








Research Article

Monitoring Population Phenology of Asian Citrus Psyllid Using Deep Learning

Maria Bibi ¹, **Muhammad Kashif Hanif** ¹, **Muhammad Umer Sarwar** ¹,
Muhammad Irfan Khan ¹, **Shouket Zaman Khan** ², **Casper Shikali Shivachi** ³,
and Asad Anees ^{4,5}

¹Department of Computer Science, Government College University, Faisalabad, 38000, Pakistan

²Department of Entomology, University of Agriculture Faisalabad Sub-Campus Burewala, Vehari 61010, Pakistan

³South Eastern Kenya University, Kitui, Kenya

⁴Cardiovascular Engineering Inc, 1 Edgewater Drive, Norwood, MA 02062, USA

⁵Department of Mathematics, Clausthal University of Technology, Erzstraße 1, D-38678 Clausthal-Zellerfeld, Germany

Correspondence should be addressed to Casper Shikali Shivachi; cshikali@seku.ac.ke

Received 28 May 2021; Revised 3 December 2021; Accepted 10 December 2021; Published 31 December 2021

Academic Editor: Muhammad Ahmad

Copyright © 2021 Maria Bibi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Asian citrus psyllid, *Diaphorina citri* Kuwayama (Liviidae: Hemiptera) is a menacing and notorious pest of citrus plants. It vectors a phloem vessel-dwelling bacterium *Candidatus Liberibacter asiaticus*, which is a causative pathogen of the serious citrus disease known as Huanglongbing. Huanglongbing disease is a major bottleneck in the export of citrus fruits from Pakistan. It is being responsible for huge citrus economic losses globally. In the current study, several prediction models were developed based on regression algorithms of machine learning to monitor different phenological stages of Asian citrus psyllid to predict its population about different abiotic variables (average maximum temperature, average minimum temperature, average weekly temperature, average weekly relative humidity, and average weekly rainfall) and biotic variable (host plant phenological patterns) in citrus-growing regions of Pakistan. The pest prediction models can be used for proper applications of pesticides only when needed for reducing the environmental and cost impacts of pesticides. Pearson's correlation analysis was performed to find the relationship between different predictor (abiotic and biotic) variables and pest infestation rate on citrus plants. Multiple linear regression, random forest regressor, and deep neural network approaches were compared to predict population dynamics of Asian citrus psyllid. In comparison with other regression techniques, a deep neural network-based prediction model resulted in the least root mean squared error values while predicting egg, nymph, and adult populations.

1. Introduction

The citrus greening disease, which is also known as Huanglongbing, is a severe affliction to citrus plants that causes significant losses to the citrus economy, caused by a phloem-dwelling bacterium *Candidatus Liberibacter asiaticus*. This incurable and economically damaging disease is transmitted by infection of a sucking pest, Asian citrus psyllid (ACP), which is a transmitting vector agent of *Candidatus Liberibacter asiaticus* phloem-inhabiting bacterium. Effective management of ACP is crucial for preventing the losses caused by Huanglongbing and ACP

complex [1]. There are three types of bacteria, *Candidatus Liberibacter asiaticus*, *Candidatus Liberibacter americanus*, and *Candidatus Liberibacter africanus*, associated with the spreading of Huanglongbing throughout the citrus-growing areas worldwide [2]. Huanglongbing is a vector-borne disease, and its causative agents grow and transmit through ACP [3].

Psyllid population growth rate is directly associated with the flush phenology (biotic factor) of host plants because female adults are only able to lay eggs on young, tender, and succulent plant leaves, and resultantly, the nymphs are more likely to hatch and grow during the season of abundant flush

growth on citrus plants. Availability of flush growth and optimum meteorological conditions leads to large infestations of ACP on citrus plants. Different meteorological conditions such as relative humidity, temperature, and rainfall in the study area are important factors influencing the existence of ACP stages in the field. Citrus host plant phenological characteristics have the tremendous ability to influence psyllid biology, survival, and resultant pest outbreaks under optimum environmental conditions [4–6].

Entomologists carried out various trials previously regarding population change of psyllid over time and these types of studies are significantly useful in future pest prediction and forecasting. Keeping in view, the significant effect of weather factors on insect populations and natural enemies' existence can be correlated with the changing pest-natural enemies' populations, which will better explain the density curves of both psyllid and associated insect enemies [7–9].

The science of ecology in natural sciences studies the mutual relationships among the biotic and abiotic components of the ecosystem to understand ecological processes and make predictions about future trends. Machine learning (ML) techniques have advantages over typical statistical approaches because these techniques are modeling ecological processes in a better way by allowing better decision-making and informed actions in the real world without (or with minimal) human involvement. ML techniques not only provide a flexible framework for the execution of data-driven tasks but also help for the integration of expert knowledge into the system [10].

The abilities of ML algorithms to model high-dimensional and nonlinear data with complex interactions, missing values, and identification of complicated structures from more complex datasets are defeating typical statistical approaches in population modeling [11]. Recent advancements in ML are deep learning (DL) techniques [12]. The DL approaches have the potential of automated feature learning, and the complex structures allow to solve more complex problems faster and accurately and reduce error while modeling regression problems and increasing accuracy in classification problems analysis in the presence of large dataset availability [13].

Machine learning techniques have been used in several studies related to pest's population prediction such as modeling the population dynamics of paddy stem borer (*Scirpophaga incertulas*) [14], the population density of *Scirtothrips dorsalis* Hood [15], risk of Melon thrip (*T. palmi*), diamondback moth (*P. xylostella*) [16], fluctuating trends of *Dendrolimus superans* population [17], population phenology of Black Planthopper (*Nilaparvata lugens*) [18], population occurrence of mosquitoes in correlation with different socioeconomic factors and landscape variables [19], *Prostephanus truncatus* infestation and accompanying damages to maize grain storage in correlation with abiotic factors [20], fluctuating trends of cotton's pest population (*Thrips tabaci* Linde) [21], and the effect of temperature and rainfall monitored by Watts and Worner [22] to the establishment of mealybug (*Planococcus citri*) and aphid (*Myzus persicae*, *Aphis gossypii*, *Eriosoma lanigerum*, and *Brevicoryne brassicae*).

The random forest regressor (RFR) model has been employed by researchers in various fields related to prediction and classification problems; for example, the authors of [23–25] used this ensemble learning approach prediction of dengue, citrus flatid planthopper, and sunn pest's nymphal stage, respectively. For earlier prediction of pest's risk, the multiple linear regression (MLR) model was adopted by numerous researchers. The authors of [26, 27] implemented MLR approach to model potential risk of black planthopper and oriental fruit fly (*Bactrocera dorsalis*) population, respectively.

Deep neural network (DNN) has broader applicability in the following agricultural domains in general. Chlingaryan et al. [10] used DNN for estimation of crop yield prediction. The authors of [28, 29] deployed DNN for the prediction of soil moisture contents, and Scher [30] used DNN for weather conditions' prediction. DNN has been also used for land cover and crop type classification, image identification, and classification of plants and weeds [31–34]. Rammer and Seidl [35] deployed DNN and RFR to predict damages that will occur in the future through bark beetle population outbreak using pest's historical data and concluded that DNN has the tremendous power to model bark beetle outbreaks' dynamics and other ecological prediction problems. This review of previous studies shows a research gap concerning the use of ML and DL models in the prediction of phenological stages of insects-pests. Keeping in view the literatures, the present study was conducted to (a) make comparative analysis of different machine and deep learning techniques to predict phenological stages of ACP and (b) monitor the cumulative effect of different weather factors and host plant phenology on psyllid phenological stages.

In present research, we made a comparative analysis of different regression-based approaches, i.e., DNN, MLR, and RFR models, to predict the population of different ACP-phenological stages using environmental variables and host-plant phenology variables as independent variables. By using the abovementioned regression approaches, we evaluated the combined effect of different independent variables on three ACP-phenological stages, i.e., eggs, nymphs, and adults separately.

2. Materials and Methods

2.1. Study Site and Data Collection. For data collection to monitor population dynamics of Asian Citrus Psyllid, two study locations, Square No.9 (31° 25' 50.4" N; 73° 03' 40.2" E; elevation 190 m) and PARS (N31°23'35.20"; E73°01'27.0"; elevation 210 m), were selected from University of Agriculture Faisalabad (UAF), Pakistan. From both study locations, 15 trees of two citrus species, sweet orange (*Citrus sinensis sensu latu*), and kinnow (*Citrus reticulata*) were randomly selected and tagged properly to monitor population fluctuations of ACP on weekly basis from a time course, 26 March 2011 to 20 April 2013. A detailed description of both study sites and ACP-phenological stages' data collection is given in [36]. We used datasets spanning 25 months to reduce experimental errors and to confirm the psyllid response in different weather conditions in different

seasons. If psyllid population increases in spring, then we repeated this for next spring to see if psyllid responds similarly.

Meteorological data during the experimental period regarding daily temperature (maximum, minimum, and average temperature), rainfall, and relative humidity on daily basis were documented from the meteorological observatory of the Crop Physiology (CP) department in the Agricultural faculty of UAF. The effect of meteorological (abiotic) factors was also monitored by calculating the percentage of branches infected with different life stages of ACP, i.e., eggs, nymphs, and adults individually and collectively.

2.2. Model Development. In this study, three models, i.e., RFR, DNN, and MLR, were employed to model population dynamics of ACP. The Google Collaboratory was used, and it is an effective cloud computing environment for developing python-based applications.

2.3. Random Forest Regressor. Random forest is an ensemble learning approach proposed in [36] and used both for regression and classification problems' analysis [37]. Each random forest is composed of a specified number of decision trees, and each decision tree trains on samples of training data by following a randomized approach called Bagging (Bootstrap aggregating). Random forest regressor returns the output in the form of the mean value calculated from the results of the prediction of all decision trees. It minimizes the effect of model overfitting by introducing randomness in variables and data instances' selection. RFRs have the capability of efficient training and testing. As each prediction is made by random forests (RFs), a built-in mechanism is usually found in RF to calculate test errors, e.g., root mean squared error (RMSE), mean absolute error (MAE), and confidence [38]. Hyperparameter tuning is an important step in the development of models. In order to train the RF, we set the value of `n_estimators` (number of decision trees) as 20 and `random_state`=42, while keeping other hyperparameters with their default values. We used RMSE as a loss function to calculate test errors. The mathematical formulation of RMSE is given as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}, \quad (1)$$

where y_i and \hat{y}_i are the actual and predicted values, respectively, and n is the number of observations. For evaluating the accuracies of the forecasting models, RMSE is a common indicator used in regression problems analysis [39].

2.4. Deep Neural Network. Artificial neural networks (ANNs) were developed in the middle of the nineteenth century. The term "deep learning" refers to training of deeper and larger ANNs. Here, deeper and larger are

concerned with more layers and more neurons as compared to ANNs [12]. DNNs are the result of recently developed improved algorithms which are optimizing the weights of the connections [40].

For predicting the population phenology of ACP, we developed a DNN comprised of one input layer consisting of six input neurons/nodes and two hidden layers with six and eight neurons, respectively. The activation function and optimizer used are ReLu and Adam, respectively. DNN architecture also consisted of one output layer with a single neuron to predict each ACP life stage, i.e., eggs, nymphs, and adults separately. We used dense layers to make the model more stable for prediction (Figure 1).

2.5. Multiple Linear Regression. To quantify the relationship between different input variables (Average Max Temp, Average Min Temp, Average Weekly Temp, Average Weekly RH, Average Weekly Rainfall, and Branches with Flush) and ACP phenological stages, Pearson correlation analysis was performed. We used Pearson correlation coefficient (R) values as criteria to select suitable input variables for developing the MLR model. The MLR model was deployed with a stepwise selection method to monitor the fluctuating trends of ACP population occurrence. Equation (2) for MLR is given below:

$$y = \beta_0 + \beta_{1x_1} + \beta_{2x_2} + \cdots + \beta_{kx_k} + \varepsilon, \quad (2)$$

where y refers to predicted or response variable. The range of predictors or controlled variables starts from x_1 to x_k . β_0 is called the intercept or constant variable and β_1 to β_k are the regression coefficients of controlled variables. ε is fitted or residual error to indicate the uncertainty in the model [41]. We normalized the dataset before fitting the MLR model on the respective dataset to monitor the population growth of ACP in relation to host plant phenology and different abiotic factors.

2.6. Feature Importance. To measure the importance of different input variables for predicting different ACP-phenological stages, we used a feature importance graph using RFR. The feature importance graphs for eggs and nymphs reveal that "branches with flush" is one of the most important variables for ACP egg and nymph growth (Figure 2).

3. Results

3.1. Effect of Abiotic Factors on Population Fluctuations of ACP. To study the impact of various abiotic factors on the population phenology of *D. citri* during the experimental time duration of 25 months on an individual and cumulative basis for different citrus species, correlation coefficient values by using Minitab software were calculated (Table 1). In the case of the ACP-eggs' population, host plant flush growth patterns and average weekly relative humidity have a significant and positive relationship with ACP-eggs' production and growth as $R=0.44$ and 0.247 and $p \leq 0.05$). ACP-nymphs' growth was found to be positively correlated with input variable branches with flush as $R=0.48$ and

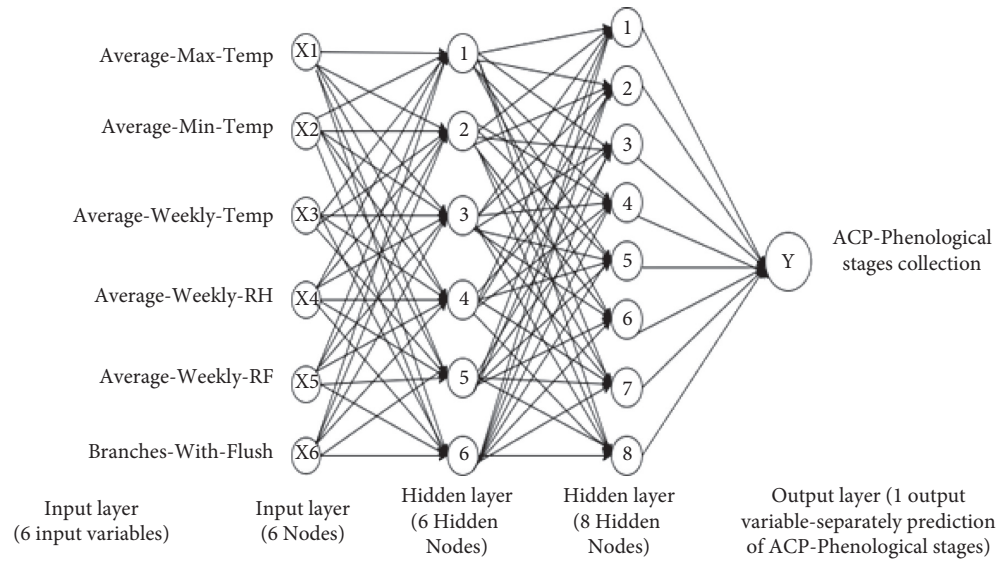


FIGURE 1: Deep neural network architecture for predicting ACP population dynamics.

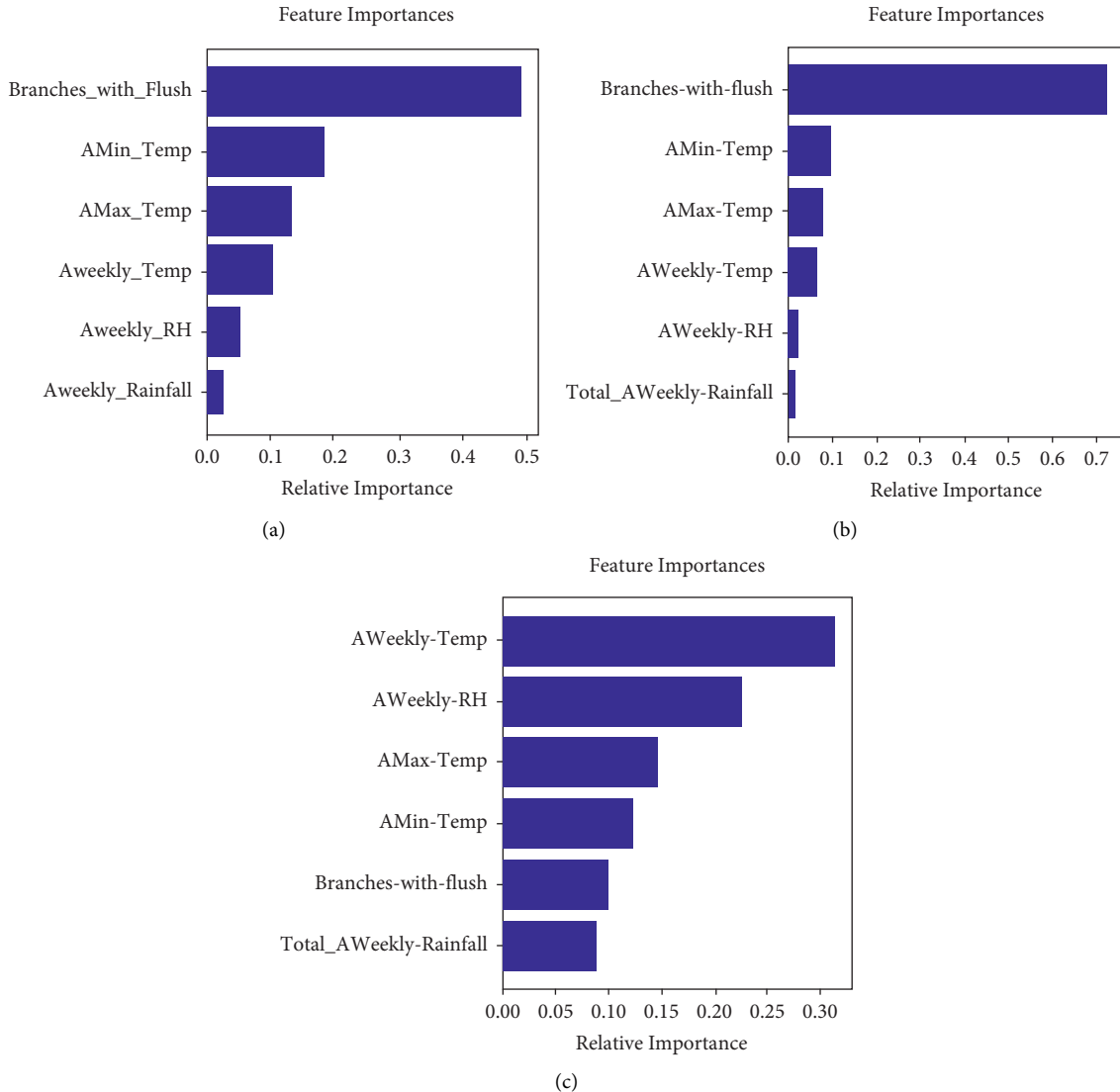


FIGURE 2: Feature importance graphs for different ACP phenological stages. (a) Input variables importance for ACP-egg's growth. (b) Input variables importance for ACP-nymph's growth. (c) Input variables importance for ACP-adult's growth.

TABLE 1: Pearson's correlation coefficient values between different ACP phenological stages and different abiotic variables and host plant flush growth patterns.

Predictor variables	Phonological stage	Correlation coefficient (R)	Sig.
Average Max Temp	Eggs	−0.052	0.663
Average Min Temp		0.011	0.929
Average Weekly Temp		−0.019	0.871
Average Weekly RH		0.247	0.035**
Average Weekly Rainfall		0.121	0.308
Branches with flush		0.44	0.000**
Average Max Temp	Nymphs	0.065	0.583
Average Min Temp		0.043	0.715
Average Weekly Temp		0.053	0.654
Average Weekly RH		−0.028	0.817
Average Weekly Rainfall		−0.036	0.762
Branches with flush		0.48	0.000**
Average Max Temp	Adults	0.233	0.047**
Average Min Temp		0.25	0.033**
Average Weekly Temp		0.244	0.038**
Average Weekly RH		−0.013	0.914
Average Weekly Rainfall		0.051	0.667
Branches with flush		−0.053	0.655

$p \leq 0.05$. Average weekly rainfall and relative humidity were nonsignificant and negatively correlated with ACP-nymphs' abundance.

It is clear from the results that average minimum temperature, average maximum temperature, and average weekly temperature showed positive and significant impact on the population of ACP-adults ($R = 0.233, 0.25$, and 0.244 and $p \leq 0.05$). Cumulatively, from the time course of March 2011 to April 2013, rainfall and relative humidity exerted a significant but negative impact on the ACP population. Meanwhile, all three temperatures exerted a positive but not nonsignificant correlation with population of ACP (Table 1).

3.2. Comparison of Different Regression Approaches to Predict ACP-Eggs' Population. To predict ACP-eggs' population, all three employed models were fitted using training data. We experimented with these regression-based approaches on eight types of datasets (Figure 3). Figure 3 shows a comparison of actual and predicted values. We have categorized models from best to worst in the context of their performance in ACP-eggs' population. In most cases, the DNN model resulted in the least RMSE value of 0.63925 while predicting the ACP-eggs' population. The RMSE value was computed by taking the mean of RMSE values obtained from results of eight datasets. The RFR model was the other best regression approach which resulted in the second least RMSE value of 0.70375. RFR is an ensemble method which is much efficient in extracting meaningful information from the given data. It was found to be true in previous studies [18, 35]. The MLR model resulted in the RMSE value of 0.7935 as it could not perform well in comparison with other approaches deployed for ACP-eggs' population prediction. These findings are consistent with result of [19, 42, 43].

In the case of ACP-nymphs' population prediction, the DNN model performed better when compared with the

other competitive approaches (Figure 4). DNN has the least RMSE value. Before the training of a prediction model, hyperparameter tuning was performed to attain the best parameters for each model. Then, models were retrained using these best parameters to obtain minimum the loss function's values. The residuals calculated by DNN, RFR, and MLR models were 1.1875, 1.38775, and 1.2715, respectively (Table 2).

ACP-adult stage considers to be more threatening for all ornamental and citrus plants. Timely identification and removal from citrus cultivars is a matter of great interest for citrus growers. While predicting the ACP-adults' population in relation to different abiotic variables (Table 1) and flush growth patterns, the DNN model resulted in an RMSE value of 3.6776 which was the least RMSE value as compared to residuals computed by RFR and MLR models (Figure 5). RFR and MLR models resulted in 6.0553 and 8.6883 residuals while predicting ACP-adults' population's fluctuating trends, respectively (Table 2).

4. Discussion and Conclusion

Pest's population prediction can be used as a tool for area-wide integrated pest management programs as it will help to reduce the applications of agrochemicals in fields [27]. Different abiotic factors can be used as independent variables for building a pest's population prediction model [44]. Along with abiotic factors, there are also some biotic factors that can be used for predicting pest population abundance, e.g., host plant phenology [27, 45]. It was found that, during seasons of abundant flush growth, more infestation of ACP-eggs and ACP-nymphs were observed in citrus orchards, and the same effects were observed in [1, 8, 46]. Proper pest management strategies will help to conserve psyllid-natural enemies by minimizing pesticides applications in fields so that they can play their role as a biocontrol agent against ACP effectively. Optimum climatic conditions and host plant phenological

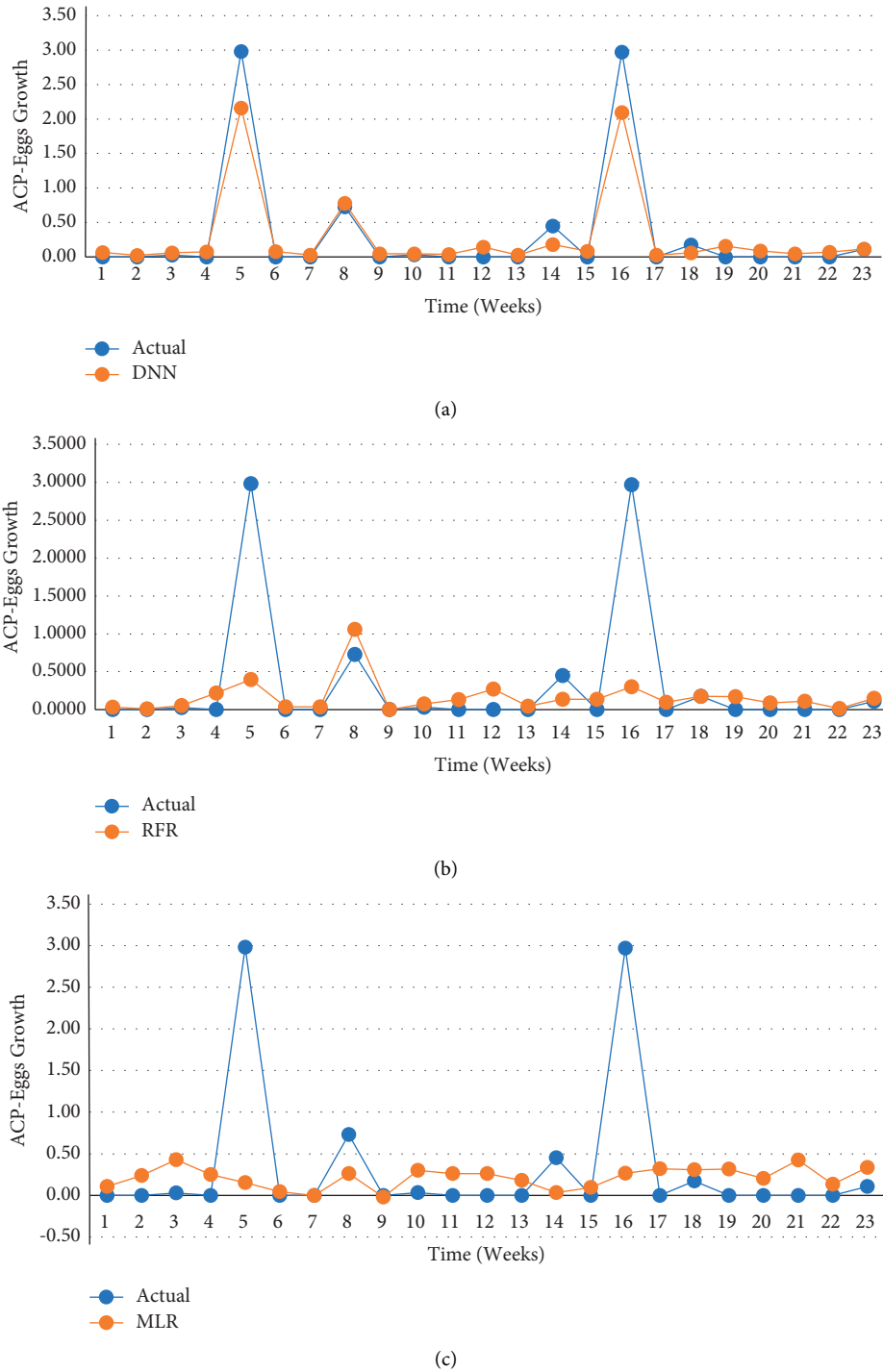


FIGURE 3: A comparison of actual and predicted ACP-eggs' values using different models. (a) DNN, (b) RFR, and (c) MLR comparison of different regression approaches to predict ACP-nymphs' population.

patterns have a great impact on ACP's survival, biology, and resultant pest abundance. The findings of this work are consistent with studies of [4, 5, 47]. ACP's population was found to decrease significantly with rainfall and relative humidity and increased with temperature. The ACP-adults' population was seen at its peak from March to April and September to October where maximum ACP-adults' population was observed in former study duration.

DNN is an appropriate choice for modeling the ACP population dynamics prediction problems as it has the potential to model complex data [35]. DL, an emerging and powerful evolution in ML, can become a powerful tool for ecologists because of its quantitative and predictive nature [48–50]. Because of generalizability of DL algorithms, they are competent models for prediction problems specifically in ecology and generally in all domains of research related to

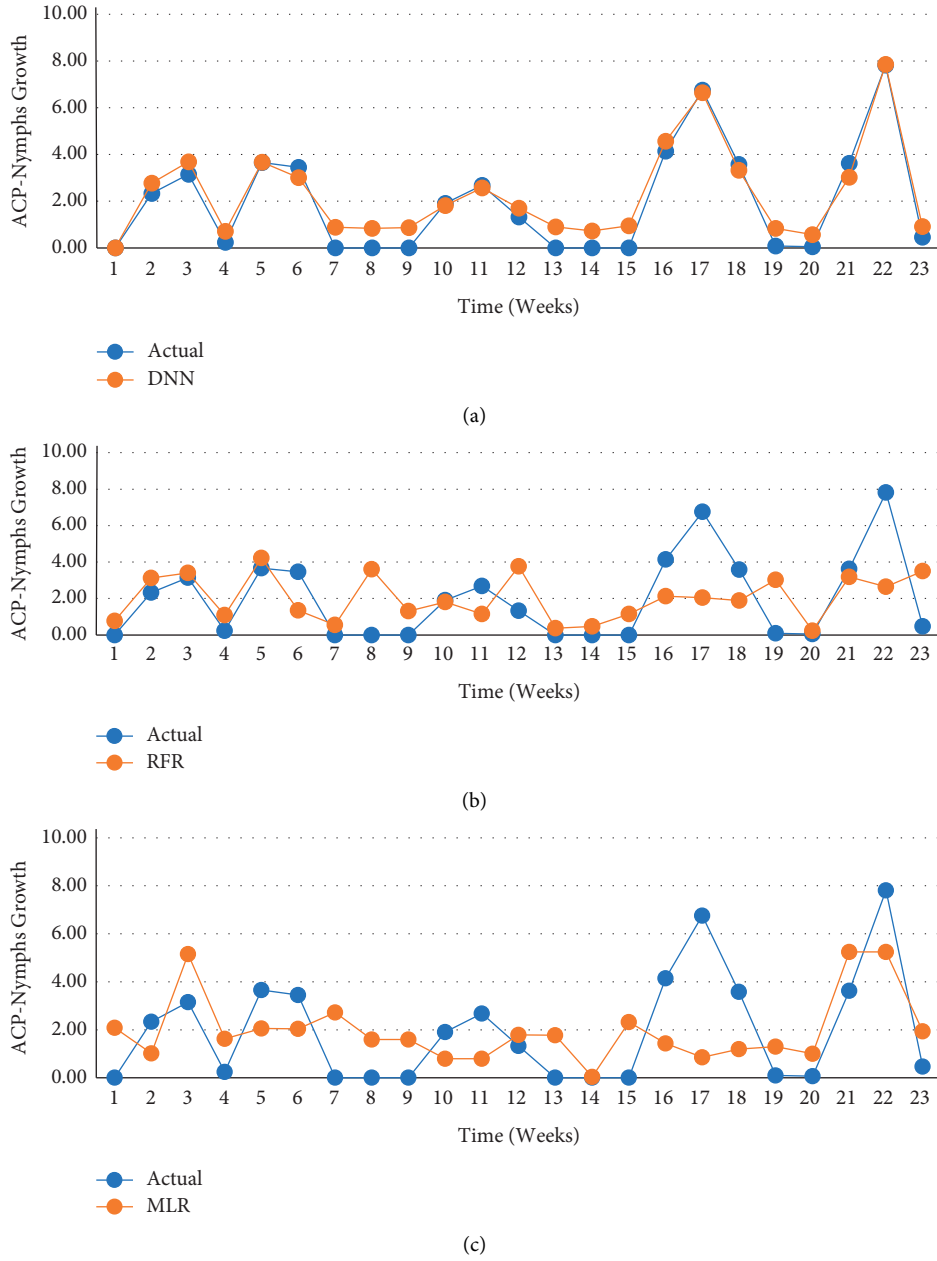


FIGURE 4: A comparison of actual and predicted ACP-nymphs' values using different models. (a) DNN, (b) RFR, and (c) MLR comparison of different regression approaches to predict ACP-adults' population.

TABLE 2: A comparison of RMSE values computed by different regression techniques for ACP-egg, ACP-nymph, and ACP-adult populations.

Models	RMSE-eggs	RMSE-nymphs	RMSE-adults
RFR	0.7037	1.3877	6.0553
MLR	0.7935	1.2715	8.6883
DNN	0.63925	1.1875	3.6776

forecasting problems. It was concluded here that DNN outperformed other classical and statistical regression techniques while modeling ACP population fluctuating trends, and we can deploy it in the future for modeling complex forecasting problems. The RFR techniques also

performed better than the statistical model MLR while predicting ACP-egg and ACP-adult population. According to [18], RFR can be more robust for prediction by acting upon some factors such as proper adjustment of hyper-parameter values and larger datasets.

In this study, various regression-based models, ranging from classical regression to deep learning-based regression, were employed for predicting the population dynamics of ACP. Current study compared predicting of the performance of various models by comparing and evaluating their resultant RMSE values. Different input variables, i.e., Average Max Temp, Average Min Temp, Average Weekly Temp, Average Weekly RH, Average Weekly Rainfall, and

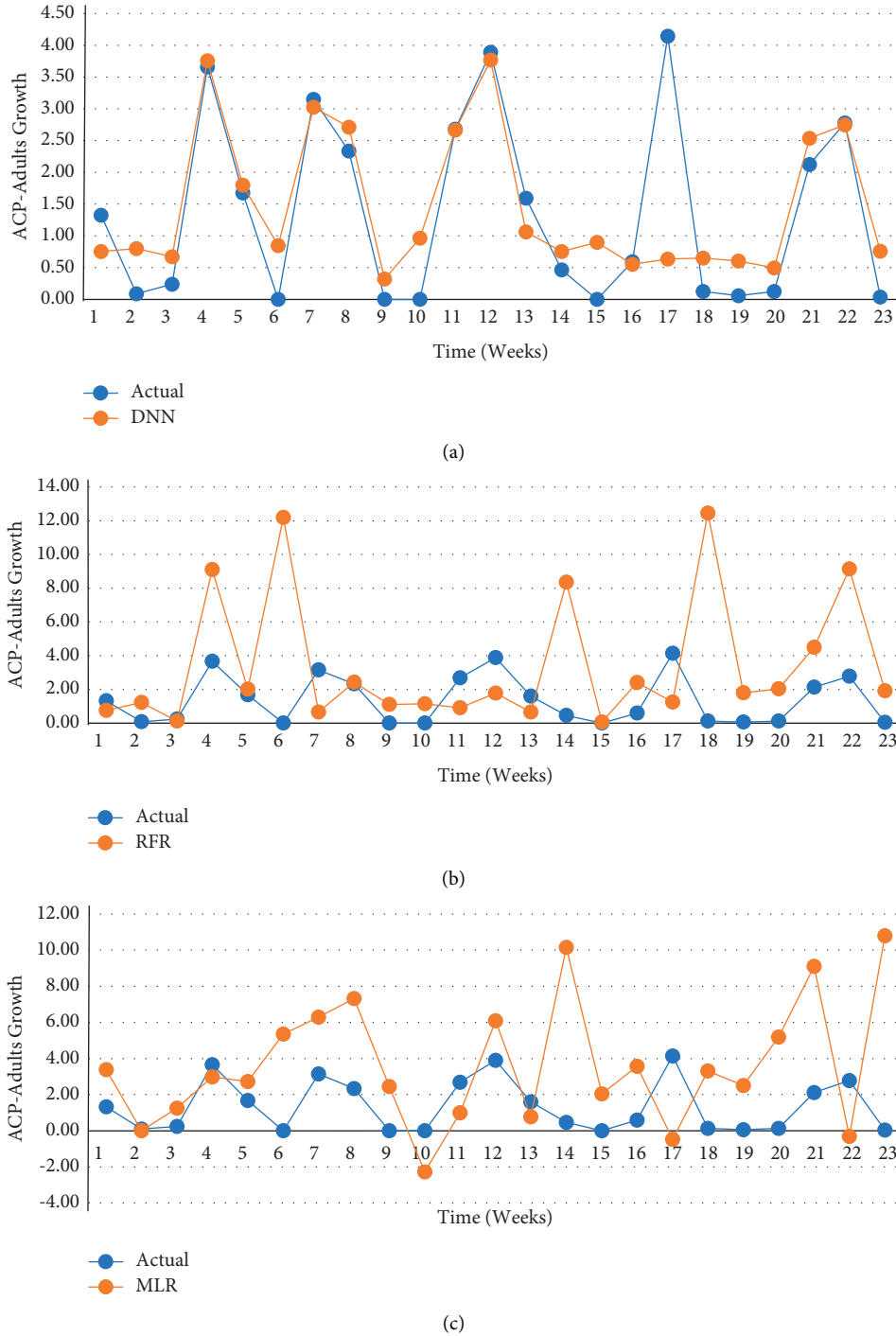


FIGURE 5: A comparison of actual and predicted ACP-adults' values computed by different regression techniques. (a) DNN. (b) RFR. (c) MLR.

Branches with flush, were used in present regression-based models. The key findings of this research can be summarized as follows. (1) The DNN model with differently tuned hyperparameters (Input, hidden, output layers, activation functions, and optimizer) is best suited for predicting population phenology of ACP. (2) A comparison of RMSE values computed by different regression-based models depicted that the DNN-based model has the potential to

model time-series forecasting problems. (3) The RFR model was another effective regression-based model and a good choice for predicting ACP-population dynamics as it resulted in the second least RMSE values for different ACP-phenological stages' population prediction. (4) For reliable predictions and optimization of different regression-based models, configurations are also crucial. (5) The model which resulted in the smallest mean-RMSE value for the

corresponding ACP-phenological stage was considered as the best prediction model.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

There are no conflicts of interest regarding the publication of this paper.

References

- [1] D. G. Hall, M. L. Richardson, E.-D. Ammar, and S. E. Halbert, "Asian citrus psyllid, *Diaphorina citri*, vector of citrus huanglongbing disease," *Entomologia Experimentalis et Applicata*, vol. 146, no. 2, pp. 207–223, 2013.
- [2] J. M. Bove, "Huanglongbing: a destructive, newly emerging, century-old disease of citrus," *Journal of Plant Pathology*, vol. 88, pp. 7–37, 2006.
- [3] L. B. Kumagai, C. S. LeVesque, C. L. Blomquist et al., "First report of *Candidatus Liberibacter asiaticus* associated with citrus huanglongbing in California," *Plant Disease*, vol. 97, no. 2, p. 283, 2013.
- [4] F. L. Russell and S. M. Louda, "Phenological synchrony affects interaction strength of an exotic weevil with *Platte* thistle, a native host plant," *Oecologia*, vol. 139, no. 4, pp. 525–534, 2004.
- [5] S. Mopper, "Adaptive genetic structure in phytophagous insect populations," *Trends in Ecology & Evolution*, vol. 11, no. 6, pp. 235–238, 1996.
- [6] Y. H. Liu and J. H. Tsai, "Effects of temperature on biology and life table parameters of the Asian citrus psyllid, *Diaphorina citri* Kuwayama (Homoptera: psyllidae)," *Annals of Applied Biology*, vol. 137, no. 3, pp. 201–206, 2000.
- [7] J. R. Stedinger, C. A. Shoemaker, and R. F. Tenga, "A stochastic model of insect phenology for a population with spatially variable development rates," *Biometrics*, vol. 41, no. 3, pp. 691–701, 1985.
- [8] M. S. Hoddle, C. D. Hoddle, S. V. Triapitsyn, S. Z. Khan, and M. J. Arif, "How many primary parasitoid species attack nymphs of *Diaphorina citri* (Hemiptera: liviidae) in Punjab, Pakistan?" *Florida Entomologist*, vol. 97, no. 4, pp. 1825–1828, 2014.
- [9] Q. Zeb, I. Khan, M. Inayatullah, P. Agriculture, Y. Hayat, and S. Ahmad-ur-Rehman, "Population dynamics of citrus whiteflies, aphids, citrus psylla leaf miner and their biocontrol agents in Khyber Pakhtunkhwa," *Sarhad Journal of Agriculture*, vol. 27, no. 3, pp. 451–457, 2011.
- [10] A. Chlingaryan, S. Sukkari, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review," *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018.
- [11] J. D. Olden, J. J. Lawler, and N. L. Poff, "Machine learning methods without tears: a primer for ecologists," *The Quarterly Review of Biology*, vol. 83, no. 2, pp. 171–193, 2008.
- [12] M. A. Nielsen, *Neural Networks and Deep Learning*, Vol. 2018, Determination press, San Francisco, CA, USA, 2015.
- [13] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
- [14] L.-n. Yang, L. Peng, L.-m. Zhang, L.-l. Zhang, and S.-s. Yang, "A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on back propagation artificial neural network and principal components analysis," *Computers and Electronics in Agriculture*, vol. 68, no. 2, pp. 200–206, 2009.
- [15] C.-N. Lin, M.-Y. Wei, N.-T. Chang, and Y.-Y. Chuang, "The occurrence of *Scirtothrips dorsalis* Hood in mango orchards and factors influencing its population dynamics in Taiwan," *Journal of Asia-Pacific Entomology*, vol. 18, no. 3, pp. 361–367, 2015.
- [16] Y. Yan, C.-C. Feng, M. P.-H. Wan, and K. T. T. Chang, "Multiple regression and artificial neural network for the prediction of crop pest risks," in *Proceedings of the International Conference on Information Systems for Crisis Response and Management in Mediterranean Countries*, pp. 73–84, Springer, Tunisia, May 2015.
- [17] W. Y. Zhang, T. Z. Jing, and S. C. Yan, "Studies on prediction models of *dendrolimus superans* occurrence area based on machine learning," *Journal of Beijing Forestry University*, vol. 39, no. 1, pp. 85–93, 2017.
- [18] S. Skawsang, M. Nagai, N. K. Tripathi, and P. Soni, "Predicting rice pest population occurrence with satellite-derived crop phenology, ground meteorological observation, and machine learning: a case study for the central plain of Thailand," *Applied Sciences*, vol. 9, no. 22, p. 4846, 2019.
- [19] S. Chen, A. Whiteman, A. Li et al., "An operational machine learning approach to predict mosquito abundance based on socioeconomic and landscape patterns," *Landscape Ecology*, vol. 34, no. 6, pp. 1295–1311, 2019.
- [20] T. Nyabako, B. M. Mvumi, T. Stathers, S. Mlambo, and M. Mubayiwa, "Predicting *Prostephanus truncatus* (Horn) (Coleoptera: bostrichidae) populations and associated grain damage in smallholder farmers' maize stores: a machine learning approach," *Journal of Stored Products Research*, vol. 87, pp. 101–592, 2020.
- [21] J. Patil and V. D. Mytri, "A prediction model for population dynamics of cotton pest (*Thrips tabaci* Linde) using multi-layer-perceptron neural network," *International Journal of Computer Application*, vol. 67, no. 4, pp. 19–26, 2013.
- [22] M. J. Watts and S. P. Worner, "Using artificial neural networks to determine the relative contribution of abiotic factors influencing the establishment of insect pest species," *Ecological Informatics*, vol. 3, no. 1, pp. 64–74, 2008.
- [23] T. M. Carvajal, K. M. Viacrusis, L. F. T. Hernandez, H. T. Ho, D. M. Amalin, and K. Watanabe, "Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines," *BMC Infectious Diseases*, vol. 18, no. 1, pp. 1–15, 2018.
- [24] D. S. Lee, Y. S. Bae, B. K. Byun, and S. Lee, "Occurrence prediction of the citrus flatid planthopper (*Metcalfa pruinosa* (Say, 1830)) in South Korea using a random forest model," *Forests*, vol. 10, no. 7, p. 583, 2019.
- [25] İ. Balaban, F. Acun, O. Y. Arpalı, and F. Murat, "Development of a forecasting and warning system on the ecological life-cycle of sunn pest," 2019, <https://arxiv.org/abs/1905.01640>.
- [26] D. S. Yadav, S. Chander, and K. Selvaraj, "Agro-ecological zoning of brown planthopper [*Nilaparvata lugens* (Stal)] incidence on rice (*Oryza sativa* L.)," *Journal of Scientific & Industrial Research*, vol. 69, pp. 818–822, 2010.
- [27] P. K. Jayanthi and V. A. Erghese, "Host-plant phenology and weather-based forecasting models for population prediction

- of the oriental fruit fly, *Bactrocera dorsalis* Hendel,” *Crop Protection*, vol. 30, no. 12, pp. 1557–1562, 2011.
- [28] K. Fang, C. Shen, D. Kifer, and X. Yang, “Prolongation of SMAP to spatiotemporally seamless coverage of continental US using a deep learning neural network,” *Geophysical Research Letters*, vol. 44, no. 21, pp. 11–30, 2017.
- [29] O. Adeyemi, I. Grove, S. Peets, Y. Domun, and T. Norton, “Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling,” *Sensors*, vol. 18, no. 10, p. 3408, 2018.
- [30] S. Scher, “Toward data-driven weather and climate forecasting: approximating a simple general circulation model with deep learning,” *Geophysical Research Letters*, vol. 45, no. 22, pp. 12–616, 2018.
- [31] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, “Deep learning classification of land cover and crop types using remote sensing data,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 778–782, 2017.
- [32] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, “Machine learning in geosciences and remote sensing,” *Geoscience Frontiers*, vol. 7, no. 1, pp. 3–10, 2016.
- [33] G. L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, “Deep learning for plant identification using vein morphological patterns,” *Computers and Electronics in Agriculture*, vol. 127, pp. 418–424, 2016.
- [34] A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, “Machine learning for high-throughput stress phenotyping in plants,” *Trends in plant science*, vol. 21, no. 2, pp. 110–124, 2016.
- [35] W. Rammer and R. Seidl, “Harnessing deep learning in ecology: an example predicting bark beetle outbreaks,” *Frontiers of Plant Science*, vol. 10, p. 1327, 2019.
- [36] L. Breiman, “Randomizing outputs to increase prediction accuracy,” *Machine Learning*, vol. 40, no. 3, pp. 229–242, 2000.
- [37] S. Nawar and A. M. Mouazen, “Comparison between random forests, artificial neural networks and gradient boosted machines methods of on-line Vis-NIR spectroscopy measurements of soil total nitrogen and total carbon,” *Sensors*, vol. 17, no. 10, p. 2428, 2017.
- [38] K. Arun and C. J. Langmead, “Structure based chemical shift prediction using Random Forests non-linear regression,” in *Proceedings of the 4th Asia-Pacific Bioinformatics Conference*, pp. 317–326, Taipei, Taiwan, February 2006.
- [39] M. Meng and C. Song, “Daily photovoltaic power generation forecasting model based on random forest algorithm for north China in winter,” *Sustainability*, vol. 12, no. 6, p. 2247, 2020.
- [40] D. Rumelhart, G. Hinton, and R. Williams, “Learning representations by back-propagation errors,” *Nature*, vol. 323, pp. 533–536, 1986.
- [41] I. M. M. Ghani and S. Ahmad, “Stepwise multiple regression method to forecast fish landing,” *Procedia-Social and Behavioral Sciences*, vol. 8, pp. 549–554, 2010.
- [42] S. Z. Khan, M. J. Arif, C. D. Hoddle, and M. S. Hoddle, “Phenology of Asian citrus psyllid (Hemiptera: liviidae) and associated parasitoids on two species of Citrus, kinnow Mandarin and sweet orange, in Punjab Pakistan,” *Environmental Entomology*, vol. 43, no. 5, pp. 1145–1156, 2014.
- [43] L. E. de Oliveira Aparecido, G. de Souza Rolim, J. R. D. S. C. De Moraes, C. T. S. Costa, and P. S. D. Souza, “Machine learning algorithms for forecasting the incidence of Coffea arabica pests and diseases,” *International Journal of Biometeorology*, vol. 64, pp. 1–18, 2020.
- [44] M. K. Sharma, A. Atsedewoin, and S. Fanta, “Forewarning models of the insects of paddy crop,” *International Journal of Biodiversity and Conservation*, vol. 3, no. 8, pp. 367–375, 2011.
- [45] G. Hu, X. N. Cheng, G. J. Qi et al., “Rice planting systems, global warming and outbreaks of Nilaparvata lugens (Stål),” *Bulletin of Entomological Research*, vol. 101, no. 2, pp. 187–199, 2011.
- [46] H. A. Narouei-Khandan, S. E. Halbert, S. P. Worner, and A. H. van Bruggen, “Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling approaches, with emphasis on the USA,” *European Journal of Plant Pathology*, vol. 144, no. 3, pp. 655–670, 2016.
- [47] M. Fiaz, M. Afzal, and M. Z. Majeed, “Influence of abiotic weather factors on population dynamics of asian citrus psyllid, diaphoria citri Kuwayama (Hemiptera: psyllidae) in central Punjab,” *Pakistan Journal of Agricultural Research*, vol. 56, no. 1, pp. 35–40, 2018.
- [48] M. R. Evans, K. J. Norris, and T. G. Benton, “Predictive ecology: systems approaches,” *Philosophical Transactions of the Royal Society of London B Biological Sciences*, vol. 367, pp. 163–169, 2012.
- [49] J. S. Clark, S. R. Carpenter, and M. Barber, S. Collins, A. Dobson, J. A. Foley et al., *Ecological Forecasts: An Emerging Imperative*. Science, Beck-Johns, Cumbria, England GB, 2001.
- [50] M. Reichstein, G. Camps-Valls, B. Stevens, J. Denzler, and N. Carvalhais Prabhat, “Deep learning and process understanding for data-driven Earth system science,” *Nature* vo.vol. 566, pp. 195–204, 2019.