International Journal of Ecology and Environmental Sciences 50: 693-698, 2024
ISSN: 2320-5199 (Online); https://doi.org/10.55863/ijees.2024.0151

Exploratory Development of Predictive Model to Study the Rice Blast Disease Development at Different Growth Stages Using Machine Learning

RITU RAJ, BALJEET KAUR* AND P. P. S. PANNU
Punjab Agricultural University, Ludhiana, 141001, Punjab, India
E-mail: rituraj1610@gmail.com, bchahal57@gmail.com, adr-nrphm@pau.edu
***Corresponding author**

ABSTRACT
Rice blast disease caused by *Pyricularia grisea* Sacc. has become an emerging constraint in Basmati rice cultivation in Punjab for the recent years. A detailed field investigation was conducted during Kharif 2015 and 2016 to study the impacts of different meteorological elements on blast disease development and compute predictive models to predict the disease ahead of its appearance in the field at different growth stages. Correlation analysis showed that maximum air temperature and relative humidity were the key elements to govern the disease in the field among all other meteorological elements. Maximum air temperature around 34°C and relative humidity above 60% were observed to be favorable for the disease spread in the field. Predictive models were developed for nursery stage (R^2=0.71), tillering stage (R^2=0.81), panicle stage (R^2=0.99) and for whole growth period (R^2=0.55) using R programming. A step-wise multilinear regression approach was adopted to identify the most appropriate predictive variables to formulate the model.

Key words: Machine learning, Multicollinearity, Predictive model, Blast disease, Rice

INTRODUCTION

Rice (*Oryza sativa* L.) is one of the most important cereal crop grown worldwide. It is a primary source of food for more than half of the world‘s population particularly in Asia, where more than 90 per cent of the world‘s rice is grown and consumed (Pooja and Katoch 2014). Globally, India accounts for largest area under rice cultivation with total area of 43.52 million hectares, generating an annual production of 159.20 million tonnes and an average productivity of 3.59 t/ha during the year 2013 (Anonymous 2014). India is accounted for the cultivation of different rice varieties across the regions and basmati rice occupies a very special status in rice cultivation. Punjab is famous for producing high quality basmati rice, which currently occupies 20 per cent of total rice area in the state (Anonymous 2013).

Rice crop production is constrained by several biotic and abiotic factors. In recent times rice blast disease has appeared as the most destructive disease of rice, which is caused by the ascomycetes fungus *Pyricularia grisea,* Teleomorph *Magnaporthe grisea* (Herbert) (Ou 1985) infecting rice plant at all the developmental stages, from seedling stage to grain formation, causing leaf blast, node blast, neck blast and panicle blast symptoms (Webster and Gunnel 1992). Rice blast disease has been regarded as one of the most important disease due to its widespread distribution and its potential to cause 70 to 80% yield losses (Hajano et al. 2011). Especially under favourable environmental conditions, it can be very disastrous (Scardaci et al. 1997). A detailed field investigation was conducted to study the impacts of different meteorological elements on blast disease development and compute predictive models to predict the disease ahead of its appearance in the field at different growth stages.

METHODOLOGY

Study location

Field experiments were conducted in the month of June during 2015 and 2016 in randomized block design, using standard agronomic practices and seedlings were transplanted in field after 30 days in small plots of 1×1m size at research farm, Department of Plant Pathology, PAU, Ludhiana. Seeds of basmati rice cultivar Pusa basmati 1121 were obtained from Department of Plant Breeding and Genetics, PAU, Ludhiana.

Effect of date of sowing on the incidence of rice blast disease at different stages

Nursery of cultivar Pusa Basmati 1121 was sown at weekly interval starting from 11th of June till 15th of

July for two successive years 2015 and 2016 to study the relation of weather parameters in the incidence of disease. Weekly observations for disease incidence were recorded regarding the development of blast symptoms. A week after inoculation the data was recorded and disease incidence and disease severity was recorded. Disease severity was recorded on 0-9 scale (Anonymous 2002) as per Table 1. Disease index % (DI) was calculated as

DI = (sum of all the numerical ratings/ (total number of leaves examined X maximum rating)) X 100

The meteorological factors were tested for a correlation analysis with periodic disease incidence during Kharif 2015 and 2016. These factors viz., maximum temperature (T_{max}), minimum temperature (T_{min}), relative humidity at morning (RH_1) and at evening (RH_2) and rainfall (RF) were also used in blast predictive model. The statistical parameters of meteorological factors i.e., mean, median, maximum, minimum, standard deviation and coefficient of variation were calculated to describe the characteristics of the weather over the Ludhiana district during *Kharif* 2015 and 2016 (Table 2).

Correlation analysis was carried out to identify any variable/variables associated with the DI. Multicollinearity or correlation among predictor variable is an important consideration in a multiple linear regression. To determine Multicollinearity among the variables, variance inflation factor for each variable were worked out using Rstudio libraries viz., *car* and *corrplot.* Akaike's Information Criterion (AIC) was used to compare the fitness of the statistical model using:

$$AIC = n \log (\hat{\sigma^2}) + 2K$$

Where, = residual sum of squares/n, n = sample size, K = the number of model parameters. AIC rank each model and from best to worst. Lowest the AIC value, better is the model.

A step-wise regression analysis was done to fit appropriate model with lowest AIC value using the function step.model of library *MASS* in Rstudio.

RESULTS AND DISCUSSION

Effect of different meteorological parameters on blast incidence

Disease incidence (DI) and different meteorological parameters were observed and recorded on weekly basis to investigate the possible impacts of those parameters on the incidence of disease. To express such relationship, correlation analysis between different meteorological parameters and DI was carried out at different stages for 2015 and 2016.

Table 1. Disease rating scale

Disease rating	Description	Host behaviour
0	No symptoms observed	Highly resistant
1	Small brown specks of pin head size or larger brown specks without sporulating center	Resistant
2	Small roundish to slightly elongated, necrotic gray spots, about 1-2 mm in diameter, with a distinct brown margin. Lesions are mostly found on the lower leaves	Moderately resistant
3	Lesion type same as in 2, but significant number of lesions are on the upper leaves	Moderately resistant
4	Typical susceptible type blast lesions 3 mm or longer, infecting less than 4 per cent of leaf area	Moderately susceptible
5	Typical blast lesions infecting 4-10% of the leaf area	Moderately susceptible
6	Typical blast lesions infecting 11-25% of the leaf area	Susceptible
7	Typical blast lesions infecting 26-50% of the leaf area	Susceptible
8	Typical blast lesions infecting 51-75% of the leaf area and many leaves are dead	Highly Susceptible
9	More than 75 per cent of the leaf area affected	Highly Susceptible

Table 2. Descriptive Statistics of weather parameters during 2015 and 2016

Parameter	Kharif 2015					Kharif 2016				
	Tmin	Tmax	RH1	RH2	RF	Tmin	Tmax	RH1	RH2	RF
Mean	24.7	33.7	82	56	31.7	26.3	34.1	84	59	31.4
Median	25.7	33.3	85	53	16.0	26.1	33.6	85	59	22.6
Maximum	27.9	38.7	94	83	181.0	29.2	39.1	92	74	152.4
Minimum	16.4	29.5	56	32	0.0	22.2	32	64	37	0.0
Standard deviation	3.1	2.4	9.8	12.8	44.4	1.79	1.83	7.22	9.04	40.0
Coefficient of variation	0.1	0.1	0.1	0.2	1.4	0.07	0.05	0.09	0.15	1.3

The results revealed if a parameter was positively or negatively correlated with DI.

The correlation analysis identified maximum temperature and relative humidity as the most influential factors impacting disease incidence. During the nursery stage in 2015, a strong negative correlation (-0.73) was observed between maximum temperature and disease incidence (DI). Conversely, all measures of relative humidity (morning, evening, and mean) exhibited positive correlations with DI, ranging from 0.70 to 0.72 (Table 3). These statistically significant findings suggest that higher relative humidity throughout the day creates a more favorable environment for disease intensification at the nursery stage. Similar negative correlations between maximum temperature and DI were observed in 2016, although the positive association with relative humidity was weaker. These results align with the work of Shafaullah et al. (2011), who reported negative correlations between maximum temperature and rice blast incidence across various rice lines/varieties.

The correlation analysis for the tillering stage in both years revealed weak negative associations between DI and maximum temperature, minimum temperature, and rainfall. In contrast, morning, evening, and mean relative humidity displayed weak positive correlations with DI during this stage. At the panicle stage, however, a stronger positive association emerged between DI and all humidity measures. Correlation coefficients for morning, evening, and mean relative humidity during 2015 were 0.81, 0.71, and 0.72, respectively. Similarly, in 2016, these coefficients ranged from 0.68 to 0.76. These findings are consistent with previous research by Prasad and Rana (2002), highlighting the significant role of relative humidity in influencing rice blast disease.

Development of predictive models to estimate blast incidence at various growth stages

Development of predictive models was attempted from the meteorological and disease data We encountered multicollinearity, a condition of high correlation among predictor variables, when developing predictive models using the collected meteorological and disease data from 2015 and 2016. This was evident from the high variance inflation

Table 3. Correlation coefficients between weather parameters and blast disease incidence in rice during kharif 2015 and 2016

Weather parameters	**Nursery stage**		**Tillering stage**		**Panicle stage**	
	2015	2016	2015	2016	2015	2016
T_{max}	-0.73*	-0.66**	-0.30	-0.13	-0.45	-0.55
T_{min}	0.22	-0.46	-0.39	-0.28	-0.44	-0.51
RH_1	0.72*	0.54	0.47	-0.01	0.81*	0.68**
RH_2	0.67*	0.44	0.36	0.30	0.71**	0.76*
RH_M	0.70*	0.50	0.45	0.27	0.72*	0.69**
Rainfall	0.12	0.35	-0.08	-0.23	-0.18	-0.31

factors (VIFs) observed for maximum temperature, minimum temperature, and all humidity variables (Fig. 1). Correlograms (Fig. 2) further confirmed these strong associations. Including all variables together in the models could have hampered their predictive ability. Therefore, to address this challenge and identify the most suitable variables for robust model construction, we employed a step-wise multilinear regression approach. This technique prioritizes variables based on their contribution to the model, ultimately leading to models with the lowest AIC (Akaike's Information Criterion) (Table 3). The resulting models achieved promising R-squared values ranging from 0.71 (nursery stage) to 0.99 (panicle stage), demonstrating their effectiveness in predicting disease incidence at various growth stages. Additionally, the low probability values for the coefficients of variables further support the appropriateness of the chosen models. Observed versus predicted plots (Fig. 3) visually confirm the good fit between the models' predictions and actual disease incidence data. Overall, the results highlight the significant influence of meteorological parameters, particularly temperature and relative humidity, on rice blast disease incidence. The development of predictive

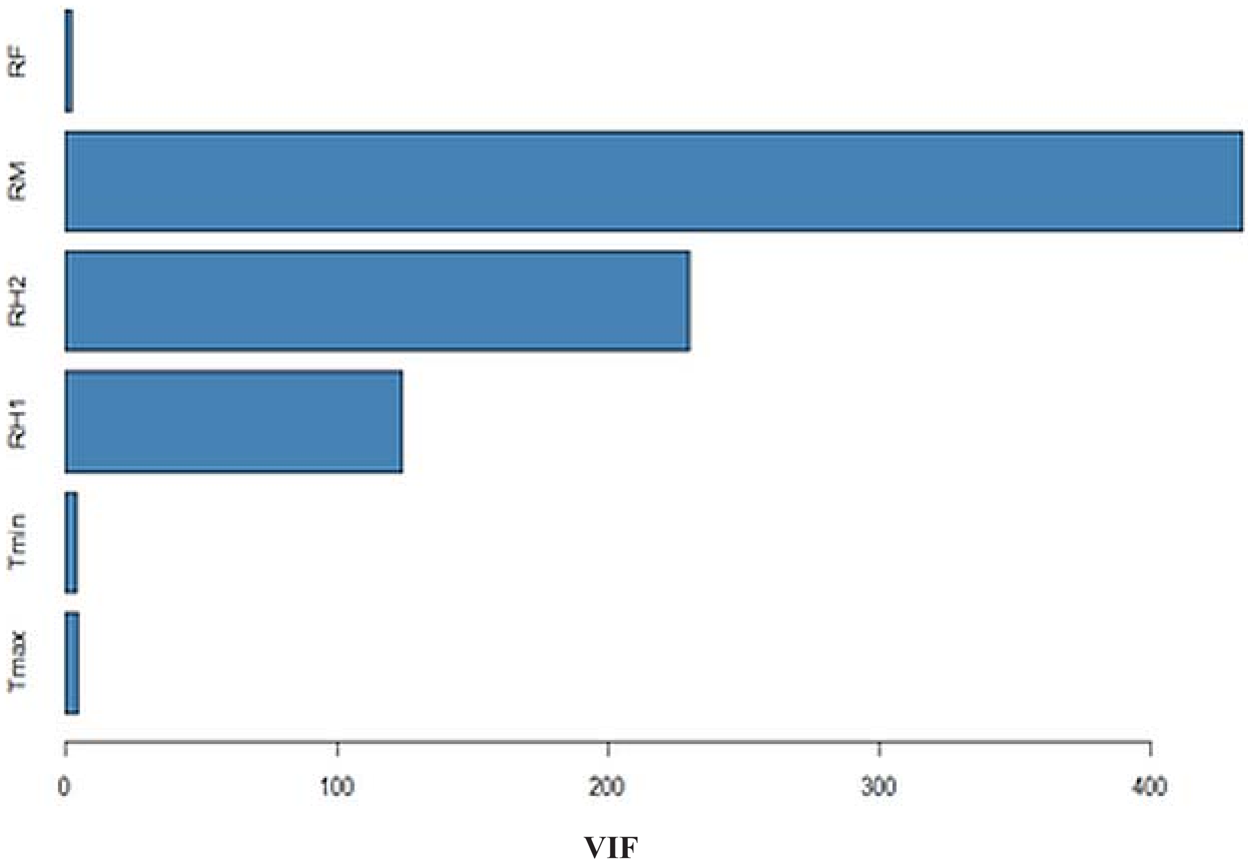

Figure 1. Variance Inflation Factors (VIF) of the predictive variables with meteorological factors

Table 3. Step-wise multi-linear regression models for the prediction of blast disease incide**nce**

Crop growth stage	**Equation**	**R^2 (Adj R^2)**	**F-stat**	**AIC**
Nursery	DI = $1108.6-22.05T_{max}-3.68_{RH1}-0.42_{RF}$	0.71(0.54)	4.16*	59.9
Tillering	DI = $1488.1+15.48T_{max}-37.66_{Tmin}-22.41RH_1+12.62RH_M$	0.81(0.69)	6.68*	68.9
Panicle	DI=$-437.8-5.86T_{max}+2.39T_{min}+8.63RH_1+19.29RH_2-16.41RH_M+1.55RF$	0.99(0.98)	71.12**	18.02
Nursery to panicle	DI=$511.14-16.41T_{min}-2.91RH_1+3.72RH_2-0.50RF$	0.55(0.47)	7.01***	184.1

*significant at 5%, ** significant at 10%, ***significant at 1%

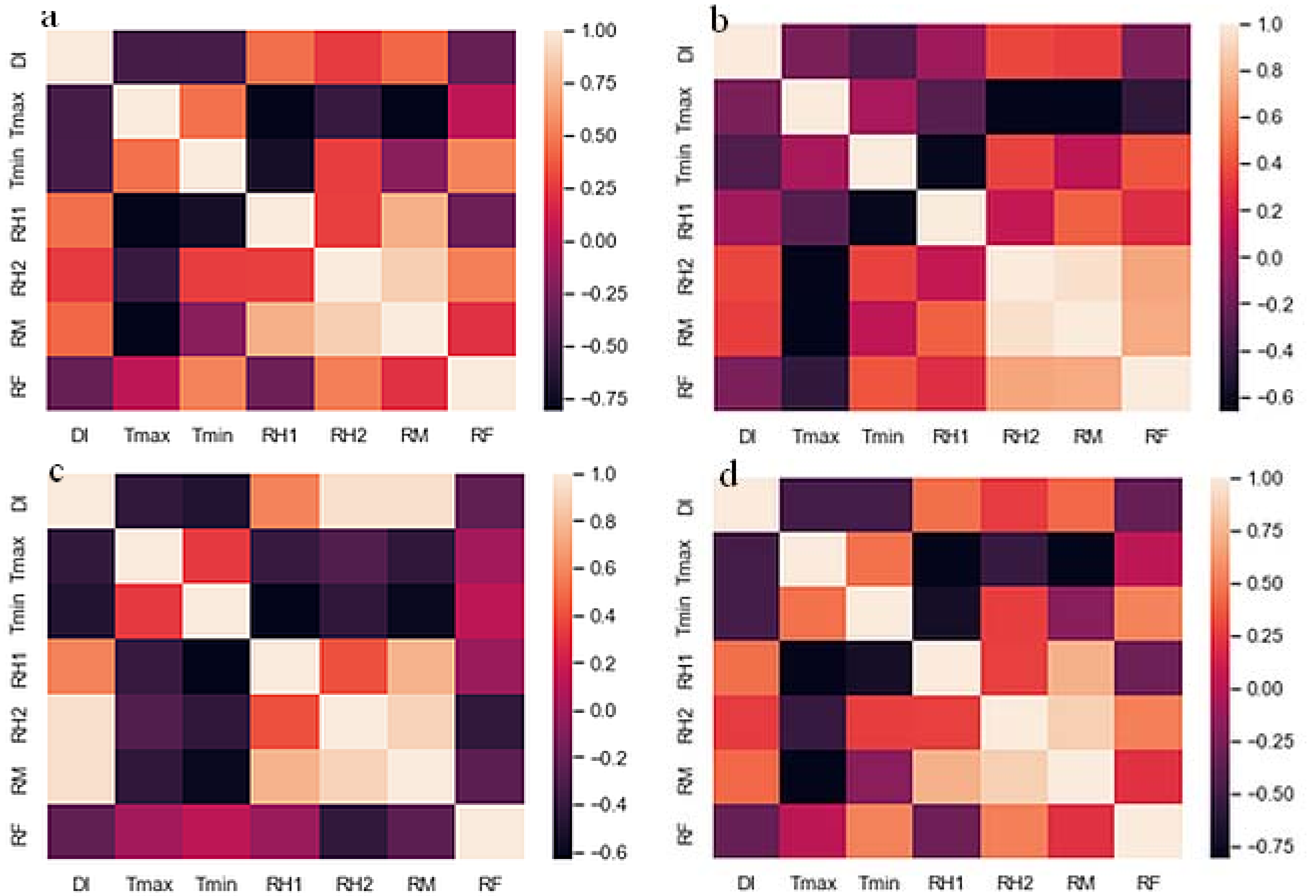

Figure 2. Correlation matrix between weather parameters and blast disease incidence a) at nursery stage, b) at tillering stage, c) at panicle stage, and d) during nursery to panicle stage

models provides valuable tools for predicting disease occurrence at different growth stages, aiding in the implementation of timely management strategies to mitigate the impact of rice blast disease on crop yield and quality (Zhang and Chen 2016).

CONCLUSION

This study investigated the influence of meteorological elements on rice blast disease development in Basmati rice and explored the potential for using predictive models to forecast outbreaks at various growth stages. The findings revealed that maximum temperature and relative humidity were the key factors impacting disease incidence. Higher temperatures were negatively correlated with disease severity, while higher relative humidity throughout the day exhibited a positive correlation. The developed predictive models achieved promising results, with R-squared values ranging from 0.71 to 0.99 for different growth stages. These models utilized a step-wise multilinear regression approach to address multicollinearity among meteorological variables, ensuring robust and reliable predictions. Overall, this study highlights the crucial role of temperature and humidity in rice blast disease development. The established predictive models offer valuable tools for Basmati rice growers to anticipate disease outbreaks at critical stages. This early warning system empowers farmers to implement timely management strategies, such as fungicide applications, to minimize yield losses and ensure Basmati rice quality.

Authors' contributions: RR planned the experiment, collected data and participated in manuscript drafting, BK participated in analysis of data and manuscript writing and PPSP is responsible for conceptualization of the study and supervising the work.

Conflict of interest: The authors declare that there is no conflict of interest related to this article.

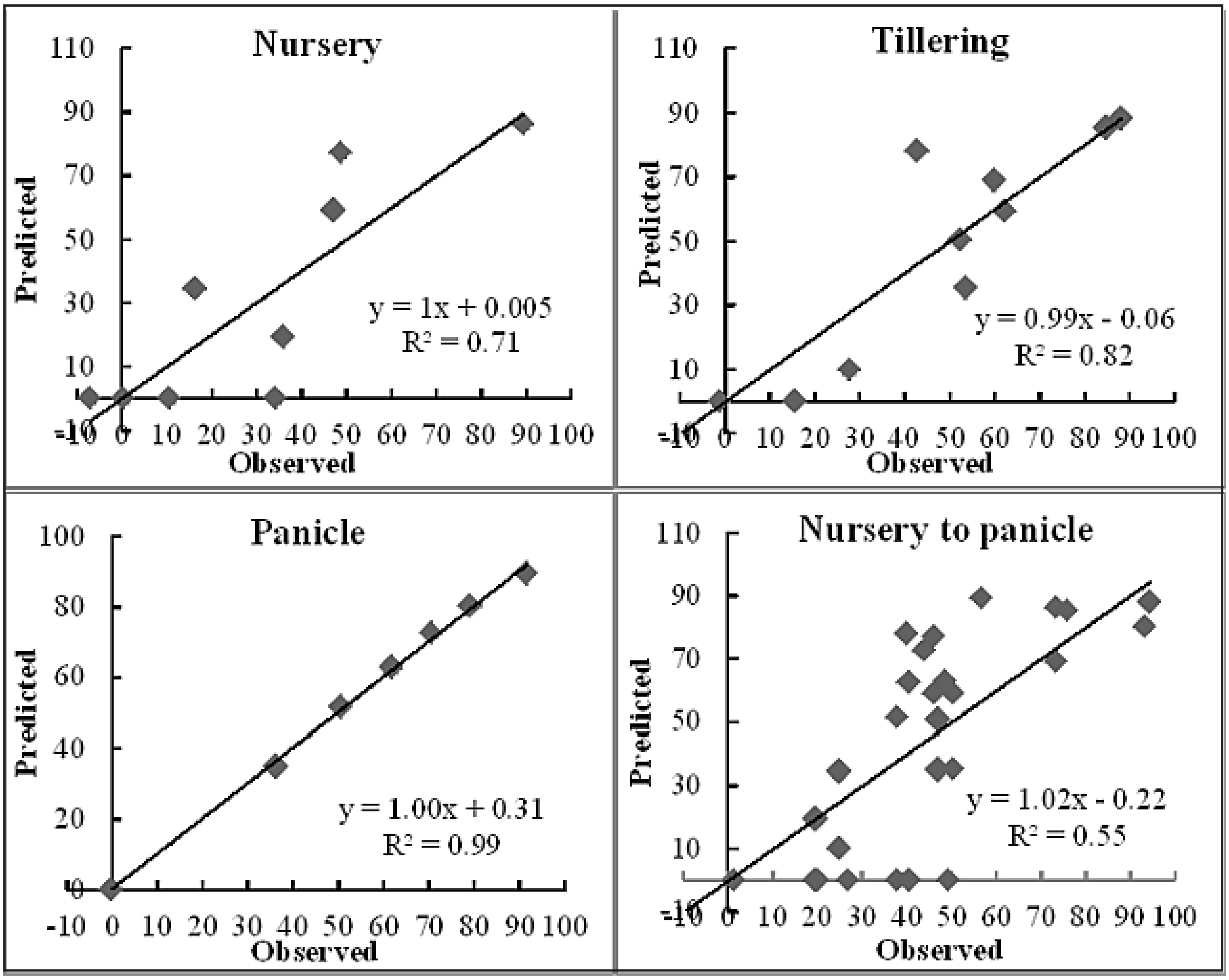

Figure 3. Performance of the predictive models for the rice blast disease incidence (DI)

REFERENCES

Anonymous. 2013. State/season-wise production of rice in India. Ministry of Agriculture, Govt. of India. *www.Indiastat.com*

Anonymous. 2014. STAT Database. Food and Agriculture Organization, Rome. www.faostat3.fao.org

Anonymous. 2002. Standard evaluation system for rice. International Rice Research Institute, Manila, Philippines. 54 pages.

Hajano, J.U., Pathan, M.A., Rajput, Q.A. and Lodhi, M.A. 2011. Rice blast mycoflora, symptomatology and pathogenicity. International Journal for Agro Veterinary and Medical Sciences, 5, 53-63.

Ou, S.H. 1985. Rice Diseases. CAB International Mycological, Institute Kew, Survey, UK.

Pooja, K. and Katoch. 2014. Past, present and future of rice blast management. Plant Science Toda*y*, 1, 165-73.

Prasad, R. and Rana, R.S. 2002. Weather relation of rice blast in mid hills of Himachal Pradesh. Journal of Agrometeorology, 4(2), 149-152.

Scardaci, S., Webster, R., Greer, C., Hill, J., Williams, J., Mutters, R., Brandon, D., McKenzie, K. and Oster, J. 1997. Rice blast: a new disease in California. Agronomy Fact Sheet Series, 1, 2-5.

Shafaullah, Muhammad, A.K., Nasir, A.K. and Yasir, M. 2011. Effect of epidemiological factors on the incidence of paddy blast (*Pyricularia oryzae*) disease. Pakistan Journal of Phytopathology, 23 (2), 108-111.

Webster, R.K. and Gunnell, P.S. 1992. Compendium of Rice Diseases. American Phytopathological Society.

Zhang, L. and Chen, S. 2016. Selection of predictor variables for regression models in agricultural research. Agricultural and Forest Meteorology, 218, 235-246.

Received:24th January 2024
Accepted: 27th May 2024