Machine Learning for Human Intelligence
2025, Vol. 3, Issue 1

Research Article

Predicting Cotton Whitefly Populations Through Deep Learning

Sana Basharat[1*], Najeeb Khan[2], and Awad Bin Naeem[3]

[1,2,3]Department of Computer Science, University of Management and Technology, Lahore, Punjab, Pakistan.; Email: sana.basharattt@gmail.com, najeebkhanmcs@gmail.com, awadbinnaeem@gmail.com
*Corresponding author: Sana Basharat (sana.basharattt@gmail.com)

Article History

Academic Editor:
Dr. Muhammad Sajid
Submitted: December 11, 2024
Revised: February 16, 2025
Accepted: March 1, 2025

Keywords:
Cotton Forecasting, Deep Learning, Whitefly Demographics, Feature Extraction

Abstract

Agriculture is the principal foundation upon which a nation's economy relies. Pakistan ranks as the fourth-largest cotton grower globally, establishing it as a principal cash crop. Their inability to determine the most suitable cotton variety for their climate is attributed to governmental rules, agricultural dangers, and a low literacy rate among the populace and farmers. This approach use a temperature-based model to categorize various plant species. The pest population is primarily concerned with its diversity, the impact of weather on its numbers, and the fluctuations between high and low populations. The main objective of this study was to provide a framework capable of managing the intricate dynamics of cotton whitefly populations. Our primary objective was to ascertain the number of insects, including their eggs and progenitors. Another objective is to ascertain information regarding the variety with a diminished bug population. To acquire preliminary insights into individuals opposed to the temperature. Cotton yield could be enhanced, and the usage of chemicals should be minimized. Consequently, farmers' income may be enhanced. We will develop the optimal ARIMAX model for predicting whiteflies on cotton. The accuracy of this model was almost equivalent to that of the statistical forecasting models. As a statistical model, ARIMAX can be utilized for forecasting purposes.

1 Introduction

The agricultural sector is anticipated to make a significant contribution to Pakistan's economy. Agriculture constitutes the primary economic activity and source of wealth for the country. Gossypium hirsutum L., the scientific designation for cotton, is the preeminent crop globally for the production of textile fibers. Wool may be transformed into soft and breathable fabrics by the process of spinning it into yarn or thread. The most prevalent natural fabric in contemporary clothing is derived from this specific material. Cotton constitutes about 40% of global natural fiber production, rendering it the preeminent source of fiber. The extensive history and production of cotton are intrinsically connected to the formative phases of human civilization. Referred to as "White Gold" in Pakistan, it is a vital agricultural crop. Consequently, it profoundly affects Pakistan's agricultural sector, irrespective of its level of development. Oil, lint, and hulls are but a few things derived from them, all produced in warmer regions. Pakistan ranks fourth globally in cotton production. Eight percent of the nation's oil supply is derived from cotton cultivation.

Ginning and turning operations utilize cotton as their principal raw material. Cotton's diverse use render it indispensable to the economy, encompassing food, fuel, fibers, and foreign exchange. The influence of cotton on our society and culture extends well beyond economic factors. The quality and quantity of cotton crops have dramatically diminished due to factors such as pests, climate change, and diseases. They are unable to determine the most suitable cotton variety for their climate due to governmental regulations, agricultural dangers, and a suboptimal literacy levels among the general populace and agriculturalists. It is estimated that there are approximately 1,326 unique species of cotton-consuming insects globally. Nevertheless, the Indian subcontinent possesses only 162 insect species that infest cotton. A study indicates that Pakistani cotton is infested with 93 pests and insects. Experiments demonstrate that contemporary technologies and innovative procedures must be employed to execute revolutionary approaches. Crop cultivation, protection, harvesting, transportation, and storage collectively influence the regulation of Pakistan's cotton sector.

Pakistan ranks as the fourth-largest producer of cotton globally, establishing it as a principal cash crop. They are unable to determine the most suitable cotton variety for their climate due to governmental regulations, agricultural dangers, and a low literacy rate among the populace and farmers. This approach use a temperature-based model to categorize various plant species. The pest population is primarily concerned with its diversity, the impact of weather on its numbers, and the fluctuations between high and low populations. The primary issue is to establish a framework capable of managing the intricate process involving the cotton whitefly population. Our primary objective was to understand the relationships between insects, their eggs, and their progenitors. Another objective is to obtain information regarding the variety characterized by a low bug population. To acquire preliminary insights into individuals opposed to the temperature. Cotton yield may be enhanced, and the utilization of pesticides should be minimized. Consequently, farmers' income may be enhanced. The primary objectives of conducting this research are:

1. To comprehensively comprehend the insect population, encompassing its eggs and progenitors.
2. For additional information about the plant species characterized by a minimal insect population.
3. To acquire initial data pertaining to the population in relation to temperature.

This document is structured as follows: Section 2 presents the literature review. Section 3 delineates the methodology of the investigation. Section 4 addresses the findings. Section 5 delineates the conclusion and outlines future work.

2 Literature Review

This section outlines the utilization of a fungal pathogen, namely the native strain V.lecanii, for the management of aphids and whiteflies in Korea. V. lecanii CS-625 exhibited the highest death rate of cotton aphids among the six isolates collected in Korea [5]. Following bioassay and field study, the strain CS-626 was chosen as the test subject, resulting in the development of a biological control agent for whiteflies. Pesticides exterminate insects. Dimethomorph and procymidone did not influence the germination of V. spores or the development of mycelium [17]. The fluctuations in insect populations are substantially affected by climatic conditions. Pest forecasting is an effective technique for monitoring and managing detrimental insects, particularly in regions where pest control is costly [4]. The Pests Warning and Quality Control of Pesticides Department of Agriculture, Government of Punjab, Multan, gathered data on pests over a five-year period from several locations within the Multan district to develop a pest forecasting model [27]. The district comprises multiple such places. The relationship between weather and infestations of sucking insects was summarized using multivariate regression and correlation analysis methods. This study analyzed the population averages of the Jassid Amrasca large tulle, along with those of other insects such the Whitefly, Thrips, Cotton Mealybug, Dusky Cotton Bug, and Cotton Leaf Curl Virus (CLCuV) [3]. The findings indicated that all examined insects exhibited a positive correlation with relative humidity. From 2006 to 2008, cotton mealybug populations proliferated rapidly; however, subsequent to that period, they began to diminish due to alterations in environmental conditions [26]. Investigations were undertaken regarding the return of the Dusky cotton insect as a novel problem in cotton cultivation. A regression analysis demonstrated a negative linear relationship between maximum temperature and whitefly population, accounting for 5.9 to 21.6% of the variance. Rainfall exerted a detrimental linear regression impact on the Jassid population, ranging from 1.3 to 3.4% [25]. The project's research was conducted in 2008 and 2009 during the Kharif seasons. The detrimental effects of feeding by sucking insect pests were examined in 19 genotypes, comprising 17 Bt hybrids, one conventional hybrid, and one variety, without the application of insecticides. Observations of sucking insects were systematically documented on a weekly basis [24].

The evolutionary history and ecological implications of the Bemisia tabaci complex have been examined in Ecuador. Whitefly samples were obtained in Ecuador from nine distinct locations, with latitudes ranging from 2 to 5 degrees south and longitudes from 78 to 81 degrees west. Genetic and adjusted pairwise distance analyses were conducted areas. Consequently, the mitotypes can be identified (832 bp) [22]. MaxEnt was employed to model mitotype distributions and predict mitotype niches in relation to environmental gradients. An imported B mitotype of B exists. tabaci, alongside the three indigenous ECU mitotypes. Mitotypes ECU1 44%, ECU2 0.74%, and ECU3 1.47% were identified in the American Tropics AMTROP species. The mitotypes exhibited divergence of up to 10%, exceeding previous AMTROP predictions [23].

During the later stages of the novel coronavirus, various factors are influencing individuals' dietary habits, and the current forecasting method is hindered by redundant data, leading to diminished predictive accuracy [21]. A predictive model was selected within the framework of big data, and its forecasting accuracy was examined to avert a discrepancy between supply and demand during a catastrophic health crisis. Following multiple enhancements and precision evaluations, the revised grey Verhulst emerged as the most dependable and optimal model for short-term forecasting [19].

Cassava production in Africa is significantly hindered by disease outbreaks and insect infestations. The cultivation of crops in Africa faces a significant threat from an estimated annual yield loss of $1.5 billion caused by the cassava mosaic virus disease (CMVD), which is induced by cassava mosaic Gemini viruses and disseminated by the whitefly Bemisia tabaci Gennadius [28]. To protect cotton crops, they utilized pesticides such as spirotetramat and buprofuzin [1]. Whiteflies were enumerated prior to the application of insecticides and again at 2, 3, and 10 days post-application. An exhaustive review of the data revealed that each pesticide generates a unique array of effects. Diafenthiuron and spirotetramat, however, markedly influenced the population, reducing it to a predetermined level [16].

A study by Shabbir et al. indicates that the Cassava mosaic disease SLCMV, an Asian affliction attributed to the Begomovirus genus and belonging to the Geminiviridae family, has recently threatened the cassava cultivation business [12]. The transmission of SLCMV remains largely uncharted in recently utilized areas. Their investigation of the transmission efficacy of three whitefly species across Asia reveals that Asian whiteflies can solely transmit SLCMV successfully. The viral load in the whitefly's entire body was found to be favorably correlated with the transmissible quantity of SLCMV by different whiteflies [10]. The efficacy of viral transmission varies across various whitefly species. This study conducts comprehensive research on whitefly transmission of SLCMV, contributing to the understanding of SLCMV in agricultural contexts and aiding in viral epidemic predictions [11].

Forty to fifty percent of the annual reduction in output caused by infestations, both pre- and post-harvest, can be ascribed to insects, pests, and diseases [9]. The decline in potential productivity is more pronounced in Asia and Africa. Transnational insect pests, inadequate access to crop protection agents, and insufficient preparedness are the contributing factors to this [8]. It is essential to devise innovative, cost-effective, climate-resilient pest control methods that are digitally optimized for real-time pest forecasting and decision support systems (DSSs) [7]. Efforts have been made to create weather-based forecasting models for pests of Chickpea and Pigeonpea, including Helicoverpa and Phytophthora blight, to predict the first emergence, peak severity/population, and crop age. Stepwise regression and multiple meteorological indices were employed to construct the models. Methods employing machine learning, artificial neural networks, and Bayesian networks. The models generally align well with all available data, demonstrating strong concordance between the forecasts and the actual conditions [6].

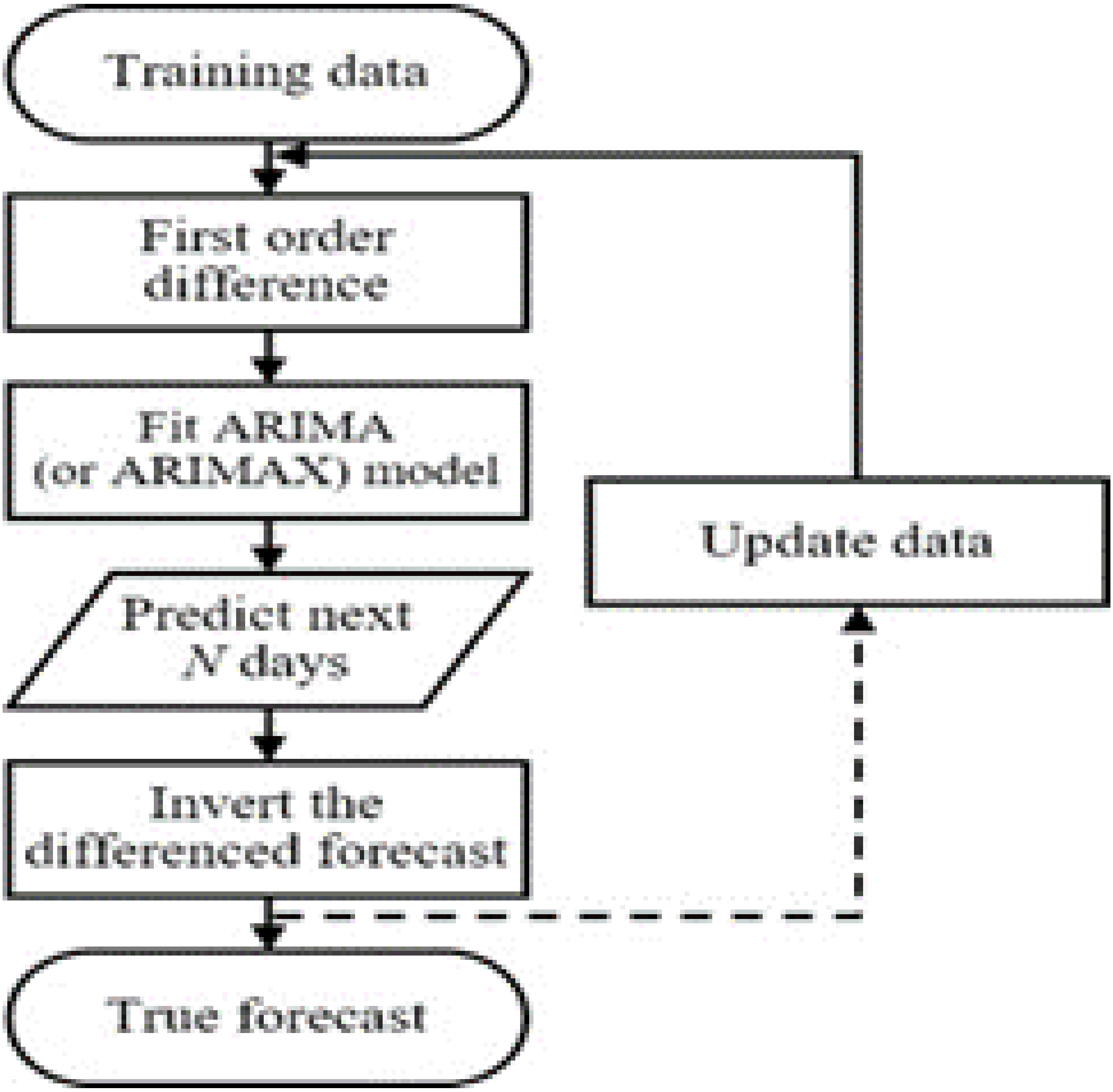

Figure 1: Model Training Flow Diagram

3 Materials and Methods

This section details an experimental technique designed to assess the classification accuracy of the proposed model. This research collected two datasets from a greenhouse, encompassing whitefly population, temperature, humidity, and light intensity, utilizing a wireless imaging apparatus. The system comprises wireless imaging devices utilizing Raspberry Pi 3 B+ embedded systems and Raspberry Pi v2 cameras. Two CNN models were utilized to analyze the yellow sticky paper photos, with the initial model excluding the insect objects and the subsequent model categorizing them. The whitefly population data was acquired by an insect pest counting method with 97% accuracy in whitefly identification.

Specifically, 10 devices were installed in greenhouse No. 1 in Yunlin, Taiwan, encompassing an area of 2208 m^2, and 7 machines were deployed in greenhouse No. 2 in Chiayi, Taiwan, covering an area of 529 m^2. Tomato seedlings constitute the primary crop in both greenhouses. The daily increase in whitefly population is determined by the raw count of whiteflies observed in a single day. The data from greenhouse No. 1 was utilized for model training and testing between May 5, 2018, and December

9, 2018. Data from February 22, 2018, to September 10, 2018, was utilized for model validation in greenhouse No. 2.

The Autoregressive Integrated Moving Average (ARIMA) model is a method for fitting time series data to generate future forecasts based on existing data. ARIMA comprises three theoretical components: auto-regressive (AR), moving-average (MA), and integrated (I) terms. ARIMAX converts ARIMA models into multiple regression models by using exogenous variables X.

3.1 Data Acquisition

Prior to model fitting, the input data must be stationary. The mean and variance of fixed data remain constant throughout time, facilitating forecasting. This study used the Augmented Dickey-Fuller (ADF) test ($\alpha = 0.05$) to ascertain the stationarity of the data. Out of a total of 199 data points, the most recent 79 from greenhouse No. 1 are utilized for model testing as given in Table 1.

Table 1: Augmented Dickey-Fuller Test for Stationarity Assessment

Variables	Count	Temperature (°C)	Relative Humidity (%RH)	Illuminance (lux)
Raw data	0.22	0.11	0.57	0.04
First-order difference	2.13×10^{-6}	3.52×10^{-19}	2.21×10^{-19}	2.83×10^{-7}

This study forecasts 7 days of expected data after each day. The forecasting algorithm was iteratively retrained to predict the subsequent day's whitefly count by incorporating freshly observed correct data into the training dataset.

3.2 Data Preparation and Standardization

Data normalization and transformation follow data pretreatment in the data preparation process. Data preprocessing encompasses 1) data reduction and 2) null value elimination. The selected pest monitoring dataset functions as the input for the data preprocessing portion of the investigation. Owing to unanticipated circumstances, such as adverse weather conditions or the absence of a scout, the raw data may exhibit missing records for some dates during the research period. An artificial intelligence method is likely to yield biased outcomes when provided with data derived from open records. Z-score normalization is utilized on the preprocessed data to achieve a consistent distribution of values among various attribute categories.

3.3 Proposed Structure

In an IoT design, sensors constitute the primary input layer, known as the perception layer. The Wi-Fi modules in the gateway layer facilitate data transmission from the input layer to the server and enable data storage in the storage layer. Ultimately, an application layer disseminates pest predictions to farmers through the RBFN algorithm, which utilizes test and training data gathered from the storage level to construct and provide forecasts.

The sensors were capable of managing the novel notion of predicting insect infestations in the field, which constitutes the strength of the challenge. A notable strength of the issue is the acquisition of real-time data from the region. Agriculturalists can implement safety protocols based on their surroundings. RFID technology at the gateway layer, sensor technology at the input layer, Wi-Fi communication modules at the gateway layer, and cloud computing at the storage layer constitute the most vital Internet of Things technologies.

A deep learning technique called RBFN was employed to obtain the specified prediction. The efficacy of RBFN in binary classification, coupled with its dataset independence, resulted in its designation as the optimal choice. Deep learning possesses a considerable superiority over machine learning. The proposed model employs RBFN for five specific reasons:

1. When trained on extensive datasets, the Radial Basis Function Network (RBFN) exhibits superior accuracy.
2. The reliability of RBFN increases with the processing of substantial data volumes.
3. As the volume of data increases, accuracy enhances.
4. RBFN methodologies possess an efficient predictive decision support system.
5. Despite a challenging issue, RBFN demonstrates a high degree of accuracy.

The RBFN algorithm comprises three components: an input layer, a hidden layer, and an output layer. The data collected by the Pest Environment Monitoring System from its sensors is located in the X input layer (PEMS). The input data includes temperature, humidity, precipitation rate, and wind speed. The output layer displays the result as either Yes or No as shown in Figure 1. This

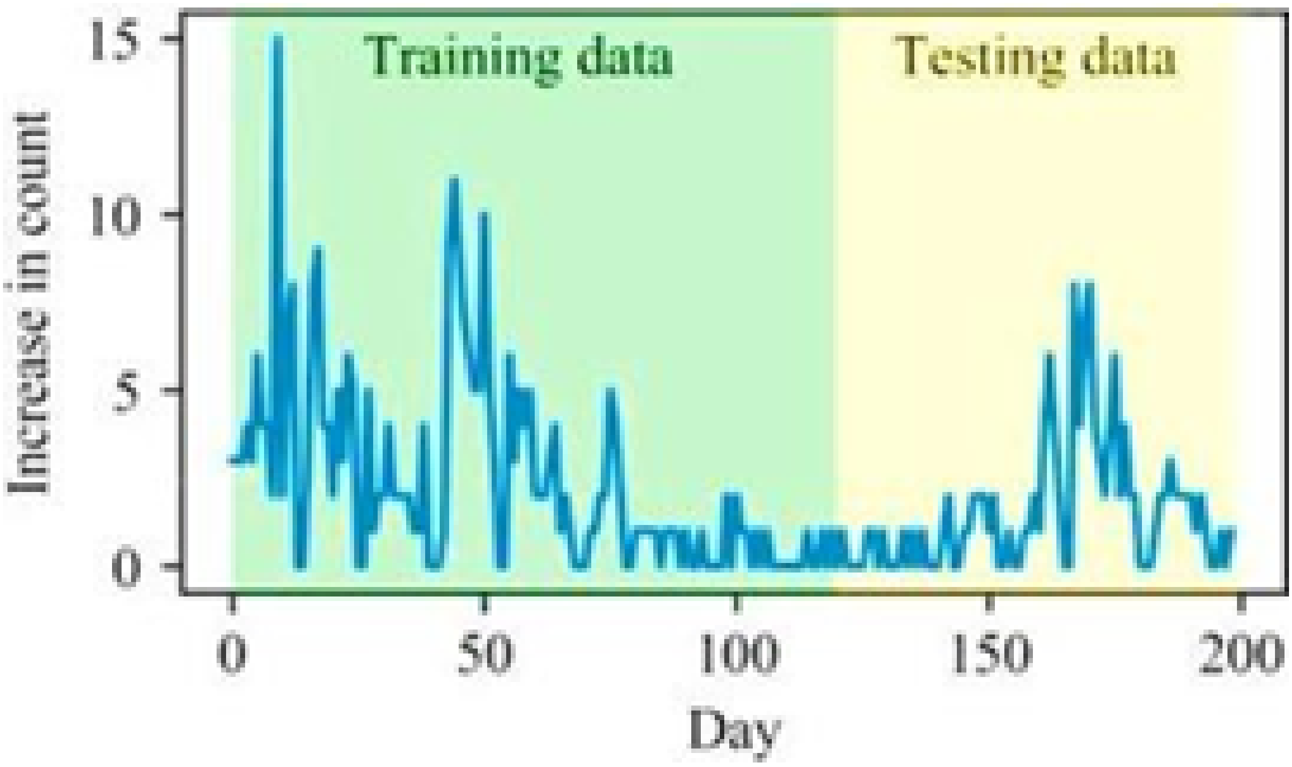

Figure 2: Raw whitefly count

section presents an example of Radial Basis Function Network (RBFN) implementation, including the experimental settings, layout of the experimental area, and deployment of the prototype model. The experiment will be conducted in the second season, spanning from May to November. The study will occur on one acre (43,560 ft^2) with dimensions of 208 feet by 208 feet. A dataset including 416 rows and 62 columns was employed to examine whitefly infestations. Each column has 416 plants. Two cotton plants fill one square foot of area. The prediction for the whitefly population encompasses a total of 12,896 cotton plants.

A sensor model is incorporated into the design to monitor temperature, humidity, precipitation, and wind speed factors. The DHT-22 is a temperature sensor that delivers highly accurate measurements of temperature and humidity in the ambient environment. A rainfall detection sensor allows the user to select between digital and analog outputs. The model employs a wind speed sensor to ascertain the prevailing wind speed in the vicinity. The microprocessor utilized in this model is identified as a WeMos D1 Wi-Fi UNO-based ESP8266 shield for Arduino.

The web application is developed using PHP, while the IoT web server employs MySQL as its database management system. The online application will collect, assess, and retain environmental data. Data obtained from sensors is collected approximately every four hours. Temperature and wind speed exhibit a positive correlation with whitefly, whereas humidity and rainfall demonstrate a negative correlation. The population of whiteflies increases with elevated temperature and wind speed, whereas it decreases with heightened relative humidity and rainfall as shown in Figure 2.

4 Results and Discussion

The test and training datasets are segregated from the other data. The training data provided to the neural network throughout the training process facilitates its learning. All three strategies evaluated for pest prediction are designed to anticipate various sorts of crop pests. The three systems were developed using three distinct neural networks. The findings indicate that the proposed intelligent system for forecasting pest population dynamics of Thrips Tabaco Linde (Thrips) on cotton (Gossypium Arboreum) exhibits superior performance relative to the other two systems examined as given in Table 2.

Table 2: Analysis of Cotton Leaf Worm Using Random Sampling

Method	Random Forest		ExtraTrees		Linear		XGBoost
	RRMSE	MAPE	RRMSE	MAPE	RRMSE	MAPE	RRMSE
Baseline	29.11%	23.66%	25.28%	20.56%	46.23%	40.39%	22.77%
Previous System	29.88%	21.87%	25.93%	19.86%	34.03%	20.26%	32.32%
Temperature	29.85%	21.16%	22.10%	19.14%	26.56%	32.12%	23.73%
Thrips	29.42%	23.60%	23.52%	20.61%	45.46%	30.235%	30.56%
Aggregate	29.46%	23.14%	26.10%	18.13%	34.39%	20.26%	26.31%

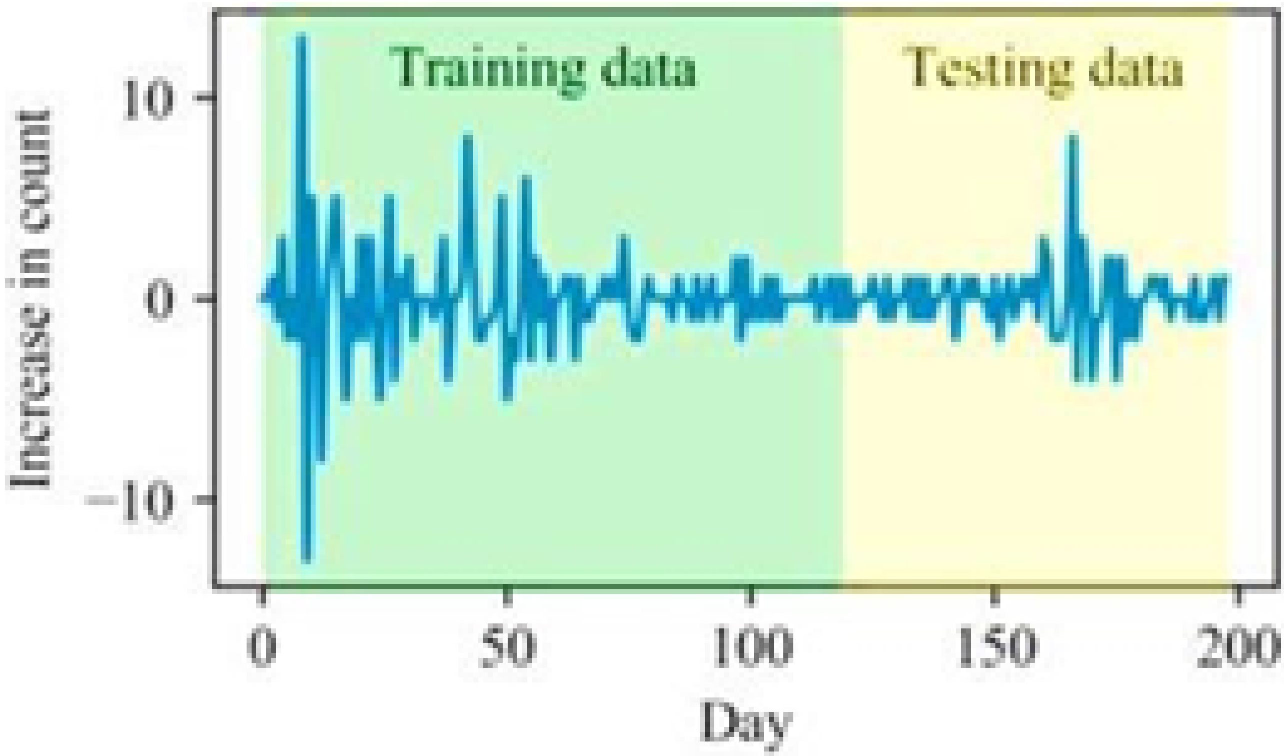

Figure 3: Difference of Whitefly Count

Table 2 presents a model (RBFN) designed to forecast whitefly infestations. Four sensors (temperature, humidity, rainfall, and wind speed) along with a microcontroller were deployed in the selected area to evaluate the outcome as given in Table 3.

Table 3: The confusion metrics of results include precision, recall, F1 score, and support

Class	Precision	Recall	F1-Score	Support
0.0	0.52	0.39	0.85	29
1.0	0.94	0.99	0.85	86
Macro average	0.73	0.69	0.85	115
Weighted average	0.82	0.85	0.85	115

The efficacy of the RBFN method is assessed utilizing the Python module "sklearn.metrics," incorporating metrics such as F1 score, precision, recall, support, Cohen's kappa, ROC AUC, and log loss as shown in Figure 3.

Table 4: Comparative Analysis of Classifier Performance

Classification Algorithm	Accuracy	Sensitivity	Precision	PPA	NPA
Support Vector Machine	98.56%	98.92%	97.82%	98.90%	97.82%
Artificial Neural Network	90.10%	93.13%	83.39%	90.69%	88.92%
Bayesian	93.90%	98.83%	86.32%	91.96%	97.86%
Decision Tree	90.10%	93.12%	83.39%	90.59%	88.92%
k-Nearest Neighbors	93.28%	96.64%	87.59%	92.97%	93.43%

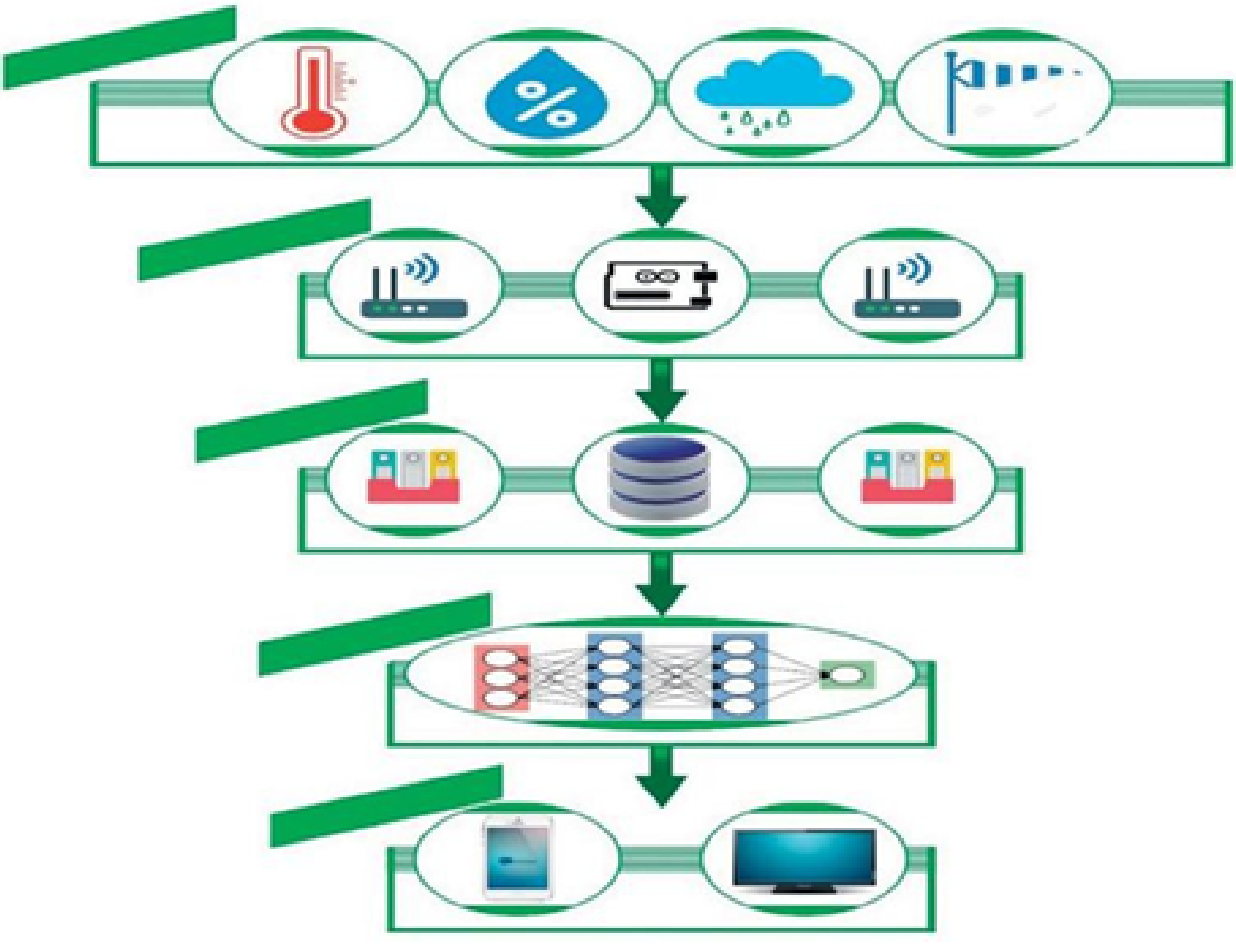

Figure 4: Proposed model

Table 4 presents the SVM classifier trained with various kernel functions in this investigation. Post-training, the experiment indicates that the SVM classifier may achieve optimal accuracy of 98.56% when employing the Radial Basis Function (RBF) kernel. The efficacy of all five trained classifiers is assessed using 130 test images (45 normal and 85 affected) that were excluded from the training dataset as shown in Figure 4.

Table 5: Forecasting Various Types of Pests and Diseases Using the LSTM Network

Measurements	Helicoverpa zea	Whitefly	Jassid	Foliar Blight
ACC	0.9315	0.9332	0.9432	0.9607
AUC	0.9768	0.9776	0.9834	0.9889
F1-score	0.8948	0.9334	0.9222	0.9281

Table 5 delineates the performance of LSTM networks across multiple datasets. This illustrates our model's exceptional generalization capability and proficiency in properly predicting illnesses and pests. The results demonstrate the LSTM network's efficacy in forecasting conditions and issues impacting cotton, establishing a theoretical basis for its prospective practical application. The suggested system employed the ARIMAX technique, and various combinations of input data were evaluated to determine the critical input data for forecasting. C represents the daily count of whiteflies, T denotes the ambient temperature (°C), H indicates relative humidity (% RH), and L signifies light intensity (lux). The smallest quantity of input training data or duration for the model was assessed to determine the method's efficacy across various greenhouses.

This study examined various training durations, including 15, 30, 60, 90, and 120 days.

Table 6: Root Mean Square Error ± Standard Deviation of Various Factors with Distinct Training Day Counts

Variables	Number of Training Days				
	15	30	60	90	120
C	3.41±3.62	2.16±1.36	2.21±1.43	2.21±1.43	2.21±1.43
C+T	2.66±1.95	2.21±1.43	1.58±1.04	1.69±1.09	1.66±1.09
C+H	2.61±1.69	1.98±1.26	1.89±0.82	1.82±1.18	1.63±0.99
C+L	2.21±1.43	2.82	1.70±1.09	1.84±1.26	1.70±1.14
C+T+H	3.14±2.30	2.21±1.48	1.62±0.72	1.84±1.26	1.56±0.91
C+T+L	3.26±2.53	2.09±1.31	1.74±1.15	1.53±1.24	1.75±1.14
C+H+L	3.65±2.87	1.95±1.27	1.82±1.26	1.92±1.27	1.72±1.13
C,T,H,L	10.56±11.88	3.32	2.92	1.67±0.95	1.62±0.96

The findings indicate that C+T+H constituted the optimal combination of input data. This indicates that with a training duration of 60 days, most models exhibit the minimal RMSE, and the RMSE of the ARIMAX models is inferior to that of the ARIMA model. The results were good at approximately 30 days, as the RMSE values were below 3, concerning the little number of input data as shown in Figure 5.

5 Conclusions and Future Research

This research employs machine learning algorithms to predict the whitefly population on cotton. The weather is considered a variable, and the population of whiteflies is forecasted. Previously, it was accomplished by the application of statistical methods. This study indicates that ARIMAX is the optimal model for predicting whiteflies on cotton, utilizing temperature and humidity as external variables. The model's accuracy was nearly equivalent to that of the statistical models employed for

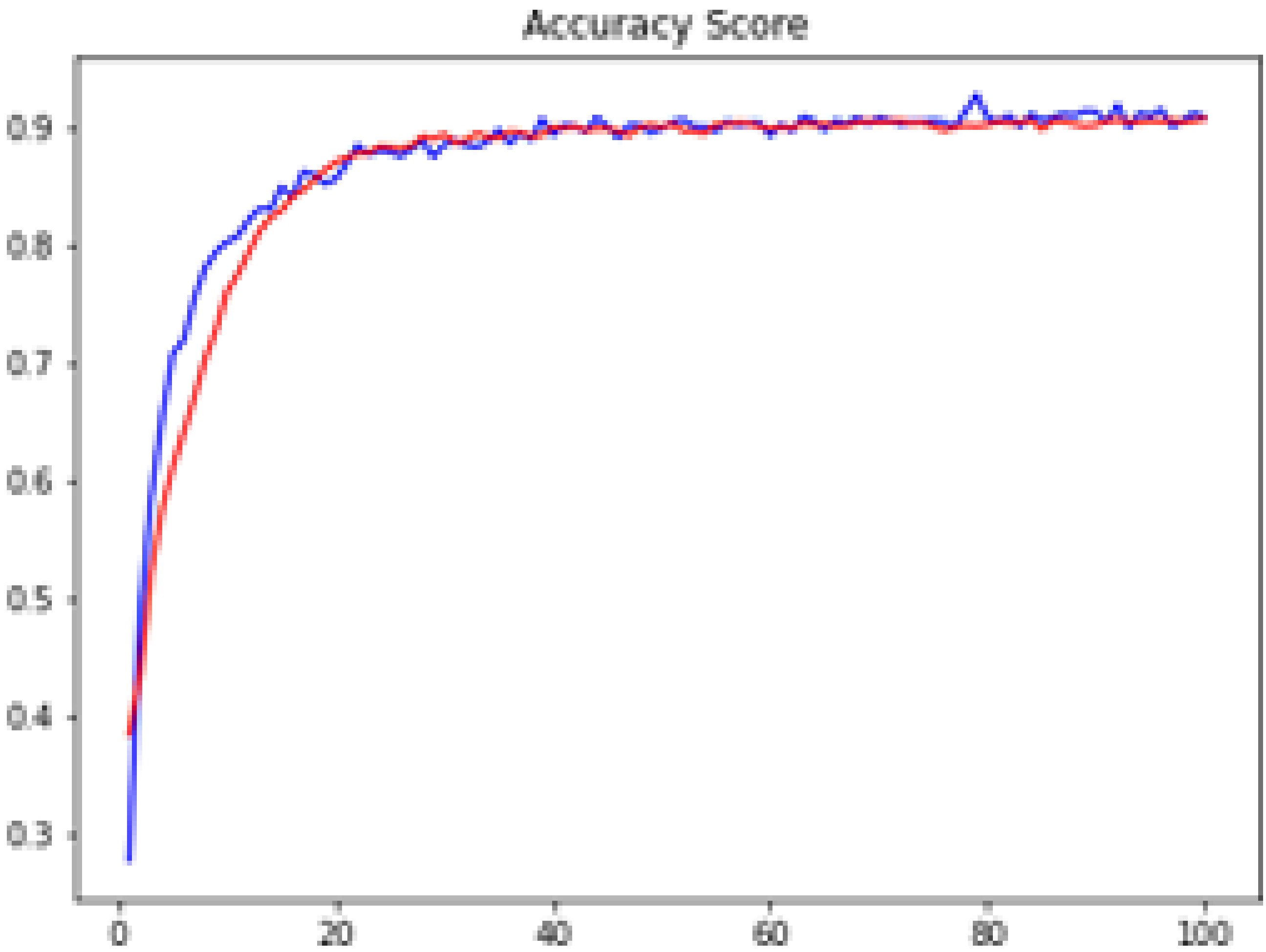

Figure 5: Accuracy and Loss Curves

predicting. The proposed methodology is more efficient than the statistical models and conserves energy. ARIMAX can be employed for forecasting due to its nature as a statistical model.

The data is collected manually in this study, requiring greater human effort and time. In the future, sensors may be employed to gather data, conserving human energy. Ultimately, a mobile application will be created that connects to the model, providing farmers with alerts on their mobile devices regarding the present whitefly population in the field, as well as forecasts of whitefly activity. To ensure the farmer is informed about the present status or population of whiteflies and can implement precautions. In the future, this approach may also be utilized for predicting additional insects and pests in cotton or alternative crops.

References

[1] Kumar, S., Jain, A., Shukla, A. P., Singh, S., Raja, R., Rani, S., ... & Masud, M. (2021). A comparative analysis of machine learning algorithms for detection of organic and nonorganic cotton diseases. *Mathematical Problems in Engineering*, 2021.

[2] Ali, S., Badar, N., & Fatima, H. (2015). Forecasting production and yield of sugar cane and cotton crops of Pakistan for 2013-2030. *Sarhad Journal of Agriculture*, 31(1), 1-10.

[3] Anwar, W., Javed, M. A., Shahid, A. A., Nawaz, K., Akhter, A., Ur Rehman, M. Z., ... & Haider, M. S. (2019). Chitinase genes from Metarhizium anisopliae for the control of whitefly in cotton. *Royal Society Open Science*, 6(8), 190412.

[4] Aswathi, V. S., & Duraisamy, M. R. (2018). Comparison of prediction accuracy of multiple linear regression, ARIMA and ARIMAX model for pest incidence of cotton with weather factors. *Madras Agricultural Journal*, 105(September 7-9), 313-316.

[5] Chen, W., Hasegawa, D. K., Kaur, N., Kliot, A., Pinheiro, P. V., Luan, J., ... & Fei, Z. (2016). The draft genome of whitefly Bemisia tabaci MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biology*, 14(1), 1-15.

[6] Chi, Y., Pan, L. L., Bouvaine, S., Fan, Y. Y., Liu, Y. Q., Liu, S. S., ... & Wang, X. W. (2020). Differential transmission of Sri Lankan cassava mosaic virus by three cryptic species of the whitefly Bemisia tabaci complex. *Virology*, 540, 141-149.

[7] Chiu, L. Y., Rustia, D. J. A., Lu, C. Y., & Lin, T. T. (2019). Modelling and forecasting of greenhouse whitefly incidence using time-series and ARIMAX analysis. *IFAC-PapersOnLine*, 52(30), 196-201.

[8] Dey, A., Bhoumik, D., & Dey, K. N. (2016). Automatic detection of whitefly pest using statistical feature extraction and image classification methods. *International Research Journal of Engineering and Technology*, 3(09), 950-959.

[9] Fenu, G., & Malloci, F. M. (2021). Forecasting plant and crop disease: an explorative study on current algorithms. *Big Data and Cognitive Computing*, 5(1), 2.

[10] Guo, R. (2021, June). Research on prediction accuracy of mathematical modeling based on big data prediction model. In *Journal of Physics: Conference Series* (Vol. 1952, No. 4, p. 042122). IOP Publishing.

[11] Amjad, A., Javed, M. S., Hameed, A., Khan, A. A., Amjad, M. R., Ali, S. A., ... & Raza, A. (2021). Improving the probiotics viability and quality characteristics of yoghurt enriched with barley bran.

[12] Swathi, M., Gaur, N., & Singh, K. Virus vector relationship of yellow mosaic virus and whitefly, Bemisia tabaci (Gennadius) in soybean. *Legume Research-An International Journal*, 1, 5.

[13] Kalkal, D., Lal, R., Dahiya, K. K., Singh, M., & Kumar, A. (2015). Population dynamics of sucking insect pests of cotton and its correlation with abiotic factors. *Indian Journal of Agricultural Research*, 49(5), 432-436.

[14] Kataria, S. K., Pal, R. K., Kumar, V., & Singh, P. (2019). Population dynamics of whitefly, Bemisia tabaci (Gennadius), as influenced by weather conditions infesting Bt cotton hybrid. *Journal of Agrometeorology*, 21(4), 504-509.

[15] Keswani, B., Mohapatra, A. G., Mohanty, A., Khanna, A., Rodrigues, J. J., Gupta, D., & De Albuquerque, V. H. C. (2019). Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms. *Neural Computing and Applications*, 31(1), 277-292.

[16] Khan, M., & Damalas, C. A. (2015). Farmers' knowledge about common pests and pesticide safety in conventional cotton production in Pakistan. *Crop Protection*, 77, 45-51.

[17] Xiao, Q., Li, W., Kai, Y., Chen, P., Zhang, J., & Wang, B. (2019). Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinformatics*, 20(25), 1-15.

[18] Khuriyati, N., Nugroho, D. A., & Wicaksono, N. A. (2020). Quality assessment of chilies (Capsicum annuum L.) by using a smartphone camera. In *IOP Conference Series: Earth and Environmental Science* (Vol. 425, No. 1, p. 012040). IOP Publishing.

[19] Kumar, S., Jain, A., Shukla, A. P., Singh, S., Raja, R., Rani, S., ... & Masud, M. (2021). A comparative analysis of machine learning algorithms for detection of organic and nonorganic cotton diseases. *Mathematical Problems in Engineering*, 2021.

[20] Xiao, Q., Li, W., Kai, Y., Chen, P., Zhang, J., & Wang, B. (2019). Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinformatics*, 20(25), 1-15.

[21] Ahmad, S. (2022). Population assessment of Bemisia tabaci (Gennadius.) and disease occurrence of Begomovirus in okra, Abelmoschus esculentus L. *Journal of the Saudi Society of Agricultural Sciences*, 21(2), 77-86.

[22] Haarith, D., Kim, D. G., Chen, S., & Bushley, K. E. (2021). Growth chamber and greenhouse screening of promising in vitro fungal biological control candidates for the soybean cyst nematode (Heterodera glycines). *Biological Control*, 160, 104635.

[23] Russia, D. J. A., Lin, C. E., Chung, J. Y., Zhuang, Y. J., Hsu, J. C., & Lin, T. T. (2020). Application of an image and environmental sensor network for automated greenhouse insect pest monitoring. *Journal of Asia-Pacific Entomology*, 23(1), 17-28.

[24] Li, Y., & Yang, J. (2020). Few-shot cotton pest recognition and terminal realization. *Computers and Electronics in Agriculture*, 169, 105240.

[25] Liu, S., Yang, Z., & Volodymyr, V. (2022). Importance of Bemisia tabaci forecasting technology: A review. *Advances in Entomology*, 10(2), 149-158.

[26] Patil, J., & Mytri, V. D. (2013). A prediction model for population dynamics of cotton pest (Thrips tabaci Linde) using multilayer-perceptron neural network. *International Journal of Computer Applications*, 67(4).

[27] Saleem, R. M., Kazmi, R., Bajwa, I. S., Ashraf, A., Ramzan, S., & Anwar, W. (2021). IoT-based cotton whitefly prediction using deep learning. *Scientific Programming*, 2021.

[28] Sseruwagi, P., Sserubombwe, W. S., Legg, J. P., Ndunguru, J., & Thresh, J. M. (2004). Methods of surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa: a review. *Virus Research*, 100(1), 129-142.

[29] Tageldin, A., Adly, D., Mostafa, H., & Mohammed, H. S. (2020). Applying machine learning technology in the prediction of crop infestation with cotton leafworm in greenhouse. *bioRxiv*.

[30] Vyas, M., Raza, A., Ali, M. Y., Ashraf, M. A., Mansoor, S., Shahid, A. A., & Brown, J. K. (2017). Knockdown of whitefly gut gene expression and mortality by orally delivered gut gene-specific dsRNAs. *PLoS One*, 12(1), e0168921.

[31] Wang, X. W., Luan, J. B., Li, J. M., Su, Y. L., Xia, J., & Liu, S. S. (2011). Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics*, 12(1), 1-12.

[32] Xiao, Q., Li, W., Kai, Y., Chen, P., Zhang, J., & Wang, B. (2019). Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinformatics*, 20(25), 1-15.

[33] Meena, R. S., Ameta, O. P., & Meena, B. L. (2013). Population dynamics of sucking pests and their correlation with weather parameters in chilli, Capsicum annum L. crop. *The Bioscan*, 8(1), 177-180.