Global NEST Journal, Vol 28, No 1, 08084
Copyright© 2026 Global NEST
Printed in Greece. All rights reserved

Prediction of Tomato Leaf Curl Virus Disease and Its Vector Based on Meteorological Conditions Using Multiple Regression Analysis

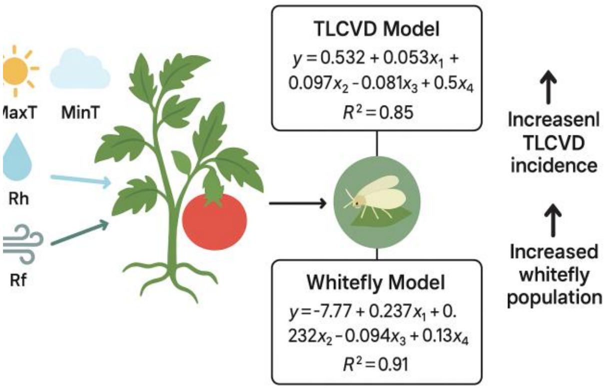
Ahmed Mahmoud Ismail[1†*], Muhammad Ahmad Zeshan[2†*], Mustansar Mubeen[2], Muhammad Aslam Khan[3], Safdar Ali[3], Yasir Iftikhar[2], Hossam S. El-Beltagi[4], Adel A. Rezk[4], Bader Alsubaie[4] and Othman Al-Dossary[4]
[1]Pests and Plant Diseases Unit, College of Agricultural and Food Sciences, King Faisal University, Al-Ahsa-31982, Saudi Arabia
[2]Department of Plant Pathology, College of Agriculture, University of Sargodha, Sargodha-40100, Pakistan
[3]Department of Plant Pathology, University of Agriculture, Faisalabad-3800, Pakistan
[4]Agricultural Biotechnology Department, College of Agricultural and Food Sciences, King Faisal University, Al-Ahsa 31982, Saudi Arabia
†These authors have contributed equally to this work

Received: 09/10/2025, Accepted: 16/01/2026, Available online: 20/01/2026

*to whom all correspondence should be addressed: e-mail: muhammad.ahmad@uos.edu.pk and amismail@kfu.edu.sa

https://doi.org/10.30955/gnj.08084

Graphical abstract



Abstract

Models developed to predict diseases associated with the tomato leaf curl virus (TLCVD) in different countries of the world. The research has been set to identify those parameters that are maximally influenced by the infectious weather variables. Therefore, the objectives of this research were to forecast the TLCVD disease together with its vector whitefly based on meteorological conditions and to validate the two models. Sown in three replications using RCBD over two seasons were the following varieties of tomato genotypes: Big Beef, Sitara-TS-101, 014276, Caldera, and Salma. Data on disease incidence and weather variables were used to develop a disease-predictive model for TLCVD by employing stepwise regression analysis. The criterion of evaluation for this model was a comparison between predicted and empirical observations. TLCVD predictive model $y= 0.532+ 0.053x_1$ (°C) $+0.97x_2$ (°C) – $0.081x_3$ (%) $+0.15x_4$ (mm) $R^2= 0.85$ y= TLCV disease, while whitefly predictive model $y= -7.77+0.237x_1$ (°C) $+0.223x_2$ (°C) $-0.094x_3$ (%) $+0.13x_4$ (°C) $+0.085x_5$ (m/s) $R^2= 0.91$. All weather parameters affected the incidence of the disease in addition to the infestation with the whitefly. A revised RMSE of 0.51 after exclusion of outliers demonstrated the accuracy in a determined range among different cultivars. Based on the current study findings, predictive models for TLCV disease and in respect of whitefly would be developed for different agroecological zones of Pakistan.

Keywords: Begomovirus, Whitefly, Empirical model, Epidemiology

1. Introduction

Tomato leaf curl virus, or TLCV, is a begomovirus from the family *Geminiviridae* that affects tomato plants particularly in conditions that are warm and comparatively less humid (Li *et al.* 2022). TLCV causes the most damaging and yield limiting tomato leaf curl disease (TLCD) worldwide (Yan *et al.* 2021). Prevailing the conducive environment, disease severity may be up to 80% causing a substantial yield loss (El-Sappah *et al.* 2022). The badly affected tomato fields result in 100% yield loss, costing tens of millions of dollars in the US alone (Kandhro 2020). More than 7 million hectares of tomato crops are affected by different strains of TLCV around the globe (Abbas *et al.* 2022). Whitefly introduces the TLCV in the phloem of the tomato plant to suck juice (Dhaliwal *et al.* 2020). The spread of viruses in the host plant depends upon the feeding times of whitefly (Legarrea *et al.* 2020). Epidemiological models that forecast the possibility of disease outbreaks may provide significant information for tomato growers enabling them to implement a timely disease management strategy (Hasanaliyeva *et al.* 2022). The practical benefit of disease predictive models is to provide advance warning about the disease. This advance warning allows for the determination of possible disease outbreaks, saving of resources, sites for cultivation, selection of cultivars, and the execution of a

Ahmed Mahmoud Ismail, Muhammad Ahmad Zeshan, Mustansar Mubeen, Muhammad Aslam Khan, Safdar Ali, Yasir Iftikhar, Hossam S. El-Beltagi, Adel A. Rezk, Bader Alsubaie and Othman Al-Dossary (2026), Prediction of Tomato Leaf Curl Virus Disease and Its Vector Based on Meteorological Conditions Using Multiple Regression Analysis, *Global NEST Journal*, **28**(1), 08084.

judicious control strategy (Saurabh *et al.* 2021). The prediction may be taken into consideration to reduce the number of insecticidal sprays under unfavorable conditions for whitefly infestation thus minimizing the monetary costs and environmental issues (Gao *et al.* 2025). The predictive models infer those insecticidal sprays would only be beneficial when the weather forecasts are in favor of whitefly infestation (Dearlove *et al.* 2024). However, when the predicted environmental conditions are not favoring the insect population and subsequent viral transmission, models would be used to reduce unnecessary insecticidal applications, thus minimizing the input expenditures (Razzaq *et al.* 2023). Numerous models are being developed to forecast the onset of the TLC disease epidemic by considering previous and current observations to predict the likelihood of epidemic onset (Ali *et al.* 2022). These models provide the basis for developing decision support systems (DSS) which predict the disease dynamics for multiple locations and their cost-effective management. The disease caused by the virus was modeled on a plant population and its dynamics were determined by a linked differential equation. The plant population was categorized into disease free, asymptomatic, and symptomatic before developing the model (Nisar *et al.* 2024). A disease predictive model was devised based on the vector population. The disease outbreak was assessed by considering the most favoring weather variables for insect infestation (Nguru and Mwongera 2023). The complications in Holt model was resolved by keeping the host and vector population constant (Asghar *et al.* 2024). TLC disease was predicted by using biological and epidemiological data because both virus and vector shared a wider host range (Nguru and Mwongera 2023). The biological data of the Holt model inferred that the vector behavior and activity, especially virus spread is a crucial aspect in the magnitude of TLCVD epidemic development (Alhakami *et al.* 2022). The spatio-temporal progress of the disease was mutually described through attributes of the pathos system and weather variables by the predictive model (Gurung *et al.* 2022; Zhao *et al.* 2025). Under the field conditions, the correlation of TLCV and insect vector was described in varied tomato germplasm. In tomato fields, the relationship of *B. tabaci* infestation virus transmission was directly proportional ($Y = -0.001x_1 + 0.03x_2 + 1.06$ and $R^2 = 0.66$) (Basha *et al.* 2024). Temperature directly showed correlation whereas relative humidity was inversely correlated with the whitefly population (Saghafipour *et al.* 2020). Considering the incidence of TLCVD in Pakistan, the development of a predictive model through regression analysis that describes the probable presence and absence of the disease may be beneficial. In Latin America, huge destruction to tomatoes was done by *B. tabaci* and TLCV for (Walerius *et al.* 2023) which a Flora Map was used to develop a prediction model for climate that destroyed tomatoes in Latin America by *B. tabaci* and TLCV.A model that can be used specifically for the prediction of whitefly and TLCD has not been developed in Pakistan. A very few predictive models have been developed for TLCD in the world and weather conditions of Pakistan are different from those specific countries so those models may not be useful for prediction in Pakistan. The objectives of this experiment were to (i) determine if regional weather conditions may be related to TLCV disease incidence and whitefly (ii) determine the disease incidence could be related to whitefly infestation (iii) validation of both models.

2. Materials and Methods

The meteorological data were acquired from the weather observatory set up at the University of Agriculture Faisalabad (UAF) Pakistan. Latitude: 31°-44' N Longitude: 73°-06' E. Altitude: 184.4m. Disease incidence and whitefly infestation were assessed from data obtained over 5 tomato genotypes cultivated on the experimental site at UAF. Complete information about weather instruments and methods utilized by Agromet is available at the website (www.uaf.edu.pk) and were complete for the entire study period, with no missing values in any meteorological parameters measured.

2.1. Study Design Statement

This study constituted an agronomic field experiment focused on plant disease modeling. The research did not involve human participants or animals, and therefore did not require formal ethical approval or registration in a clinical trials registry.

2.2. Prediction of TLCV Disease and Whitefly Infestation

The experiment for predictive models of TLC disease incidence and whitefly infestation was initiated by sowing five tomato genotypes (Big Beef, Sitara-TS-101, 014276, Caldera and Salma) at the research field of the Department of Plant Pathology, UAF for 2 years (2023 and 2024) under randomized complete block design (RCBD) with 3 replications. The cultivars were planted at 30 cm plant to plant and 70 cm row to row distance, across 15 m long row. Data on TLC disease incidence and *B.* tabaci infestation was thus recorded weekly for all genotypes from BBCH growth stage 12 (two true leaves) until BBCH 89 (fully ripe fruit). The data of whiteflies were collected early in the morning being the ideal time point for its counting. Total 180 weekly observations come from five varieties × two years × roughly eighteen sampling dates.

2.3. Analysis of Data

Those data were analyzed with least significant difference (LSD at $P<0.05$) tests and thus could serve for ANOVA and comparison of the disease incidence under different environmental conditions. SAS 9.3 statistical software was used for data analyses. The correlation analysis was conducted to determine the effects of environmental factors viz. wind speed, rainfall, relative humidity, and low and high temperature on disease incidence and vector population (Calinski *et al.* 1981). Stepwise regression analysis was applied to the environmental variables which showed a positive relationship between TLC disease and whitefly (Cao *et al.* 2024; Hu *et al.* 2025). The weather variables which had a significantly positive influence on TLCV and whitefly were illustrated on graphs and favorable ranges for disease incidence and whitefly infestation were described. All the predictions of the model and fitted values are presented on the original scale of the dependent variable (disease incidence) to simplify straight clarification. Normal distribution of residuals was Shapiro-Wilk tests for validation of ordinary regression analysis.

2.4. Analysis of residuals and correction of autocorrelation

The model was subjected to Durbin-Watson analysis to find out the independence of residuals and upon the detection of positive autocorrelation it was further refitted for AR (1) analysis.

2.5. Assessment of Disease and Vector Predictive Models

The method devised was used for the evaluation of the regression model for TLC disease and whitefly (Lobin *et al.* 2022; Shahbaz *et al.* 2023). 1) To compare disease incidence or vector infestation along with weather variables and the physical theory of the model. 2) To compare the recorded and estimated values. 3) To collect fresh data and evaluate predicted values. Formulae for calculating root mean square error (RMSE) and percentage error for evaluation of predictive methods read as follows (Wallach and Goffinet 1989):

$$\mathrm{RMSE} = \sum \mathrm{i} = 1 \text{ to } \mathrm{n} = \left[(\mathrm{Oi} - \mathrm{Pi})2/\mathrm{n}\right]0.5$$

$$\text{Error } (\%) = \text{Recorded data} - \text{Forecasted data/Recorded data} \times 100$$

The number of observations, n, while Oi are observed data points and Pi predicted data points. Root mean square error value ± 20 indicates that prediction assessment is good (Willmott 1982). Model accuracy was evaluated by RMSE and uncertainty for a single prediction was quantified using a 95% prediction interval which was calculated as:

$$y\text{^} \pm t0.025, n - p\,\sigma\text{^}\,\mathrm{sqrt}(1 + X0^{T}(\mathrm{X}^{T}\mathrm{X}) - 1X0$$

- Where y^ is the predicted TLCV or whitefly
- t0.025, n-p is the critical t-value for 95% confidence with n-p degrees of freedom,
- σ^ is the residual standard error from the model,
- X is the model matrix from the training data,
- x0 is the vector of predictor values for the new forecast

Reliability of the model was validated by robustness tests using quantile regression and ordinary least squares (OLS) (Zeng *et al.*, 2025).

2.6. Criteria for selection of variables and modelling

The models for vector and disease were developed by multiple regressions with stepwise variable selection. The significance levels for variable entry (SLENTRY) and retention (SLSTAY) were both set at α = 0.15. Weather variables i.e. maximum temperature, minimum temperature, relative humidity, rainfall, and wind speed were used in modeling.

2.7. Justification of the model

A linear regression model commonly known as linear probability model (LPM) was used to predict TLC disease incidence. Although LPM has limitations for a bounded outcome (0 and 1), it was selected for multiple reasons as stated below.

(i) The coefficients provide a clear effect of each weather variable on disease incidence that is very valued for agriculturists.
(ii) It has model parsimony and provides a linear equation to develop a simple and working prediction that can be used in field settings.
(iii) This model was empirically fit as it did not predict beyond the 0-100% range and its assumptions are enough for interpretation of a linear regression.

2.8. Management of multicollinearity

The values of variance inflation factor (VIF) were calculated for quantification of multicollinearity in regression analysis. This was necessary to find out the extent of correlation between different predictors. VIF was calculated by following formula

$$VIF = 1/1 - R^{2}$$

Due to having a unique impact on disease development all the meteorological predictors have been retained in the model despite high multicollinearity (VIF > 12). Both stages of temperature (maximum and minimum) have a varied effect on virus replication in whitely infestation. Similarly, Rh and Rf have a distinct position in affecting weather conditions while wind speed has a role in whitefly movement. We have used ridge coefficients for stabilization of multicollinearity and to maintain the ecological validity of the model (Ma & AAppolloni 2025).

3. Results

3.1. B. tabaci Prediction by Using Weather Variables

R2 measures, mean square error (MSE), and Cp value were used in the assessment of the model's predicting capability. The model with Cp=6.0 and MSE=0.35 was selected for the prediction of whitefly (*B. tabaci*). $y = -7.77+0.237x_1+0.232\,x_2-0.094\,x_3+0.13\,x_4+0.085\,x_5$

Co-efficient of determination R^2= 0.91

All environmental variables affected the whitefly infestation on tomato genotypes. The model inferred about 0.237 units change in whitefly infestation due to a single modification of maximum temperature and 0.23 units change in minimum temperature while a single unit change in relative humidity showed 0.094 units change in whitefly infestation. Whitefly infestation changed 0.13 and 0.085 units due to a single unit change in rainfall and wind velocity, respectively (**Table 3**).

3.2. Evaluation of Whitefly Predictive Model Based upon the Physical Theory

The coefficient of determination was used to find the effectiveness of the model. In the current experiment, it was 0.92 which is very good for in-vivo studies because conditions are continuously changing. The standard error of the estimate is rather lower than most other estimates (1.3) (**Table 1**). The model depicted a significant value of F-distribution at a 5% probability level (**Table 2**). The data on environmental conditions and the *B. tabaci* population contributed significantly to the predictive model. The vector predictive model proved best for forecasting with reference to physical theory.

The analysis was subjected to Durbin-Watson test which gave 0.84 value that is below 1.5 indicating positive autocorrelation hence model was further refitted autoregressive model of order 1 AR (1) analysis. In final analysis, it was inferred that a unit change in maximum temperature would result in 0.201 units change in whitefly infestation. The weather variables predicted whitefly population significantly except wind speed. All environmental variables except wind speed remained significant predictors. The value of Durbin-Watson after analysis of AR (1) was 1.92 that is very near to ideal value of 2 indicating that residuals are almost independent.

Table 1. Statistical summary of whitefly predictive model AR (1) corrected model

R^2	**0.94**	**0.93**
Adjusted R^2	0.93	0.92
MSE	0.37	0.39
SE	1.36	1.41
Durbin-Watson	0.84	1.92
Entries	180	180

**Sample size n =180*

Table 2. ANOVA for whitefly infestation during 2023 and 2024

Source	Degree of freedom	SS	MS	F	P
Regression	6	728.204	139.321	413.82	0.0001**
Error	173	59.423	0.337		
Total	179	800.642			

*** = Significant @ 5% level of probability*

Table 3. Detailed overview of predictive model

Factors	Co-efficient	SE	Type II SS	t statistics	P
Intercept	-7.758	1.294	12.691	34.09	0.0001**
Maximum temperature	0.237	0.036	14.832	41.92	0.0001**
Minimum temperature	0.232	0.022	32.474	91.78	0.0001**
Relative humidity	-0.094	0.006	87.929	252.34	0.0001**
Rainfall	0.13	0.031	4.311	12.18	0.0006**
Wind speed	0.085	0.033	2.409	7.01	0.0091**

Table 3a. Management of multicollinearity in predictive model

Factors	Ridge Coefficients	VIF	AR(1) Coefficients
Intercept	-0.048		−6.892
Maximum temperature	0.082	3.71	0.201
Minimum temperature	0.078	1.19	0.198
Relative humidity	-0.185	12.5	−0.082
Rainfall	0.028	12.5	0.108
Wind speed	0.015	12.5	0.071

Model fitted with an AR(1) for autocorrelation (Durbin-Watson = 1.92).

Table 4. B. tabaci population predictive model

Variables	Number	R^2	CP	F	P
Maximum temperature	1	0.73	414.75	504.04	0.001**
Relative humidity	2	0.84	201.07	102.11	0.001**
Minimum temperature	3	0.92	17.98	172.39	0.001**
Rainfall	4	0.92	10.81	8.87	0.003**
Wind speed	5	0.92	6.00	6.81	0.009**

*** = Significant at 5% probability level*

3.3. Evaluation of Model by the Homogeneity of Regression

Developed predictive model of *B. tabaci* population by stepwise regression analyses, having significant variables (**Table 4**). *B. tabaci* population was significantly correlated with environmental variables. The important weather conditions in a predictive model would include minimum temperature, relative humidity, and maximum temperature. The effect of rainfall (Rf) and wind speed (Ws) appeared as non-significant. The regression model accounted for variability between 73% and 94%. The *B. tabaci* predictive model was found highly reliable by with high R2, low CP, and SE values. Normal probability plot of residuals and disease vs fit, best demonstrated the full model of the two years (2023 and 2024) (**Figure 1**). The likelihood design is usually recommended for a goodness fit of a hypothesized distribution and is a common way to check the non-normality of the data (Backhaus *et al.* 2021). The conventional maximum likelihood method, while modeling a full 2-year model, showed the maximum number of data points clustered around a reference line, deviating on both lower and higher sides by a few data points, thus destroying the ordinary distribution of those points; hence causing error in the model. Residuals are calculated as differences between observed and predicted values. The following scatter plot may give an indication of how these residuals are distributed. It demonstrated that most data points were clustered pretty much evenly around the reference mark confirming very good fit of the regression model while some of the points on reference line were not distributed in such proximity causing errors in prediction.

3.4. Assessment of Predictive Model by the Comparison of Observed and Predicted Values

The predictions were considered as good if the value of percent error and RMSE is less than 20. In the current experiment, most predicted values acquired from 2 years model on 5 cultivars, gave an error (%) and RMSE below 20 (**Tables 5 and 6**). In the development of whitefly infestation on 5 tomato cultivars, relative humidity, as well as lowest temperature, was epidemiologically important environmental variables. The significant weather variables were used in stepwise regression analysis to develop a single variable model. The predicted values by these single variety models for the *B. tabaci* population were in close compliance with recorded values in 5 tomato cultivars, lines such as Big Beef, Sitara-TS-101 along with Caldera, 014276, and Salma.

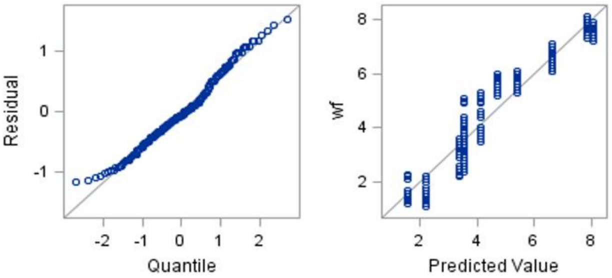


Figure 1. Model diagnostic plots. (Left) Quantile-Quantile (Q-Q) plot evaluating the normality of residuals. (Right) Plot of residuals against predicted values to evaluate homoscedasticity. The predicted (fitted) values are on the original scale of the dependent variable, representing the predicted proportion of disease incidence. (Shapiro-Wilk test: W = 0.985, p = 0.127). Sample size n =180; outliers were removed by percentage error

3.5. Assessment of Model by the Comparison of Observed and Predicted Values

Current suite of these experiments shows that for most predictions using 2-year model predictions over 5 cultivars, prediction accuracies are within ±20% with an overall mean RMSE across all predictions of less than ±20 (180) (**Table 5**).

Table 5. The variety-wise regression equation for whitefly infestation & TLCVD

Regression equations for whitefly infestation & TLCVD	**% Error**		
$y = b_o + b_1x_1 + b_2x_2 + b_3x_3$ (WF)	**Observed**	**Predicted**	
$Y = b_o + b_1 x_1 + b_2 x_2 + b_3 x_3$ (TLCV)			
Big Beef = $-0.487+0.341x_1 – 0.124x_2$ (WF)			
Big Beef = $-0.26 + 0.27X_1 – 0.085 x_2$ (TLCV)			
(x_1=MinT, x_2= Rh) (WF)	3.60	3.59	0.28
TLCV	1.20	1.18	1.67
(WF)	1.30	1.27	2.31
TLCV	1.30	1.28	1.54
(WF)	5.40	5.06	6.30
TLCV	2.10	1.83	12.86
(WF)	3.60	3.05	15.2
TLCV	3.50	3.35	4.29
Caldera = $-0.412+0.342x_1 – 0.122x_2$ (WF)			
Caldera = $-0.22 + 0.19 x_1 – 0.079 x_2$ (TLCV)			
(x_1=MinT, x_2= Rh) (WF)	1.40	1.39	0.71
TLCV	1.50	1.45	3.33
(WF)	3.90	3.83	1.79
TLCV	1.60	1.47	9.38
(WF)	5.60	5.21	6.96
TLCV	3.60	3.45	4.17
(WF)	3.80	3.26	14.21
TLCV	2.30	1.63	29.13
Sitara-TS-101 = $0.059+0.333x_1 – 0.121x_2$ (WF)			
Sitara-TS-101 = $-0.162 + 0.193x_1 – 0.74x_2$ (TLCV)			
(x_1=MinT, x_2= Rh) WF	1.50	1.46	2.67
TLCV	1.70	1.69	0.59
WF	4.50	4.16	8.17
TLCV	1.80	1.67	7.22

WF	5.90	5.43	7.97
TLCV	3.80	3.61	5.00
WF	4.20	3.58	14.76
TLCV	2.50	1.83	26.81
014276 = $-9.056+0.272x_1+0.225x_2-0.091x_3$ (WF)			
014276 = $0.18 + 0.21 x_1 - 0.089 x_2$ (TLCV)			
(x_1= MaxT, x_2= MinT, x_3= Rh) WF	6.80	6.75	0.74
TLCV	4.10	4.02	1.95
WF	5.90	5.51	7.07
TLCV	4.20	4.08	2.86
WF	1.90	1.73	8.95
TLCV	1.90	1.73	8.95
WF	4.30	3.77	12.33
TLCV	2.60	2.06	20.77
Salma = $0.722+0.308x_1-0.111x_2$ WF			
Salma = $0.26+ 0.22 x_1- 0.094x_2$ (TLCV)			
(x_1= MinT, x_2= Rh) WF	2.20	2.15	2.73
TLCV	4.30	4.21	2.09
WF	7.00	6.92	1.14
TLCV	1.90	1.76	7.37
WF	6.00	5.77	3.83
TLCV	2.20	1.95	11.36
WF	5.20	4.68	11.11
TLCV	2.80	2.18	22.14

Table 5a. 95% Prediction Intervals for Whitefly infestation Forecasts

Time Period	Predicted whitefly	Prediction interval 95% Lower	Prediction interval 95% Upper
Early season	1.4	0.7	2.1
Mid Season	3.2	2.5	3.9
Late Season	7.5	6.8	8.2
High Pressure	5.8	5.1	6.5

Predicted WF ± 0.675 (t-value × residual SE); represents range for future observations with 95% confidence

Table 5b. 95% Prediction Intervals for TLCV Disease Forecasts

Time Period	Predicted TLCD	Prediction interval 95% Lower	Prediction interval 95% Upper
Early season	0.4	0.1	0.7
Mid Season	2.3	1.8	2.8
Late Season	3.2	2.7	3.7
High Pressure	4.1	3.6	4.6

Predicted TLCV ± 0.845 (t-value × residual SE); represents range for future observations with 95% confidence

The variables with significant influence were used to develop the *B. tabaci* predictive model for two years in five tomato genotypes (**Table 6**). The high temperature, low temperature along with relative humidity played an important role in stepwise regression analysis for *B. tabaci* prediction in case of Caldera variety. The maximum temperatures, rainfalls, and wind speeds drastically affected the five different genotypes. The above-mentioned model explains more than 80 percent variability in whitefly infestation in all genotypes. The model, consisting of the three environmental parameters, was able to predict the population of *B. tabaci* with a reasonable R^2 and equally low C (p) and RMSE values obtained. Model robustness was described after the exclusion of outliers of more than 25% errors and a revised RMSE was calculated. The 95% prediction intervals quantified uncertainty for forecasts of whitefly in such a way that a mid-season prediction of 3.2 whiteflies has a 95% prediction interval of 2.5 to 3.9 for future. Similarly for TLCD, a prediction value of 2.3 represent 95% interval of 1.8 to 2.8 for future (**Table 5a & 5b**). The model predictions for disease and vector fall in biologically probable ranges (0-100% disease incidence and ≥0 whiteflies per plant, respectively), indicating that model does not produce unrealistic outputs.

Revised RMSE values calculated after excluding outliers with prediction errors >25%. Original RMSE values shown for transparency. Outliers were identified in cultivars Caldera, Sitara-TS-101, and 014276 during early to mid-season measurements.

3.6. Development of TLC Disease Predictive Model

Using 2 years of TLC disease incidence data and weather variables in stepwise regression analysis, the predictive model was developed. The predictive model $y = 0.532+ 0.053 x_1 + 0.97x_2-0.081x_3+0.15x_4$ was used to predict possible TLC disease development under weather conditions. In this equation, y = TLC disease incidence, x_1=

MaxT, x_2= MinT, x_3= Rh and x_4= Rf. According to the analysis, single unit variation in MaxT should mean a change of 0.053 units in TLCVD. In case of minimum temperature, it would be an amount of 0.97 units and, with single unit rise in Rh, there would be a decrease of 0.081 in TLCVD incidence. With a one-unit change in rainfall, the disease would experience an effect of 0.15 units. The 2 years predictive model accounted 85% variability associated with TLCVD incidence (**Table 7**). The stepwise regression analysis considered a few important factors like environmental variables, minimum temperature, rainfall, maximum temperature, and relative humidity. The model including all these variables accounted for the variability in disease development by 65 to 85%.

Table 6. Variety wise *B. tabaci* and TLCVD predictive model

Environmental parameters	R^2	Adjusted -R^2	CP	RMSE	Revised RMSE	Pr> F
Big Beef (WF)	0.94	0.93	1.537	0.64	0.49	
TLCV	0.93	0.91	0.139	0.51	0.41	
MinT						0.0195*
Rh (%)						0.0453*
Caldera (WF)	0.93	0.92	2.099	0.65	0.52	
TLCV	0.89	0.87	0.204	0.61	0.49	
MinT						0.0155*
						0.009*
Rh (%)						0.0034*
						0.00*
Sitara-TS-101 (WF)	0.91	0.89	3.049	0.71	0.53	
TLCV	0.86	0.83	0.240	0.67	0.51	
MinT						0.0152*
						0.002*
Rh (%)						0.0366*
						0.001*
014276 (WF)	0.94	0.91	3.429	0.63	0.51	
TLCV	0.92	0.91	0.083	0.55	0.44	
MaxT						0.2341
MinT						0.0045*
						0.003*
Rh (%)						0.0334*
						0.001*
Salma (WF)	0.90	0.88	2.661	0.71	0.53	
TLCV	0.93	0.92	0.096	0.53	0.42	
MinT						0.0035*
						0.001*
Rh (%)						0.0292*
						0.001*

**= Significant at 0.05 Above values are for WF; Values given below are indicating TLCV*

Table 7. TLCVD predictive model

Variables	No. in model	R^2	CP	F	P
Rh (%)	1	0.65	228.27	319.11	0.001*
MinT (°C)	2	0.82	28.54	169.65	0.001*
Rf (mm)	3	0.85	4.18	26.23	0.001*
MaxT (°C)	4	0.85	4.09	4.58	0.004*

**Significant at 5% probability*

**Sample size n =180*

The linear model described a major portion of the variance in TLC disease incidence (R^2 = 0.85). The coefficient (0.97) for MinT is indicating that average 1°C increase will enhance disease incidence by 0.97% while other variables will be constant. The model depicts the average effect on observed data range and is used for specific weather conditions of this study where predictions remain in possible limits.

The standardized coefficients (β) depicted the weightage of each predictor among which Rh was the most influential on disease incidence having β value of -0.261, showing one standard deviation increase in Rh cause a highest decrease in disease. Rh was followed by Rf, MinT and MaxT with β values 0.166, 0.097 and 0.026, respectively. The analysis indicated that Rh and Rf have more influence in disease prediction than temperature in this region.

3.7. Comparison of TLC Disease and Regression Coefficients with Physical Theory

The R^2 value for TLCVD predictive model was 0.86 (**Table. 8**). Standard error of the estimate was low (0.29) while the F value was significant at a 5% probability level. There was a vital impact which has been shown by the TLCVD predictive model with low standard error <1 value. TLC disease predictive model is considered as best based upon regression co-efficient.

Table 8. Summary of TLC disease predictive model AR (1) Corrected model

R^2	**0.86**	**0.83**
Adjusted R^2	0.85	0.83
MSE	0.28	0.32
SE	0.29	0.57
Durbin Watson	1.24	1.87
Assessment	180	

Table 9. Analysis of variance for TLC disease regression analysis

Sources	Df	Ss	ms	F	P
Model	4	289.68	72.42	244.91	0.001*
Error	175	51.75	0.29		
Total	179	341.42			

**Significant at 5% probability level*

Table 10. Overview of the predictive model

Factors	Coefficients	St. Coefficients (β)	SE	Type II SS	t stat	P
Intercept	0.53		1.18	0.07	0.22	0.001*
MaxT (°C)	0.053	0.026	0.033	1.74	4.63	0.004*
MinT (°C)	0.097	0.097	0.018	7.13	24.14	0.001*
Rh (%)	-0.081	-0.261	0.006	74.05	250.44	0.001*
Rf (mm)	0.15	0.166	0.028	8.78	29.67	0.001*

**Significant at 5% probability level*

Table 10a. Management of multicollinearity in predictive model

Factors	Ridge Coefficients	VIF	AR(1) Coefficients
Intercept	-0.048		0.841
Maximum temperature	0.082	6.66	0.042
Minimum temperature	0.078	5.56	0.083
Relative humidity	-0.185	2.86	-0.069
Rainfall	0.028	6.66	0.121

The Durbin-Watson value was 1.24, hence the model was subjected to AR (1) and value was 1.87 that indicates almost independent variables (**Tables 9 and 10**).

3.8. Variety-wise Model for TLC Disease Prediction

Five genotypes were directly affected by relative humidity (Rh) and minimum temperature (MinT) in the development of TLC disease. Single-variety models were created through stepwise regression analysis based on these conditions, and the TLCVD values predicted by those variety-wise models were found compliant with experimental values recorded on five genotypes. These models thus developed through stepwise regression helped derive statistically significant components on five tomato cultivars for incidence of TLC disease prediction over a period of 2 years. (**Table 6**). Among those critical weather variables indicated in the resultant models are relative humidity (Rh) and minimum temperature (MinT). The above varieties/lines, including Big Beef, Caldera, Sitara-TS-101, 014276, and Salma, showed a minimal effect of temperature, wind speed, and rainfall. More than 85% of the variability in TLC disease incidence among all genotypes was accounted for by this model, in which these components are included. The observed and expected points follow a nearly normal distribution, showing closeness to the reference line. Thus, the model is fit (**Figure 2**).

4. Discussion

As plant diseases force the world to face social, economic, and environmental issues, it would be beneficial to determine the progress of diseases by developing statistical models for efficient disease control (Morris *et al.* 2022). The spatio-temporal plant disease dynamics are described by predictive models formulated by integrating the host-pathogen interaction and meteorological variables (Gurung *et al.* 2022). Disease predictive models find the mathematical association of a dependent in addition to various predictor or independent variables to measure the future values of the dependent variable (Lee and Yun 2023). The major aspects of this research were to describe the influence of weather variables on TLCV disease in Faisalabad (Pakistan) by using regression analysis for two years. The overall analysis of the model showed that all five environmental variables were relevant to disease development. The model inferred that temperature and relative humidity have more influence on disease development than other variables. TLCVD predictive model is designed to forecast the possible outbreak of disease based upon weather variables in 5 tomato genotypes. Five varieties/lines have no difference in disease incidence values. In the findings, the temperature did not affect the tomato leaf curl disease epidemic development contrary to the current study (Saghafipour *et al.* 2020). Low rainfall and high temperature are good predictors of TLCV disease in the climate probability model (Walerius *et al.* 2023) and a similar trend was observed for disease development in the present study. The most common predictive models used regression and empirical approaches where disease measurement data (prevalence, incidence, and severity) is integrated with environmental variables (Shoaib *et al.* 2023). At different temperature ranges, temperature significantly influences disease resistance to viruses and affects host-pathogen interactions (Singh *et al.* 2023). Findings showed that TLCV disease incidence in plants increased with high temperature or humidity, and also showed adverse infection at 25ºC temperature and 79.73% relative humidity (Li *et al.* 2022). Positive correlations of temperature and relative humidity with tomato leaf curl incidence were also observed by scientists (Avedi *et al.* 2022). High temperature and low rainfall also have a positive correlation, they increase tomato leaf curl disease incidence (Kahsay and Makinde 2023). Wind speed was not significantly correlated with tomato leaf curl disease incidence. The following result verifies the study (Avedi *et al.* 2022) which reported a non-significant correlation between wind speed and disease incidence and that wind speed only helps in the spread of the vector whitefly. Whitefly infestation was predicted based on weather conditions that showed more than 90% variation in 2 consecutive growing seasons. The comparison of predictive models gave accurate results for forecasting whitefly and virus transmission. The major advantage of this predictive model for application in Faisalabad (Pakistan) would be to provide the farmers an awareness of the possible onslaught of Tomato leaf curl disease and enable them to the timely applicant of insecticide to minimize the whitefly population. The relationship between phloem contacts and transmission efficiency was positively significant (Legarrea *et al.* 2020). Whitefly population as well as relative humidity showed negatively substantial relationship with each other ($Y= -0.0291X_2 + 3.997X - 162.71$ and $R^2 = 0.71$) (Basha *et al.* 2024). Whitefly infestation was predicted based on environmental variables with a 65% coefficient of determination (Parola-Contreras *et al.* 2022). Disease prediction modeling techniques were developed to forecast probable outbreaks of plant viruses vectored by insects (Lee *et al.* 2022). Weather variables were used to develop a model for TLCD prediction and concluded that resistant germplasm along with new chemistry insecticides is the best disease control tactic (Gurung *et al.* 2022). After the stepwise regression analysis, it was found that the population of *B. tabaci* was significantly impacted by high temperature and low temperatures, relative humidity, wind speed, and rainfall. The findings are in accordance with Anco *et al.* (2020) who after the stepwise regression analysis determined favorable weather conditions for yellow mosaic virus and *B. tabaci* infestation. If the model predicts the probability of TLCVD development and whitefly infestation, then a management schedule may be deployed. More research is needed on the prediction of TLCVD and its vector for making precise decisions. Model verification was done by associating a real system with the model's structure it characterizes (Yağ and Altan 2022). The model should simulate the framework of the real system to pass the structure authentication process (Zhang *et al.* 2020). The verification sites were under natural infection and the variety used was the same in each site throughout the cropping season. Tomato leaf curl disease incidence (%) recorded every two weeks from transplanting from April to July 2018 served as the observed disease progress that was compared to the disease progress simulated by the Phil Curl model. It is a pre-requisite to consider the time of the year while making predictions through weather variables. Weather data in each month has a different influence on the vector, host, and virus. This model presents a reliable prediction about the major growing season of tomatoes in (Faisalabad) Pakistan. Among five weather variables, temperature and relative humidity are affecting the vector and virus more than the rest. The major tomato crop is transplanted in late march or early April in Pakistan and thus it remains safe from the harmful effects of frost and heat at the seedling stages. Rain, that contains free moisture, might well create unfavorable circumstances for whitefly activity and, as a result, for TLCVD. Whitefly infestation is more under low relative humidity and high temperatures which cause more severe TLCV disease. The vector predictive models assume that the abundance of whitefly will result in more transmission of TLCV transmission. These assumptions work well in Faisalabad conditions where weather is favorable for whitefly infestation. These findings give valuable information in deciding protective measurements against vectors and using defense-stimulating treatments against the virus.

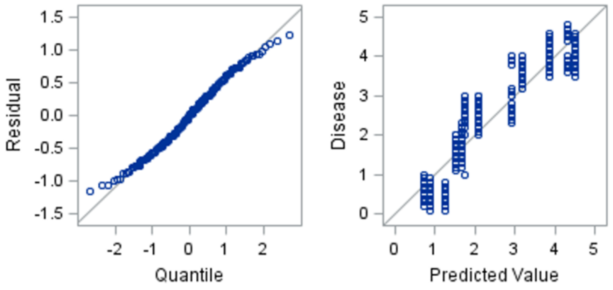


Figure 2. Model diagnostic plots. (Left) Quantile-Quantile (Q-Q) plot evaluating the normality of residuals. (Right) Plot of residuals against predicted values to evaluate homoscedasticity. The predicted (fitted) values are on the original scale of the dependent variable, representing the predicted proportion of disease incidence. (Shapiro-Wilk test: W = 0.991, p = 0.453). Sample size n =180; outliers were removed by percentage error

The decision to retain all meteorological variables despite statistical collinearity was grounded in plant pathological theory. TLCV dynamics are driven by the independent—though sometimes correlated—effects of temperature on pathogen development, humidity on vector survival, rainfall on physical disruption, and wind on dispersal. A simplified model that omitted any of these components would fail to capture the complex epidemiology of this whitefly-vectored disease.

It was noted that minimum temperature and relative humidity were key prediction components in both models of disease and vector. The issue of multicollinearity in these variables was solved by using ridge regression indicating the stability of the parameters.

4.1. Policy and climate disturbance in predictive models

Environmental sustainability and digital innovations could be linked through capacity building of human resource by training and technical education. The effectiveness of a particular policy depends upon the access of regional market and availability of resource (Lei and Zhang 2026).

According to (Lei and He 2025) abrupt changes in weather conditions have significant impact on the agricultural system by affecting the spread of pests and diseases. They stressed the integration of climate change with disease predictive models keeping in view its striking effects on biological and economic resilience as the algorithmic approaches need human input for their sustainability.

Lei and Xu (2024) stated that extreme weather anomalies like typhoons accelerate the loss of resilience in agricultural settings necessitate the need for integration of climate change with predictive models. Metropolitan climate regulator increases commercial environmental concern, display the effect of policy and human factors on climate based predictive models in agriculture (Lei and Xu 2025).

4.2. Limitations

Although the model in current study has sound prediction capacity for the specific conditions, it is vital to admit its limitations. The basic limitation is regarding data collection from a single site during two years. This means the model may capture living and non-living parameters at the selected site, while its performance may vary in years with atypical weather patterns. The generalized effectiveness of the model may be compromised due to non-validation of with different years and environmental factors. There is a precaution to apply this model for long term predictions in other regions without further validation.

The present model was developed based upon two years data from tomato growing seasons. Under the scenario of climate change the genetic resistance of the germplasm may be compromised due to mutation in pathogen that may affect the accuracy of the model with the passage of time. The model drift is a vital point to be considered for deploying the agricultural decision support.

Despite having good fitness for the observed data because of high R^2 values, the model has very limited capacity to the non-linear relationships between disease and weather variables. A linear relation is unable to describe the specific effect of temperature on vector population and disease incidence. In future, models should be focused to use non-linear modeling techniques i.e. machine learning algorithms such as random forests, support vector machines (SVM), or neural networks for improved predictions.

There was a limitation in preliminary analysis that it was not subjected to autocorrelation. However, after finding Durbin-Watson values the model was refitted to autoregressive (AR1) for validation of regression in time series analyses of whitefly and TLC disease. In future studies related to predictive modelling these analyses should be incorporated.

The values with more than 25% prediction error have been removed in revised RMSE calculation for 3 cultivars. These outliers might be appeared due to low disease incidence or during the periods of disease progression where weather variables could not be properly influencing the model.

4.3. Future Work

Consequently, later on researches should be focused on collection of data from multiple environments for external validation and to enhance the model's parameters for wider application. The process of removing outliers would be beneficial in future calculation of observed and predicted values particularly for the mentioned.

The reliability of the model can be enhanced by opting advanced AI-optimization hybrid frameworks. Application of whale optimization algorithm (WOA) with artificial neural networks (ANNs), have increased the model generalization (doi:10.1007/s11540-024-09819-9) that can also be used for TLCV and whitefly predictive models.

4.4. Specific Recommendations & Practical Considerations

For future validation studies of these predictive models, the data should be recorded from more than 30 sites in all the agro-ecological regions of Pakistan for at-least 4 years in all tomato growing seasons. The data of weather variables should be collected on daily basis while weekly data of TLCV and whitefly infestation is necessary so that multiple conditions can be added in the model. The data of multiple variables from varied locations and at different time intervals would be helpful in generalization of the models in all regions and seasons.

These predictive equations can guide the farmers and extension workers about the need for treatment interventions as the adult whitefly exceeds 3 per leaf and disease incidence recorded 10%. Exceeding the threshold beyond these limits would suggest for application of insecticides following weekly monitoring.

The predictions based upon weather variables give a lead time of 1-2 weeks before symptom appearance depending upon the availability of favourable conditions for the whitefly infestation and subsequent TLCV transmission. The lead time window is very critical for the growers to implement management tactics before waiting for the severe disease outbreak.

Authors Contribution

Ahmed Mahmoud Ismail: Formal analysis, Writing-review & editing**. Muhammad Ahmad Zeshan**: Methodology, Investigation, Writing-original draft**. Mustansar Mubeen:** Data curation, Software, Writing-original draft, Writing-review & editing**. Muhammad Aslam Khan and Safdar Ali:** Conceptualization, Supervision, Writing-review & editing. **Yasir Iftikhar:** Validation, Writing-review & editing: **Hossam S. El-Beltagi and Adel A. Rezk:** Resources, Writing-review & editing**. Bader Alsubaie:** Project administration,

Writing-review & editing**. Othman Al-Dossary:** Funding acquisition, Writing-review & editing. All authors read and approved the final version of the submitted manuscript. All authors listed have made significant, direct, and intellectual contributions to the work and approved it for submission.

Funding

Authors extend their thankfulness to the Deanship of Scientific Research, Vice-Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, for supporting work via grant number KFU242973.

Conflict of interest

The research has been conducted without any involvement by commercial or financial ties that could be interpreted as a potential conflict of interest.

Data availability

The datasets generated during and/or analyzed during the current study are available in the Zenodo repository: https://doi.org/10.5281/zenodo.17519667.

References

Abbas M, Li Y, Elbaiomy RG, et al (2022) Genome-Wide Analysis and Expression Profiling of SlHsp70 Gene Family in Solanum lycopersicum Revealed Higher Expression of SlHsp70-11 in Roots under Cd2+ Stress. Front Biosci - Landmark 27:186. https://doi.org/10.31083/j.fbl2706186

Alhakami H, Umar M, Sulaiman M, et al (2022) A Numerical Study of the Dynamics of Vector-Born Viral Plant Disorders Using a Hybrid Artificial Neural Network Approach. Entropy 24:1511. https://doi.org/10.3390/e24111511

Ali Y, Raza A, Aatif HM, et al (2022) Regression Modeling Strategies to Predict and Manage Potato Leaf Roll Virus Disease Incidence and Its Vector. Agric 12:550. https://doi.org/10.3390/agriculture12040550

Anco DJ, Rouse L, Lucas L, et al (2020) Spatial and temporal physiognomies of whitefly and tomato yellow leaf curl virus epidemics in southwestern Florida tomato fields. Phytopathology 110:130–145. https://doi.org/10.1094/PHYTO-05-19-0183-FI

Asghar A, Ikram R, Khan A, et al (2024) Applications of artificial neural network to solve the nonlinear Cassava mosaic disease model. Eur Phys J Plus 139:1012. https://doi.org/10.1140/epjp/s13360-024-05829-1

Avedi EK, Adediji AO, Kilalo DC, et al (2022) Incidence, severity and distribution of yellow leaf curl disease of tomato in Kenya. African Crop Sci J 30:1–11. https://doi.org/10.4314/acsj.v30i1.1

Backhaus K, Erichson B, Gensler S, et al (2021) Multivariate Analysis. Springer Fachmedien Wiesbaden

Basha SJ, Veeraiah D, Dev BS, et al (2024) Tomato Guard: A Predictive Model for Disease Prediction and Treatment Recommendation in Tomato Plants. In: Proceedings of the 3rd International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2024. IEEE, pp 1102–1107

Calinski T, Steel RGD, Torrie JH (1981) Principles and Procedures of Statistics: A Biometrical Approach. McGrew-Hill publishing

Cao X, Huang M, Wang S, et al (2024) Tomato yellow leaf curl virus: Characteristics, influence, and regulation mechanism. Plant Physiol Biochem 213:108812. https://doi.org/10.1016/j.plaphy.2024.108812

Dearlove EL, Chandler D, Edgington S, et al (2024) Improved control of Trialeurodes vaporariorum using mixture combinations of entomopathogenic fungi and the chemical insecticide spiromesifen. Sci Rep 14:15259. https://doi.org/10.1038/s41598-024-66051-8

Dhaliwal MS, Jindal SK, Sharma A, Prasanna HC (2020) Tomato yellow leaf curl virus disease of tomato and its management through resistance breeding: a review. J Hortic Sci Biotechnol 95:425–444. https://doi.org/10.1080/14620316.2019.1691060

Gao Y, Alyokhin A, Prager SM, et al (2025) Complexities in the Implementation and Maintenance of Integrated Pest Management in Potato. Annu Rev Entomol 70:45–63. https://doi.org/10.1146/annurev-ento-120523-023156

Gurung B, Dutta S, Singh K, et al (2022) Development of weather-based forewarning model for tomato leaf curl infestation. J Agrometeorol 24:424–426. https://doi.org/10.54386/jam.v24i3.1818

H. El-Sappah A, Qi S, A. Soaud S, et al (2022) Natural resistance of tomato plants to Tomato yellow leaf curl virus. Front Plant Sci 13:1081549. https://doi.org/10.3389/fpls.2022.1081549

Hasanaliyeva G, Si Ammour M, Yaseen T, et al (2022) Innovations in Disease Detection and Forecasting: A Digital Roadmap for Sustainable Management of Fruit and Foliar Disease. Agronomy 12:1707. https://doi.org/10.3390/agronomy12071707

Hu J, Wang D, Zeng H, et al (2025) From disclosure to decision: how climate risk shapes firms' human capital strategies in China. Theor Appl Climatol 156:556. https://doi.org/10.1007/s00704-025-05767-7

Kahsay BN, Makinde OD (2023) Ecoepidemiological Model and Optimal Control Analysis of Tomato Yellow Leaf Curl Virus Disease in Tomato Plant. J Appl Math 2023:4066236. https://doi.org/10.1155/2023/4066236

Kandhro AG (2020) Spatial distribution of chili leaf curl disease and its correlation with environmental factors in Southern Irrigated Plain Zone of Pakistan. Pure Appl Biol 9:2543–2553. https://doi.org/10.19045/bspab.2020.90270

Lee H, Stephanus AP, Fowles TM, et al (2022) Insect vector manipulation by a plant virus and simulation modeling of its potential impact on crop infection. Sci Rep 12:8429. https://doi.org/10.1038/s41598-022-12618-2

Lee S, Yun CM (2023) A deep learning model for predicting risks of crop pests and diseases from sequential environmental data. Plant Methods 19:145. https://doi.org/10.1186/s13007-023-01122-x

Lei X, He S (2025). Climate shocks and innovation persistence: evidence from extreme precipitation. Humanit Soc Sci Commun 20; 12(1):1-4

Lei X, Xu X (2024). Storm clouds over innovation: Typhoon shocks and corporate R&D activities. Econ Lett 1; 244:112014.

Lei X, Xu X. (2025). From City Wisdom to Corporate Action: Climate-Adaptive City Construction and Corporate Environmental Responsibility. Weather Climate Soc 17(4):1027-41.

Lei X, Zhang Z (2026). The algorithm can't see everything: Human capital as the hidden variable in environmental technology success. J Environ Manag 397, 128287.

Legarrea S, Barman A, Diffie S, Srinivasan R (2020) Virus accumulation and whitefly performance modulate the role of alternate host species as inoculum sources of tomato yellow leaf curl virus. Plant Dis 104:2958–2966. https://doi.org/10.1094/PDIS-09-19-1853-RE

Li F, Qiao R, Yang X, et al (2022) Occurrence, distribution, and management of tomato yellow leaf curl virus in China. Phytopathol Res 4:28. https://doi.org/10.1186/s42483-022-00133-1

Lobin KK, Jaunky VC, Taleb-Hossenkhan N (2022) A meta-analysis of climatic conditions and whitefly Bemisia tabaci population: implications for tomato yellow leaf curl disease. J Basic Appl Zool 83:57. https://doi.org/10.1186/s41936-022-00320-8

Ma S, Appolloni A (2025) Can financial flexibility enhance corporate green innovation performance? Evidence from an ESG approach in China. Journal of Environmental Management 387:125869. https://doi.org/10.1016/j.jenvman.2025.125869

Morris CE, Géniaux G, Nédellec C, et al (2022) One Health concepts and challenges for surveillance, forecasting, and mitigation of plant disease beyond the traditional scope of crop production. Plant Pathol 71:86–97. https://doi.org/10.1111/ppa.13446

Nguru W, Mwongera C (2023) Predicting the future climate-related prevalence and distribution of crop pests and diseases affecting major food crops in Zambia. PLOS Clim 2:e0000064. https://doi.org/10.1371/journal.pclm.0000064

Nisar KS, Farman M, Abdel-Aty M, Ravichandran C (2024) A review of fractional-order models for plant epidemiology. Prog Fract Differ Appl 10:489–521. https://doi.org/10.18576/pfda/100313

Parola-Contreras I, Guevara-González RG, Contreras-Medina LM, et al (2022) Analysis of the Population Dynamics of Whitefly (Bemisia tabaci [Hemiptera: Aleyrodidae]) under Greenhouse Conditions. Environ Entomol 51:294–302. https://doi.org/10.1093/ee/nvab128

Razzaq A, Zafar MM, Ali A, et al (2023) Biotechnology and Solutions: Insect-Pest-Resistance Management for Improvement and Development of Bt Cotton (Gossypium hirsutum L.). Plants 12:4071. https://doi.org/10.3390/plants12234071

Saghafipour A, Zahraei-Ramazani A, Vatandoost H, et al (2020) Relationship between some environmental and climatic factors on outbreak of whiteflies, the human annoying insects. J Arthropod Borne Dis 14:78–87. https://doi.org/10.18502/jad.v14i1.2714

Saurabh S, Mishra M, Rai P, et al (2021) Tiny flies: A mighty pest that threatens agricultural productivity—a case for next-generation control strategies of whiteflies. Insects 12:585. https://doi.org/10.3390/insects12070585

Shahbaz MU, ul Haq ME, Kamran M, et al (2023) Prediction of cotton leaf curl virus disease at different potassium dosses based on abiotic environmental factors and sowing dates. Arch Phytopathol Plant Prot 56:1205–1220. https://doi.org/10.1080/03235408.2023.2276099

Shoaib M, Shah B, EI-Sappagh S, et al (2023) Corrigendum: An advanced deep learning models-based plant disease detection: a review of recent research (Frontiers in Plant Science, (2023), 14, (1158933), 10.3389/fpls.2023.1158933). Front Plant Sci 14:1158933. https://doi.org/10.3389/fpls.2023.1282443

Singh BK, Delgado-Baquerizo M, Egidi E, et al (2023) Climate change impacts on plant pathogens, food security and paths forward. Nat Rev Microbiol 21:640–656. https://doi.org/10.1038/s41579-023-00900-7

Walerius AH, Pallini A, Venzon M, et al (2023) Use of Geostatistics as a Tool to Study Spatial-Temporal Dynamics of Leucoptera coffeella in Coffee Crops. Agric 13:438. https://doi.org/10.3390/agriculture13020438

Wallach D, Goffinet B (1989) Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecol Modell 44:299–306. https://doi.org/10.1016/0304-3800(89)90035-5

Willmott CJ (1982) Some comments on the evaluation of model performance. Bull - Am Meteorol Soc 63:1309–1313. https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2

Yağ İ, Altan A (2022) Artificial Intelligence-Based Robust Hybrid Algorithm Design and Implementation for Real-Time Detection of Plant Diseases in Agricultural Environments. Biology (Basel) 11:1732. https://doi.org/10.3390/biology11121732

Yan Z, Wolters AMA, Navas-castillo J, Bai Y (2021) The global dimension of tomato yellow leaf curl disease: Current status and breeding perspectives. Microorganisms 9:740. https://doi.org/10.3390/microorganisms9040740

Zhang N, Yang G, Pan Y, et al (2020) A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades. Remote Sens 12:1–34. https://doi.org/10.3390/rs12193188

Zhao G, Zhai X, Wu Y, Ma S (2025). Spatial and Temporal Evolution Patterns and Influencing Factors of Logistics Expansion in the Beijing-Tianjin-Hebei-Shanxi Region. Promet Traffic Transportation, 37(6), 1525–1544. https://doi.org/10.7307/ptt.v37i6.899

Zeng H, Liu H, Yan H, Ma S (2025) Biodiversity risk and global stock markets: A cross-national heterogeneity analysis based on quantile-on-quantile methods. Borsa Istanbul Review. https://doi.org/10.1016/j.bir.2025.10.013