

Research Article

Predictive analysis of cotton leaf curl disease using growth models in Haryana's Hisar Region

Manoj Kumar*

Department of Mathematics and Statistics, Department of Agrometeorology, CCS Haryana Agricultural University, Hisar-125001 (Haryana), India

Anil Kumar Saini

Department of Plant Pathology, CCS Haryana Agricultural University, Hisar-125001 (Haryana), India

Sunil Kumar

Department of Genetic and Plant Breeding, CCS Haryana Agricultural University, Hisar-125001 (Haryana), India

Anurag

Department of Agrometeorology, CCS Haryana Agricultural University, Hisar-125001 (Haryana), India

Sarita Rani

Department of Mathematics and Statistics, CCS Haryana Agricultural University, Hisar-125001 (Haryana), India

*Corresponding author: E-mail: m25424553@gmail.com

Article Info<https://doi.org/10.31018/jans.v17i3.6711>

Received: April 01, 2025

Revised: August 13, 2025

Accepted: August 25, 2025

How to CiteKumar, M. *et al.* (2025). Predictive analysis of cotton leaf curl disease using growth models in Haryana's Hisar Region. *Journal of Applied and Natural Science*, 17(3), 1253 - 1261. <https://doi.org/10.31018/jans.v17i3.6711>**Abstract**

Cotton leaf curl virus disease (CLCuV) incidence has increased in the northern states of India (Punjab, Haryana, and Rajasthan) since 1993. CLCuV reduces the number of harvestable bolls by 15-87 % and seed cotton yield by 11-92%, depending on the extent and timing of infection. Cotton leaf curl virus disease (CLCuV) is among the most damaging cotton diseases. It was first found on *Gossypium barbadense* in Nigeria (Africa) in 1912. CLCuV was first reported in India in Sri Ganganagar (Rajasthan) in 1993 and in Punjab (India) in 1994. Various mathematical models have been employed to create models which describe epidemic dynamics. In this study, an attempt was made to compare three models—namely, the Monomolecular, Logistic, and Gompertz models—using secondary data from 2017 to 2022 for the prediction of Cotton Leaf Curl Disease (CLCuD). The studied models were evaluated using goodness-of-fit criteria, i.e., the Coefficient of determination and root mean square error. It was observed that the Coefficient of Determination (R^2) was found to be in the range of 0.98 to 0.99 for all years, whereas the root mean square value was observed in the range of 2.06 to 16.25. The value of the Coefficient of determination (R^2) was very high, and the mean square error was very low for the years 2017, 2019, 2020, and 2022, respectively, for the logistic model compared to other studied models. It was observed that 98 to 99 % the variation in disease intensity was explained by time, and an error was also observed to be minimal for all years. It was also concluded that, for the years 2017 to 2022, except for 2018 and 2021, the logistic model was found to be the best fit for predicting disease severity. In contrast, for 2018 and 2021, the Gompertz model was found to be the best. The analysis was performed using R and Excel software.

Keywords: Cotton leaf curl viral disease (CLCuV), Growth model, Coefficient of determination, root mean square error, prediction**INTRODUCTION**

Cotton is one of the most important cash crops globally and is widely referred to as "White Gold" due to its significant economic value. In addition to its primary use as a source of fibre and lint for the textile industry, cotton also provides seed oil and serves as a food source

in various forms. However, the cultivation of cotton has been severely affected by Cotton Leaf Curl Disease (CLCuD), particularly in the Indian subcontinent (Bird Maramorosch, 1978; Berger, 1981; Campbell and Mad-den, 1990; Kapur *et al.*, 1994; Van Maanen and Xu 2003; Xu, 2006; Jamadar, *et al.*, 2009; Kumar *et al.*, 2019; Sain *et al.*, 2020; Monga and Sain, 2021)

CLCuD is a devastating viral disease of cotton caused by the Cotton Leaf Curl Virus (CLCuV) and is transmitted by whiteflies (*Bemisia tabaci*) in a persistent and circulative manner. The disease was first reported in Nigeria in 1912 and has since become a significant constraint to cotton production in South Asia.

Symptoms of the disease include upward or downward curling of leaves, yellowing, stunted plant growth, vein thickening, and leaf enation. The incidence and severity of CLCuD are significantly influenced by climatic factors, including temperature, rainfall, and relative humidity. The disease is known to thrive under high temperature conditions (30–40°C), high humidity levels (above 70%), and moderate to heavy rainfall. These conditions, especially prevalent during the monsoon season, create a conducive environment for the spread and intensification of the disease (Buttar and Singh, 2017; Chugh *et al.*, 2020; Kumar *et al.*, 2019; Sain *et al.*, 2024).

In the past, several statistical and epidemiological models have been developed to predict the incidence of Cotton Leaf Curl Disease (CLCuD) and its whitefly vector, based on climatic and ecological variables (Fantaye, A.K., 2022; Sattar *et al.*, 2017). Early prediction of CLCuD is crucial for understanding disease dynamics, facilitating timely intervention, and ensuring sustainable cotton production, which is vital to meet the demands of a rapidly growing population.

In this context, the present study aimed to analyze the progression behavior of CLCuD using nonlinear growth models. The objective was to develop an early prediction framework that supports effective disease management strategies and helps minimize crop losses due to this economically significant disease.

MATERIALS AND METHODS

A Cotton Leaf Curl Disease (CLCuD)-susceptible cultivar, HS-6, was sown at a spacing of 67.5 × 30 cm in the cotton research area of CCS HAU, Hisar. Observations on disease incidence, percent disease index (PDI), and whitefly population were recorded at 7-day intervals. CLCuD first appeared during the 22nd to 23rd Standard Meteorological Week (SMW) and concluded between the 40th and 44th SMW in different years. For statistical analysis, data from the 24th to the 40th Standard Meteorological Week (SMW) were considered, as presented in Table 1, following the methodology based on the foundational principles outlined by Vander Plank (1963). Disease progression was conceptualized as a dynamic interaction among host, pathogen, and environment, considering both monocyclic and polycyclic epidemics. Disease intensity was quantified using incidence, severity, and the area under the disease progress curve (AUDPC). To describe the temporal development of the disease, three nonlinear models—monomolecular, logistic, and ompertz—were fitted

to the data. Comparative analyses were conducted to evaluate the performance of each model using R² and RMSE values. Overall, this approach follows Vander Plank's emphasis on integrating quantitative disease modeling with control strategies and aligns with modern trends in statistical and machine learning-based disease forecasting. A higher incidence and disease index were recorded during periods of elevated temperatures (approximately 33–39°C), accompanied by high relative humidity (83–94%) and consistent rainfall, particularly between the 28th and 36th SMWs. Data analysis was conducted using Microsoft Excel and R software (<https://www.R-project.org>). Since various environmental and biological factors influence plant disease epidemics, mathematical modelling was employed to understand the progression of the disease. These models help simulate disease behaviour in response to external variables, such as the presence of inoculums, host susceptibility, and weather events, and assess the effectiveness of cultural and control measures.

For modelling disease progression, nonlinear mechanistic growth models were employed due to their ability to provide insights into the biological and epidemiological processes governing disease spread. Unlike empirical models, mechanistic models provide a framework for understanding the underlying dynamics of epidemics and guiding effective disease management strategies.

The growth models considered in this study included:

Monomolecular Model

This model is suitable for monocyclic epidemics, where no secondary spread of disease occurs during a growing season (Forrest, 2007). It assumes that the rate of disease increase is proportional to the remaining susceptible host tissue. The model is described by:

$$Y(t) = c - (c - b) \cdot e^{(-a \cdot t)} + e(t) \quad (1)$$

Where:

Y(t): Disease intensity at time t,

c: Carrying capacity or maximum disease level,

b: Initial level of disease,

a: Rate of disease increase,

e(t): Random error term

Logistic Model

The logistic model is widely used for polycyclic diseases (e.g., CLCuD), where secondary infections occur within a single growing season (Forrest, 2007). It is particularly suitable for modeling epidemic development in such cases.

The integrated form of the logistic model is given by:

$$Y(t) = c_1 + b \cdot e^{(-a \cdot t)} \quad (2)$$

This model produces an S-shaped (sigmoid) curve that is symmetric about the point of inflection, indicating

balanced acceleration and deceleration in disease progression.

Gompertz Model

The Gompertz model is another nonlinear model suitable for polycyclic diseases and is often used as an alternative to the logistic model (Forrest, 2007). While it also produces an S-shaped curve, the Gompertz curve is asymmetric, with a faster initial increase in disease and a gradual decline thereafter.

The model is particularly useful in biological applications due to its flexibility in capturing asymmetry in the development of epidemics. Its absolute rate curve reaches a maximum earlier than the logistic model and then declines more slowly.

Before fitting the model, the normality assumption of the data was tested using the Shapiro-Wilk test and quantile-quantile (Q-Q) plots. Additionally, descriptive statistics were calculated to analyze the behaviour of the mean and variance over time.

These models were employed to fit disease progression data for the years 2017 to 2022, using parameter optimization via the Excel Solver, and were evaluated using goodness-of-fit measures such as the Coefficient of determination (R^2) and root mean square error (RMSE).

RESULTS AND DISCUSSION

The results of this study are divided into three sections, viz. Normality testing and descriptive analysis of data, correlation analysis with weather variables, Model Fitting and Evaluation using Goodness-of-Fit criteria.

Normality testing and descriptive analysis of data

Based on the results of the Shapiro-Wilk test, there was sufficient evidence to reject the null hypothesis (i.e., that the data follow a normal distribution), since the p-value (< 0.0005) was less than the alpha level (0.05) for all years. The Q-Q plots for the years 2017 to 2022 (Fig. 1) further confirmed that the data did not satisfy the normality assumption.

The descriptive statistics presented in Table 2 indicate that, for all years, the variance exceeded the mean, which also suggests non-normality. Furthermore, when the disease intensity of cotton leaf curl disease was plotted against time (Fig. 2), the data followed an S-curve, indicating a growth pattern.

As a result, it was concluded that the data are not normally distributed, making simple linear regression techniques unsuitable for analyzing the relationship between disease intensity and time. Therefore, more appropriate nonlinear growth models were considered. The models fitted to the disease intensity data (for the years 2017-2022) included the Monomolecular, Logistic, Gompertz, and Negative Exponential growth models.

Correlation analysis with weather variables

Table 3 shows that for the years 2017 to 2022, disease severity had a positive and highly significant correlation with relative humidity, with correlation coefficients of 0.59, 0.73, 0.74, 0.77, 0.80, and 0.70, respectively. Conversely, a negative and highly significant correlation was observed with maximum temperature, with coefficients of -0.66, -0.61, -0.75, -0.79, and -0.62, except for the year 2017, which showed a correlation of -0.45.

Additionally, disease severity showed a negative, significant correlation with wind speed for all years except 2020 and 2021. These relationships are visually represented in the correlation plot (Fig. 3), which was created using R code, which is given below:

R Code for Generating Multiple Correlation Plots (Year-wise)

Load the dataset

sevdata <- read.csv(file.choose(), header = TRUE)

View(sevdata)

Set plotting layout

par(mfrow = c(3, 2)) # To plot six correlation matrices in a 3x2 grid

2017 Data (Rows 1 to 17)

d1 <- sevdata[1:17,]

M1 <- cor(d1)

p.mat <- cor_pmat(d1)

corrplot(M1,

type = "upper",

method = "square",

addCoef.col = "black",

tl.col = "black",

tl.srt = 45,

p.mat = p.mat)

2018 Data (Rows 18 to 34)

d2 <- sevdata[18:34,]

M2 <- cor(d2)

p.mat <- cor_pmat(d2)

corrplot(M2,

type = "upper",

method = "square",

addCoef.col = "black",

tl.col = "black",

tl.srt = 45,

p.mat = p.mat)

2019 Data (Rows 35 to 51)

d3 <- sevdata[35:51,]

M3 <- cor(d3)

p.mat <- cor_pmat(d3)

corrplot(M3,

type = "upper",

method = "square",

addCoef.col = "black",

tl.col = "black",

tl.srt = 45,

p.mat = p.mat)

2020 Data (Rows 52 to 68)

```

d4 <- sevddata[52:68, ]
M4 <- cor(d4)
p.mat <- cor_pmat(d4)
corrplot(M4,
  type = "upper",
  method = "square",
  addCoef.col = "black",
  tl.col = "black",
  tl.srt = 45,
  p.mat = p.mat)
# 2021 Data (Rows 69 to 85)
d5 <- sevddata[69:85, ]
M5 <- cor(d5)
p.mat <- cor_pmat(d5)
corrplot(M5,
  type = "upper",
  method = "square",
  addCoef.col = "black",
  tl.col = "black",
  tl.srt = 45,
  p.mat = p.mat)
# 2022 Data (Rows 86 to 102)
d6 <- sevddata[86:102, ]
M6 <- cor(d6)
p.mat <- cor_pmat(d6)
corrplot(M6,
  type = "upper",
  method = "square",
  addCoef.col = "black",
  tl.col = "black",
  tl.srt = 45,
  p.mat = p.mat)

```

Model fitting and evaluation using Goodness-of-Fit Criteria

The parameter coefficients for the various fitted models were computed using data from 2017 to 2022, as presented in Table 4. These parameter estimates were obtained using the Solver tool in Microsoft Excel. To evaluate the performance of each model, Two goodness-of-fit metrics were used: the Coefficient of determination (R^2) and the root mean square error (RMSE). The R^2 values ranged from 0.98 to 0.99 across all years, indicating an excellent fit of the models to the data. The RMSE values were observed to range from 2.06 to 16.25, further supporting the reliability of the models.

Among the models analyzed, the Logistic model consistently outperformed others in the years 2017, 2019, 2020, and 2022, as it exhibited higher R^2 values and lower RMSE compared to other models. This suggests that 98% to 99% of the variation in disease intensity could be explained by time, with minimal error during these years. However, for the years 2018 and 2021, the Gompertz model provided a better fit than the logistic model, as indicated by superior goodness-of-fit statistics.

The fitted values of the Logistic and Gompertz models are visually presented in Fig. 4, and their numerical values are provided in Table 5 for the respective years. The findings of the present study on modeling Cotton Leaf Curl Disease (CLCuD) using nonlinear growth models like Logistic and Gompertz are well supported by similar research endeavors in recent years. Notably, Kumar *et al.* (2021) modelled powdery mildew in cotton and reported that the disease progression over time

Table 1. Observed Cotton leaf curl disease (CLCuD (%)) with standard meteorological week

S.No.	Std Met. Week	2017	2018	2019	2020	2021	2022
1	24	22.30	0.66	5.66	0.33	0.00	9.40
2	25	42.30	1.00	9.00	10.00	4.33	18.50
3	26	71.30	6.66	9.66	17.67	5.66	25.30
4	27	85.30	9.33	22.00	28.67	7.66	28.50
5	28	89.33	21.00	32.66	47.00	18.33	32.25
6	29	96.70	36.00	49.00	66.6	67.33	58.5
7	30	99.00	51.60	61.00	75.33	72.00	67.33
8	31	99.00	69.66	75.33	96.33	91.00	70.50
9	32	100.00	70.00	82.66	96.66	96.00	75.25
10	33	100.00	75.66	94.33	98.00	97.00	76.75
11	34	100.00	78.00	94.33	98.00	97.00	78.25
12	35	100.00	78.33	94.66	98.00	97.00	79.25
13	36	100.00	80.00	95.66	98.00	97.00	80.50
14	37	100.00	82.67	96.66	98.00	97.00	82.25
15	38	100.00	92.00	96.66	98.00	97.00	83.75
16	39	100.00	95.33	97.00	98.00	97.00	84.20
17	40	100.00	95.33	97.00	98.00	97.00	85.00

Source: Experimental Data, Department of Pathology, CCS HAU, Hisar

Table 2. Descriptive statistics of the Cotton leaf curl disease severity from 2017 to 2022 for the Hisar region in Haryana

Year	Mean	Variance
2017	80.21	1087.34
2018	51.96	1431.17
2019	71.49	1205.50
2020	70.93	1452.43
2021	66.62	1736.70
2022	58.04	921.88

Table 3. Correlation coefficient of Cotton leaf curl disease severity with weather variables from the years 2017 to 2022 for the Hisar region in Haryana

Year	Tmax	Tmin	Rhm	Rhe	WS	BS	Total Rainfall
2017	-0.452	-0.008	0.594*	0.167	-0.510*	0.221	-0.493*
2018	-0.664**	-0.666**	0.731**	0.261	-0.484*	0.285	-0.135
2019	-0.619**	-0.050	0.748**	0.412	-0.485*	-0.046	-0.379
2020	-0.75**	-0.496*	0.776**	0.351	-0.465	-0.437	-0.116
2021	-0.791**	-0.412	0.809**	0.752**	-0.310	-0.604*	0.343
2022	-0.623**	-0.348	0.702**	0.453	-0.626**	0.071	0.223

Tmax: Maximum Temperature, Tmin: Minimum Temperature, rhm: Relative Humidity Morning, Rhe: Relative humidity evening, WS; Wind Speed; BS; Bright Sun shine; Rainfall

Table 4. Parametric Coefficient and goodness of fit criteria for the various fitted models for the years 2017 to 2022

Parameters	Logistic	Monomolecular	Gompertz
2017			
A	99.64	101.05	100.17
B	3.43	0.81	1.60
K	0.99	0.49	0.73
R ²	0.99	0.97	0.96
MSE	2.06	6.09	2.09
2018			
A	87.12	133.15	90.56
B	44.30	1.08	7.13
K	0.67	0.09	0.41
R ²	0.96	0.95	0.99
MSE	22.30	51.41	16.25
2019			
A	97.55	122.83	99.73
B	24.25	1.06	5.50
K	0.62	0.13	0.41
R ²	0.99	0.95	0.99
MSE	2.60	64.31	8.75
2020			
A	98.90	109.87	100.17
B	25.88	1.10	5.68
K	0.79	0.20	0.53
R ²	0.99	0.94	0.98
MSE	6.04	27.25	12.71
2021			
A	99.60	117.08	97.06
B	563.64	1.16	74.62
K	1.33	0.16	0.99
R ²	0.98	0.88	0.98
MSE	17.32	181.88	15.97
2022			
A	83.30	95.04	85.89
B	8.04	0.96	2.57
K	0.53	0.16	0.34
R ²	0.98	0.95	0.97
MSE	12.39	30.98	17.16

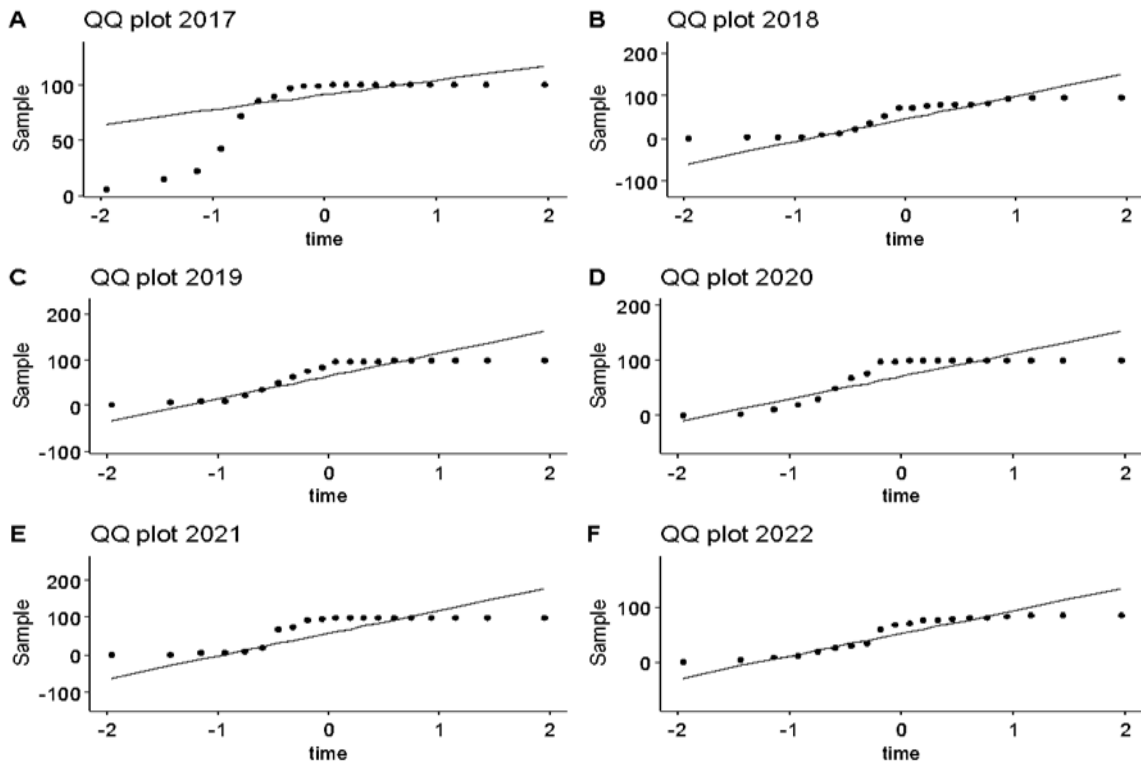


Fig. 1. QQ plot of data for the years 2017 to 2022

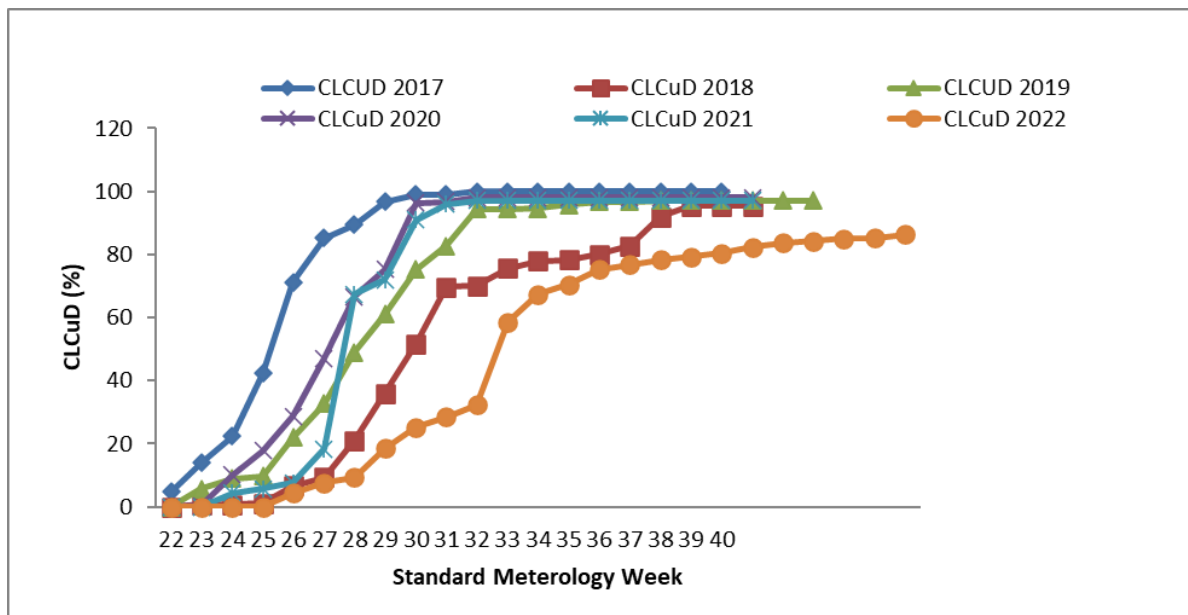


Fig. 2. Disease Progress Curve of cotton leaf curve disease (%) for the year 2017 to 2022

followed S-shaped curves, specifically the Logistic and Gompertz models. Their interpretation highlighted that such patterns emerge due to the biological interaction between pathogen multiplication and host susceptibility over time—a phenomenon also observed in the current study. This reinforces the biological relevance and appropriateness of selecting Logistic and Gompertz models for CLCuD data, as both diseases exhibit sigmoidal progression due to infection dynamics.

Moreover, Sain *et al.* (2024) advanced the modelling

approach by applying machine learning tools to predict CLCuD under varying ecological conditions. They split the data into training and validation sets and compared the performance of several ML models. The Best Subset Feature (BSF) model emerged as the most accurate with the highest R^2 value, indicating its strong predictive capacity, while the Artificial Neural Network (ANN) model performed relatively poorly. This highlights a key distinction: while traditional growth models (like those used in the current study) offer high inter-

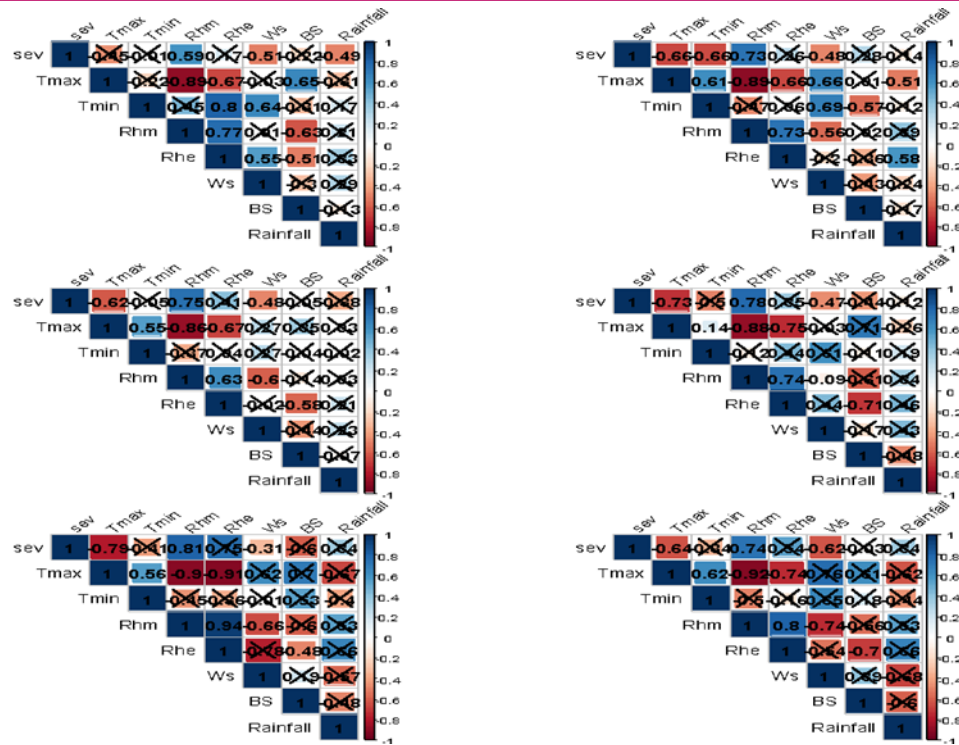


Fig. 3. Correlogram of disease severity with weather variables for the years 2017 to 2022

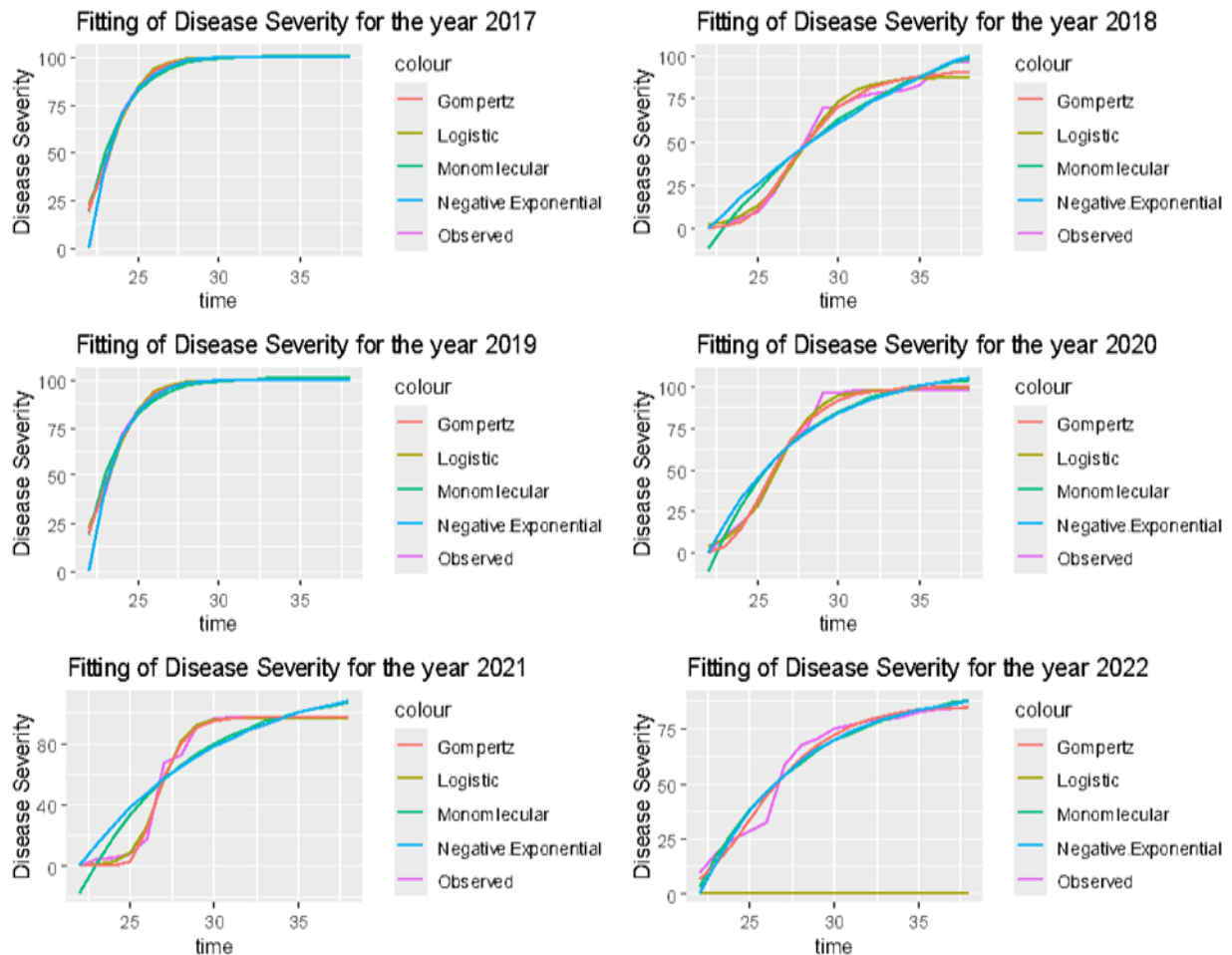


Fig. 4. Fitting of observed and predicted values for the years 2017 to 2022

Table 5. Observed and Predicted values for different best-fitted models

2017		2018		2019	
Observed	Logistic	Observed	Gompertz	Observed	Logistic
22.30	22.51	0.66	0.07	5.66	3.86
42.30	44.06	1.00	0.82	9.00	7.01
71.30	68.05	6.66	4.05	9.66	12.38
85.30	85.10	9.33	11.67	22.00	20.92
89.33	93.75	21.00	23.44	32.66	33.05
96.70	97.39	36.00	37.13	49.00	47.83
99.00	98.80	51.60	50.29	61.00	62.79
99.00	99.33	69.66	61.44	75.33	75.34
100.00	99.53	70.00	70.11	82.66	84.31
100.00	99.60	75.66	76.49	94.33	90.02
100.00	99.63	78.00	81.02	94.33	93.39
100.00	99.64	78.33	84.15	94.66	95.29
100.00	99.64	80.00	86.28	95.66	96.33
100.00	99.64	82.67	87.71	96.66	96.89
100.00	99.64	92.00	88.67	96.66	97.20
100.00	99.64	95.33	89.31	97.00	97.36
100.00	99.64	95.33	89.74	97.00	97.45
2020		2021		2022	
Observed	Logistic	Observed	Gompertz	Observed	Logistic
0.33	3.68	0.00	0.17	9.40	9.21
10.00	7.78	4.33	0.65	18.50	14.54
17.67	15.69	5.66	2.40	25.30	22.04
28.67	29.09	7.66	8.52	28.50	31.62
47.00	47.40	18.33	25.92	32.25	42.48
66.60	66.30	67.33	56.19	58.50	53.23
75.33	80.89	72.00	81.20	67.33	62.53
96.33	89.85	91.00	92.00	70.50	69.69
96.66	94.59	96.00	95.35	75.25	74.72
98.00	96.90	97.00	96.27	76.75	78.03
98.00	97.99	97.00	96.52	78.25	80.12
98.00	98.49	97.00	96.58	79.25	81.40
98.00	98.71	97.00	96.60	80.50	82.17
98.00	98.82	97.00	96.60	82.25	82.63
98.00	98.86	97.00	96.60	83.75	82.90
98.00	98.89	97.00	96.60	84.20	83.06
98.00	98.90	97.00	96.60	85.00	83.16

pretability and are suitable for modelling disease over time with limited variables, ML models can harness multidimensional data and offer enhanced prediction under complex conditions, provided that careful feature selection and validation are undertaken. Although the present study did not use ML tools, its findings are consistent with the temporal prediction aspect of Sain *et al.*'s work and can serve as a foundational step for more advanced modelling.

In parallel, Buttar and Pritpal Singh (2017) developed an environment-based prediction model for CLCuVD by incorporating four years of disease incidence data, whitefly population data, and meteorological variables in Punjab. Their model demonstrated moderate predictive power, with R^2 values of 0.82 and 0.78, and a reasonable match between the observed and predicted slopes (5.85 and 6.68). Their findings underscore the significance of environmental factors in disease devel-

opment and the practical utility of these models for disease forecasting and targeted pesticide application. These insights align well with the correlation analysis section of the current study, where relative humidity and maximum temperature were found to have significant relationships with disease severity. This not only validates the inclusion of weather variables in predictive modelling but also confirms the consistency of environmental influences across geographies and study designs.

Overall, the comparison underscores that while machine learning models like those in Sain *et al.* (2024) provide sophisticated, high-dimensional predictions, and environment-based models like Buttar and Singh (2017) aid in practical decision-making, the current study's nonlinear regression models offer a robust, interpretable framework, particularly effective in characterizing disease progression over time. Together, these

approaches form a complementary suite of tools for understanding, predicting, and managing CLCuD, with potential for integration in future studies for better decision support in disease surveillance and control strategies.

Conclusion

The present study successfully validates the use of nonlinear growth models over simple linear regression for predicting CLCuD intensity. The Logistic and Gompertz models emerged as the most reliable tools for different years, showing superior statistical fit and alignment with the biological behaviour of diseases. These findings are strongly supported and contextualized by earlier research on CLCuD and similar plant disease systems, emphasizing the need for weather-integrated, growth-based modelling frameworks for effective disease forecasting and management. To address this, various nonlinear growth models were fitted to the data. Based on model performance, it was concluded that the logistic growth model provided the best fit for the years 2017, 2019, 2020, and 2022, whereas the Gompertz model was found to be most appropriate for 2018 and 2021. Using the best-fitted models for each year, disease predictions were computed to support early detection and effective management of Cotton leaf curl disease (CLCuD).

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

- Berger, R. D. (1981). Comparison of the Gompertz and logistic functions to describe plant disease progress. *Phytopathology*, 71, 716–719. <http://dx.doi.org/10.1094/Phyto-71-716>
- Bird, J., & Maramorosch, K. (1978). Viruses and Virus diseases associated with whiteflies. *Advances in Virus Research*, 22, 55–110. [https://doi.org/10.1016/S0065-3527\(08\)60772-1](https://doi.org/10.1016/S0065-3527(08)60772-1)
- Buttar, D.S. and Singh, Pritpal (2017). Cotton leaf curl virus disease (CLCuVD) predictive model based on environmental variables. 87(5):25-32. <https://doi.org/10.56093/ijas.v87i5.70199>
- Campbell, C. L., & Madden, L. V. (1990). *Introduction to plant disease epidemiology*. John Wiley and Sons, New York, NY, USA.
- Chugh, R. K., Kumar, M., & Kumar, S. (2020). Growth modeling for prediction of cotton leaf curl disease (CLCuD). *Journal of Cotton Research and Development*, 34(2), 291–300.
- Fantaye, A.K. (2022). Modelling and Stability Analysis of Cotton Leaf Curl Virus (CLCuV) Transmission Dynamics in Cotton Plant. *Journal of Applied Mathematics*: 1-12. <https://doi.org/10.1155/2022/6988197>
- Forrest, F. W. N. (2007). The role of plant disease epidemiology in developing successful integrated disease management programs. In Ciancio, A., & Mukerji, K. G. (Eds.), *General concepts in integrated pest and disease management* (Vol. 1, pp. 45–79). Springer, Dordrecht, The Netherlands.
- Jamadar, M. M., Venkatesh, H., Balikai, R. A., & Patil, D. R. (2009). Forecasting of powdery mildew disease incidence on ber (*Ziziphus mauritiana* Lam.) based on weather. *Acta Horticulturae*, 840, 63.
- Kapur, S. P., Singh, J., Chopra, B. L., Sohi, A. S., Rawal, H. S., & Narang, D. D. (1994). Cotton leaf curl disease in Punjab. *Plant Diseases Research*, 9(1), 86–90.
- Kumar, P., Kumar, M., & Rani, S. (2019). Estimation of annual compound growth rates of guava (*Psidium guajava* L.) fruit in Haryana using nonlinear model. *Journal of Applied and Natural Science*, 11, 778–784. <https://doi.org/10.31018/jans.v11i4.2175>
- Kumar, S. K., Sain, D., & Monga, D. (2019). Study on correlation between population of viruliferous whitefly and the percent intensity of cotton leaf curl disease in cotton. *International Journal of Current Microbiology and Applied Sciences*, 8, 922–937.
- Monga, D., & Sain, S. K. (2021). Incidence and severity of cotton leaf curl virus disease on different BG II hybrids and its effect on the yield and quality of cotton crop. *Journal of Environmental Biology*, 42, 90–98. <http://doi.org/10.22438/jeb/42/1/MRN-1296>
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Sain, S. K., Monga, D., Mohan, M., Sharma, A., & Beniwal, J. (2020). Reduction in seed cotton yield corresponding with symptom severity grades of cotton leaf curl disease (CLCuD) in upland cotton (*Gossypium hirsutum* L.). *International Journal of Current Microbiology and Applied Sciences*, 9, 3063–3076. DOI: 10.20546/ijcmas.2020.91.1.372
- Sain, S.K., Paul, Debashis, Kumar, Pradeep, Kumar, Ashok, Mohan, Man, Monga, Dilip, Prakash, A.H., Prasad, Yenumula (2024). Cotton leaf curl disease (CLCuD) prediction modeling in upland cotton under different ecological conditions using machine learning tools. *Ecological Informatics*, 81, 1-10 (<https://doi.org/10.1016/j.ecoinf.2024.102648>)
- Sattar, M. N., Iqbal, Z., Tahir, M. N., & Ullah, S. (2017). The prediction of a new CLCuD epidemic in the Old World. *Frontiers in Microbiology*, 8, Article 631. <https://doi.org/10.3389/fmicb.2017.00631>
- Van Maanen, A., & Xu, X. M. (2003). Modelling plant disease epidemics. *European Journal of Plant Pathology*, 109, 669–682. <https://link.springer.com/article/10.1023/A:1026018005613>
- Vander Plank, J. E. (1963). *Plant Disease Epidemics and Control*. Academic Press, New York, USA.
- Xu, X. (2006). Modelling and interpreting disease progress in time. In Cooke, B. M., Gareth Jones, D., & Kaye, B. (Eds.), *The Epidemiology of Plant Disease* (pp. xx–xx). Springer, Dordrecht, The Netherlands.