a multiple linear regression model for the prediction of BLB disease incidence in rice. We suggest the use of this model in early prediction of BLB in Tamil Nadu and thus contribute to the effective management of the disease.

5. CONCLUSION

The prediction model proposed in the study is based on an observation period of 10 years for weather variables, including maximum and minimum temperature, morning and evening relative humidity and rainfall. ng bacterial leaf blight (BLB) incidence in rice using meteorological variables. The proposed model provides an initial quantitative framework for understanding the relationship between weather parameters and disease development and may serve as a basis for developing location-specific disease forecasting strategies. However, further validation using independent datasets collected across multiple years, locations, and environmental conditions is necessary to assess its reliability, and generalizability. Despite this limitation, the findings provide a valuable foundation for future research on weather-based disease prediction and demonstrate the potential of integrating meteorological data with statistical modelling to support timely disease management of BLB incidence of rice in Tamil Nadu.

Conflict of Interest : The authors declare that they have no conflict of interest regarding the publication of this paper.

REFERENCES

1. Ahmad, T., Haider, M., Randhir, T., Randhir, R., & Ahmad, S. (2023). Spatial analysis of mfactors influencing bacterial leaf blight in rice production. *Brazilian Journal of Biology*, *83*, e264249.
2. Arivelarasan, T., Manivasagam, V., Geethalakshmi, V., Bhuvaneswari, K., Natarajan, K., Balasubramanian, M., Gowtham, R., & Muthurajan, R. (2023). How far will climate change affect future food security? An inquiry into the irrigated rice system of Peninsular India. *Agriculture*, *13*(3), 551.
3. Dash, S., Saraswat, V., Panda, S., & Sharma, N. (2013). A study of changes in rainfall and temperature patterns at four cities and corresponding meteorological subdivisions over coastal regions of India. *Global and Planetary Change*, *108*, 175-194.
4. De Wolf, E. D., & Isard, S. A. (2007). Disease cycle approach to plant disease prediction. *Annu. Rev. Phytopathol.*, *45*(1), 203-220.
5. Haque, M. M., Mita, M. M., Bashar, S., Masud, M. M., Protic, I. A., Mozumdar, U. K., & Islam, M. R. (2022). Weather parameters largely regulate the outbreak of the rice bacterial blight: An endemic disease poses a threat to food security in Bangladesh. *Asian-Australasian Journal of Food Safety and Security*, *6*(1), 35-47.
6. Marcoulides, K. M., & Raykov, T. (2019). Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and psychological measurement*, *79*(5), 874-882.

7. Naqvi, S. A. H., Perveen, R., Rehman, A. U., Chohan, S., & Abbas, S. H. (2016). Bacterial leaf blight of rice: a disease forecasting model based on meteorological factors in Multan, Pakistan. *Journal of Agricultural Research (JAR). 54*(4), 707-718.
8. Reddy, A., Mackenzie, D., Rouse, D., & Rao, A. (1979). Relationship of bacterial leaf blight severity to grain yield of rice. *Phytopathology*, *69*(9), 967-969.
9. Reddy, B. R., Kumari, V. M., Naidu, G. M., Harathi, P., & Kumar, M. M. (2015). Statistical analysis and forecasting of rice bacterial leaf blight (BLB) based on climate factor in SPSR, Nellore District of Andhra Pradesh. *Journal of Agricultural Science*, *1*(1), 73-77.
10. Rehman, H., & Atiq, R. (2022). A disease predictive model based on epidemiological factors for the management of bacterial leaf blight of rice. *Brazilian Journal of Biology*, *84*, e259259.
11. Rice, I. N. f. G. E. o. (1996). *Standard evaluation system for rice*. IRRI, International Rice Research Institute.
12. Sellaperumal, P., Kaliaperumal, R., Dhanaraju, M., NS, S., P, S., S, S., Singaram, M., AP, S., Marimuthu, R., & Rangasamy, B. (2025). Time series analysis of Sentinel 1 A SAR data to retrieve annual rice area maps and long-term dynamics of start of season. *Scientific Reports*, *15*(1), 8202.
13. Shakoor, U., Saboor, A., Baig, I., Afzal, A., & Rahman, A. (2015). Climate variability impacts on rice crop production in Pakistan. *Pakistan Journal of Agricultural Research*, *28*(1).
14. Shekhar, S., Sinha, D., & Kumari, A. (2020). An overview of bacterial leaf blight disease of rice and different strategies for its management. *Int. J. Curr. Microbiol. App. Sci*, *9*(4), 2250-2265.
15. Shivappa, R., Navadagi, D. B., Baite, M. S., Yadav, M. K., Rathinam, P. S., Umapathy, K., Pati, P., & Rath, P. C. (2021). Emerging minor diseases of rice in India: losses and management strategies. In *Integrative advances in rice research*. IntechOpen.
16. Surendran, U., Kumar, V., Ramasubramoniam, S., & Raja, P. (2017). Development of drought indices for semi-arid region using drought indices calculator (DrinC)–a case study from Madurai District, a semi-arid region in India. *Water Resources Management*, *31*(11), 3593-3605.
17. Vasanthi, R., Sivasankari, B., Gitanjali, J., & Paramasivam, R. (2017). Efficiency Analysis of Paddy Production in Tank Irrigated Systems of Southern Zone in Tamil Nadu, India. *Int. J. Curr. Microbiol. App. Sci*, *6*(6), 1161-1167.
18. Wang, C., Tariq, R., Ji, Z., Wei, Z., Zheng, K., Mishra, R., & Zhao, K. (2019). Transcriptome analysis of a rice cultivar reveals the differentially expressed genes in response to wild and mutant strains of Xanthomonas oryzae pv. oryzae. *Scientific Reports*, *9*(1), 3757.
19. Yabe, A. (1987). Current status and future prospects of research on bacterial blight of rice. *Annual review of phytopathology*.