Webology (ISSN: 1735-188X)
Volume 19, Number 2, 2022

Machine Learning Time Series Models For Tea Pest Helopeltis Infestation In India

Dwijendra Nath Dwivedi[a,*], Pradish Kapur[b], Nipun Kapur[c]

[a] Krakow university of Economics, Rakowicka 27, Kraków, 31-510, Poland

[b] PES University Electronic City, Hosur Road, Bangalore – 560100, India

[c] Indian Institute of Technology Roorkee, Roorkee - Haridwar Highway, 247667, Uttarakhand, India

Abstract

Understanding pests and their life cycle are a complex phenomenon due to their nonlinear relation with environmental factors and the interdependence of the environmental factors themselves. We focus our attention on ML (Machine Learning) models due to their ability to simulate non-linear phenomena effectively. We use Stacked Models, that consist of a neural network model (NN) and a time series model (TS) to simulate the life cycle of pests which varies in duration with the seasons. Several Machine Learning models were developed for predicting Helopeltis (tea pest, though impacts several other crops as well) infestation. The idea was not to depend heavily on the good scenario data, but rather which is readily available even for small holders and planters who are not too much reliant on technology to capture the data on pest infestation. Evaluating the various developed models shows that Neural Network models with better accuracies can be developed with real world data from the field. Thus, being suitable for tea gardens without too much reliance on technologies and extensive data capturing processes. This has been possible through special treatment of available data as well as ability of stacked models to provide good results in the solution space with a generally smaller set of data having lesser number of parameters.

Keywords: Machine Learning, Pesticides, Artificial Intelligence, Agriculture, Temperature

INTRODUCTION

Tea industry is facing several challenges world over since last few years (Kramer & Ware, 2021; Duncan et al., 2018; Brouder et al., 2017). These include increased costs of production, climate change and associated aspect of pest infestations. We focus our attention to pest infestation which result in huge production losses up to 55% and in some instances 100% loss as well (Kachhawa and Kumawat, 2018; Ahuja et al., 2013; Samanta and Gosh, 2012). The pest infestation in turn increases per unit production costs as a result of reduced yields and/or pesticide costs and cost of labor involved in spraying of the pesticides (Roy et al., 2020; Khanali et al., 2021). Considering that about 70% of the global tea is produced by small holders with limited means, pest infestations create several challenges for most in the tea industry (Voora et al., 2019). Further, of late several restrictions on the use of pesticides in tea has also decreased total exports by some of the countries traditionally known for tea exports (Gurusubramanian, 2008; Voora et al., 2019).

Towards solving some of the above challenges, lots of efforts are being spent world over on the automation of the tea industry, though with some success only. Mechanized Harvesting to Drone based pesticide sprays. Each of the automation initiatives has some pros and cons. E.g. Mechanical harvesting impacts separate plucking of buds hence segregation into good quality highly prized teas like White tea. Similarly, it has been a challenge to use pesticide sprays using drones due to several factors from costs to effectiveness in removing pests (Delavalpour et al.,2021; Morley et al., 2017). Another aspect of automation results in ethical issue of job loss and unemployment, particularly in countries like Kenya, where not too many job alternatives are available (Mulinya, 2021; Gebre, 2017). Note that tea industry employs 13 Mn people globally (Voora et al., 2019). These contradictory aspects bring back into the focus on increasing profitability in the business, be it through segregation of high end teas to environmentally friendly ways in operations. Ultimately, it would be a combination of cost (mostly mechanization driven) and profitability benefiting the tea industry.

Literature Study

Attempts have been made in past to use pest forecasting to understanding nature of infestation and hence more efficient pest control with lower costs and lower pesticide content in tea (Kallor et al., 2020; Azrag et al., 2018). Many of the early forecasts were simple correlation studies or regression models (Ahmed et al., 2012). Over the years with advancement of new age technologies people started realizing that methods like Neural Network would be better suited to model nonlinear phenomena of pest infestations (Zhang et al., 2017). This because in nature most of the phenomena are highly nonlinear and have a complex interplay of various factors. They (Zhang et al., 2017) also affirm that Support Vector Machine Model for accurate and effective predictions of insect pest occurrence areas as a reliable predictive tool.

On the other hand, recently Xiao et al., (2018) used LSTM models for modelling pest infestation in Cotton. They used data of eight different parameters to predict occurrence of pests and diseases. The weather features consist of Maximum Temperature (Max T (°C)), Minimum Temperature (Min T (°C)), Relative Humidity in the morning (RH1 (%)), Relative Humidity in the evening (RH2 (%)), Rainfall (RF (mm)), Wind Speed (WS (comp)), Sunshine Hour (SSH (hrs)) and Evaporation (EVP (mm)). They found out that the experimental results showed that LSTM performed better on the prediction of occurrence of pests and diseases in cotton fields, and yielded an Area Under the Curve (AUC) of 0.97. Though as described above almost 70% of global tea production comes from the small tea planter. These small players do not have technologies to capture data for several of the parameters as mentioned above.

Infact there is a big technological gap in adoption of scientific practices of tea cultivation by small tea growers (Parasar et al., 2020)

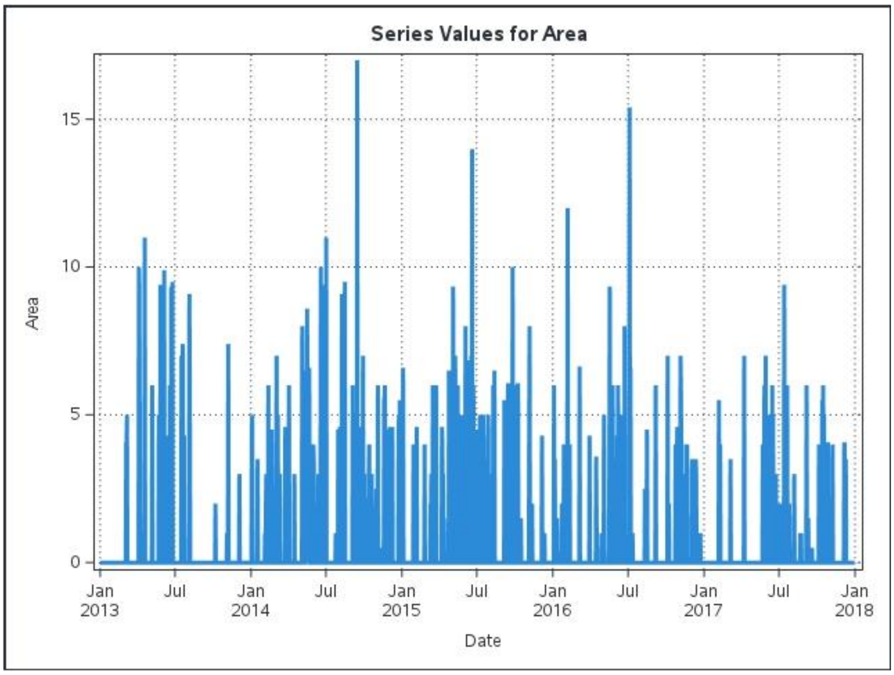
Considering above aspects we decided to use different Neural Network models towards forecasting of pests in the tea industry. More specifically, we strive to evaluate various neural network models specifically for Helopeltis infestation. Also, in spite of attractiveness and ability of using neural nets in the tea industry for developing complex nonlinear models challenge is faced due to unavailability of granular data for many of the weather parameters especially for small tea growers. E.g. Calculating exact intensity of pest infestation pest traps at various places in the tea garden is required. Then the sample taken are extrapolated for entire tea garden. Similarly, small growers typically don't have the resources to measure parameters like Evaporation or Wind Speed etc. We decide to use just two parameters daily temperatures (maximum and minimum) and rainfall for modelling infested area by Helopeltis. Several tools are available in the market with their own pros and cons. We decided to use SAS Deep Learning Python (DLP y) and stacked model for the same.

Data Source:

The data are from Teok Tea gardens of Assam in India covering around 250 hectares of tea gardens. It consists of diurnal data of rainfall, minimum and maximum temperatures and estimated value of the area infested with pests. The duration of the data is from 2013 through 2017.

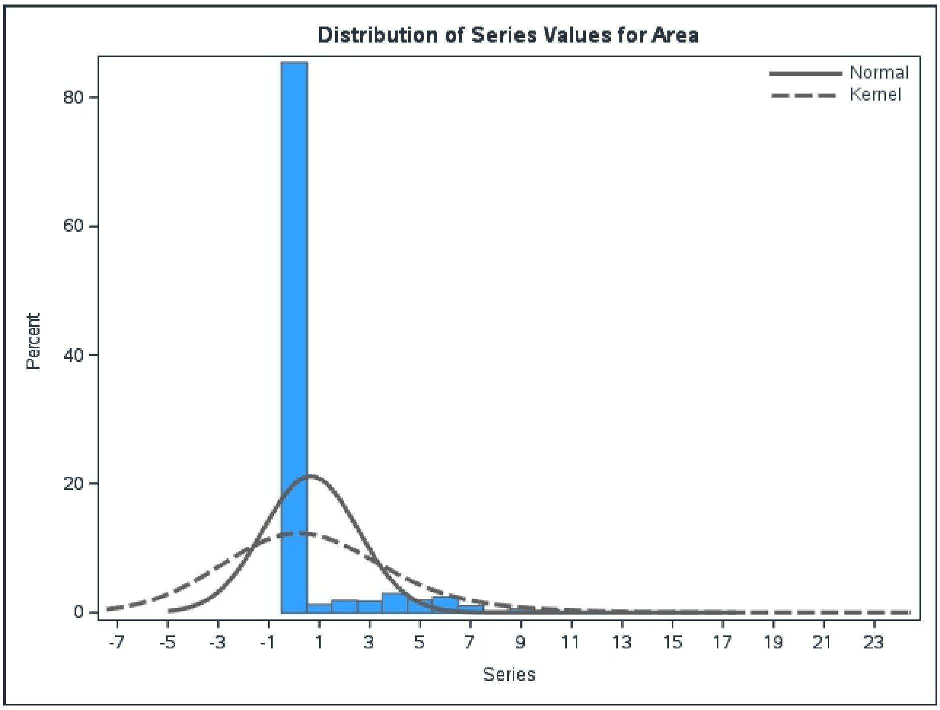
Data Exploration:

Pest control data has a unique random behavior as this data is not continuous in time series. Pest control is performed only certain number of days and that is reflected in the data as per below chart:



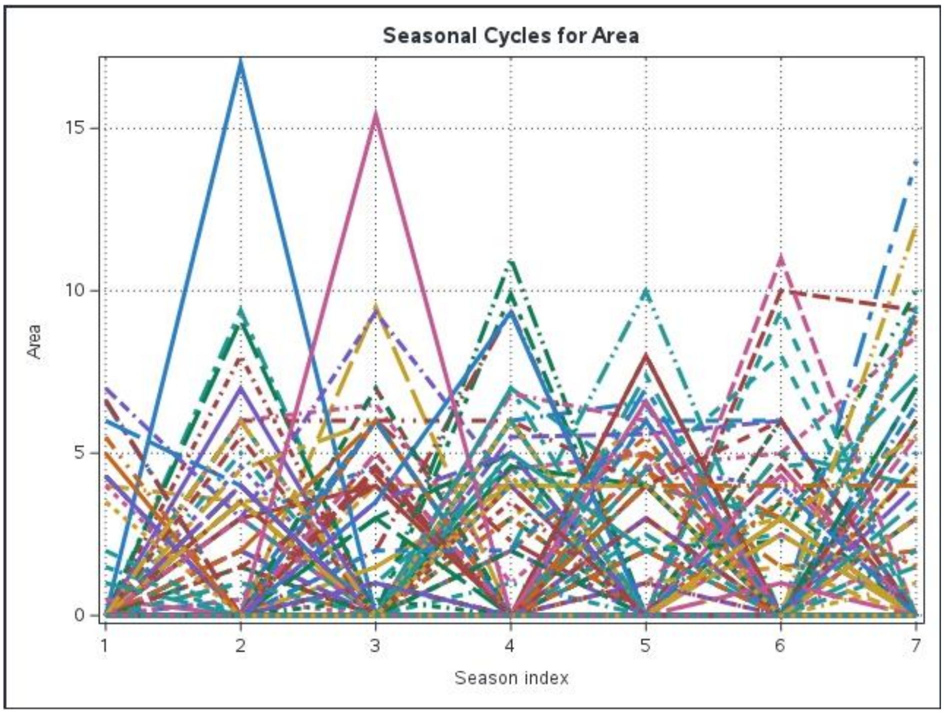
Graph 1: data for pest control (area) over time series (source: Author)

The distribution of the data is a bit not so continuous as most frequent is the day when there is no pest control performed in the tea gardens.



Graph 2: data for pest control (area) over time series (source: Author)

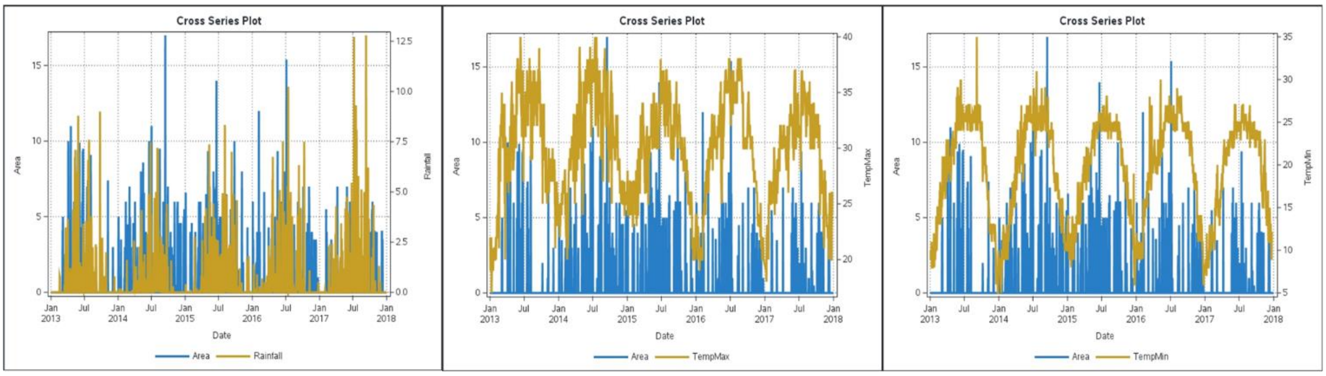
There is hardly any seasonality in the terms of days of the week and that reflects from the data.



Graph 3: Seasonal index for the days of the week (source: Author)

Cross Series Plots:

We also tried to understand the behavior of the input datasets that are maximum temperature, Minimum Temperature and rainfall.



Grap4: Cross Series plots with independent variables (Minimum Temperature, Maximum temperature and rainfall) (source: Author)

Models and Results:

SARIMA Analysis: We employ a SARIMA design to the time series data. We choose the multiplied seasonal best fit model using the Akaike Information Criterion to select the most preferred model with the least value. The SARIMA is an extension of ARIMA that explicitly models the seasonal element in univariate data. A seasonal ARIMA model is an extension of ARIMA and is formed by including additional seasonal terms in the ARIMA models. The seasonal part of the model consists of terms that are like the non-seasonal components of the model, but involve back shifts of the seasonal period.

Stacked Model: The stacked modeling strategy generates forecasts using stacked models that consist of a neural network model (NN) and a time series model (TS). The time series model applies to the residuals of the neural network forecasts. This modeling strategy captures the nonlinear relationship between the dependent and independent variables as well as time series characteristics in the data, such as seasonality and trend.

Model Comparison uses weighted MAPE: The weighted version of a statistic of fit is calculated using the input data to assign weights to each series, with higher weights being assigned to series with higher volume.

$$\text{Series weight} = \left|\frac{1}{T}\Sigma_{t=1}^{T} y_t\right|$$

Weighted mean absolute percent error (WMAPE)

The weighted mean of the MAPE values for each time series, $\frac{\Sigma \text{MAPE} \times \text{SeriesWeight}}{\Sigma \text{SeriesWeight}}$

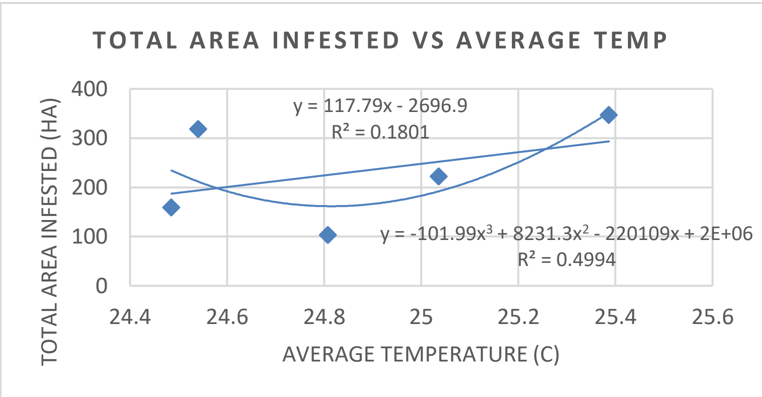
Model Name	Status	Weighted Mean Absolute Error
Stacked Model (NN + TS)	Successful	0.8512

Panel Series Neural Network	Successful	0.9454
Seasonal Model	Successful	1.0558
Temporal Aggregation Model	Successful	1.0843
RNN Forecasting	Successful	1.0984
Auto-forecasting	Successful	3.7881

Table 1: Model comparison using Weight MAPE

Discussions and Conclusion:

A comparative analyses of the various Neural Net models towards simulating area infested by Helopeltis in tea gardens highlight that Stacked Models perform better than most of the other models. The Stacked modeling strategy captures the nonlinear relationship between the area infested by pests and independent variables of temperature as well as time series characteristics in the data, such as seasonality (season dependent growth of pest Helopeltis). This feature of modeling the season dependent variation of growth of Helopeltis (Kalloor, et.al., 2020) as well as the nonlinear relation between temperatures and area infested is in line with the observed real phenomena (Figure below). Thus, giving better results.



Graph 5: Nonlinear relationship (source: Author)

Acknowledgement:

Acknowledge Mr. Jagjeet Kandhal and Mr. Kaiylanjeet Bohra for thankfully sharing the data. SAS software has been used for the analysis and modeling.

References:

[1] A. Bondori, A. Bagheri, M. Sookhtanlou, and C. A. Damalas, "Modeling farmers ' intention for safe pesticide use: the role of risk perception and use of information sources," pp. 66677–66686, 2021.

[2] C. A. Damalas, "Farmers ' intention to reduce pesticide use: the role of perceived risk of loss in the model of the planned behavior theory," pp. 35278–35285, 2021.

[3] A. B. Tambe, B. M. R. Mbanga, D. L. Nzefa, and M. G. Name, “Pesticide usage and occupational hazards among farmers working in small-scale tomato farms in Cameroon,” pp. 0–6, 2019.

[4] G. K. N. Gallia and C. Kephaliacos, “Ecological - economic modeling of pollination complexity and pesticide use in agricultural crops,” J. Bioeconomics, vol. 23, no. 3, pp. 297–323, 2021, doi: 10.1007/s10818-021-09317-9.

[5] S. W. Bengal, T. Dutta, and C. Nayak, “S-Transferase Enzyme Activities and Their Correlation with Genotypic Variations Based on GST M1 and GST T1 Loci in Long Term-Pesticide-Exposed Tea Garden Workers of,” 2019, doi: 10.1007/s13530-019-0389-1.

[6] S. Xie et al., “Does a dual reduction in chemical fertilizer and pesticides, improve nutrient loss and tea yield and quality? A pilot study in a green tea garden in Shaoxing, Zhejiang Province, China,” pp. 2464–2476, 2019.

[7] V. S. Le, D. Lesueur, L. Herrmann, L. Hudek, L. Ngoc, and Q. Lambert, “Sustainable tea production through agroecological management practices in Vietnam: a review,” Environ. Sustain., no. 0123456789, 2021, doi: 10.1007/s42398-021-00182-w.

[8] H. Niwa, “Detection of organic tea farms based on the density of spider webs using aerial photography with an unmanned aerial vehicle (UAV),” Landsc. Ecol. Eng., vol. 17, no. 4, pp. 541–546, 2021, doi: 10.1007/s11355-021-00454-x.

[9] M. Suganthi, S. Event, and P. Senthil kumar, “Comparative bioefficacy of Bacillus and Pseudomonas continues against Helopeltis severe in tea (Camellia sinensis (L.) O . Kuntze,” Physiol. Mol. Biol. Plants, vol. 26, no. 10, pp. 2053–2060, 2020, doi: 10.1007/s12298-020-00875-2.

[10] A. K. Pal, “Bio-control of Pests in Tea: Effect of Environmental,” Int. J. Apple. Comput. Math., vol. 5, no. 3, pp. 1–9, 2019, doi: 10.1007/s40819-019-0666-3.

[11] Z. Qianjiang, “Demand for Pesticides in China to Reach,” 2015.

[12] “Africa, Asia Pacific Canada, China Europe / United Kingdom, India Latin America.”

[13] B. J. Tilley and C. Carr, “New plant benefits need to sell,” no. October, 2012.

[14] P. Mallik and T. Ghosh, “Impact of climate on tea production: a study of the Dooars region in India,” Theor. Apple. Climatol., no. Cline 2008, 2021, doi: 10.1007/s00704-021-03848-x.

[15] M. Ariyani, M. M. Pitoi, T. A. Koesmawati, H. Maulana, and E. S. Endah, “Pyrethroid residues on tropical soil of an Indonesian tea plantation: analytical method development, monitoring, and risk assessment,” vol. 7, 2020.

[16] A. K. Prasad, S. Roy, S. Sen, S. Neave, A. Nagpal, and V. Pandit, “Impact of different pest management practices on natural enemy population in tea plantations of Assam special emphasis on spider fauna,” pp. 629–635, 2020.

[17] R. Sun, W. Yang, Y. Li, and C. Sun, “Multi - residue analytical methods for pesticides in teas: a review,” Eur. Food Res. Technol., vol. 247, no. 8, pp. 1839–1858, 2021, doi: 10.1007/s00217-021-03765-3.

[18] D. Zhang et al., “Detection of systemic pesticide residues in tea products at trace level based on SERS and verified by GC – MS,” pp. 7187–7196, 2019.

[19] A. Heshmati, F. Mehri, and A. M. Khaneghah, “Simultaneous multi-determination of pesticide residues in black tea leaves and infusion: a risk assessment study,” pp. 13725–13735, 2021.

[20] D. K. Soydan, N. Turgut, M. Yalç, and C. Turgut, “Evaluation of pesticide residues in fruits and vegetables from the Aegean region of Turkey and assessment of risk to consumers,” pp. 27511–27519, 2021.

[21] K. Ayouba and S. Vigeant, “Can We Really Use Prices to Control Pesticide Use? Results from a Nonparametric Model,” pp. 885–900, 2020.

[22] H. Machekano, W. Masamba, B. M. Mvumi, and C. Nyamukondiwa, “Cabbage or ‘ pesticide ’ on the platter? Chemical analysis reveals multiple and excessive residues in African vegetable markets,” 2019.

[23] L. Guo, A. Cao, M. Huang, H. Li, and A. Cao, “Effects of haze pollution on pesticide use by rice farmers: fresh evidence from rural areas of China,” pp. 62755–62770, 2021.

[24] M. O. Kwakye, B. Mengistie, and J. Ofosu, “Pesticide registration, distribution and use practices,” Environ. Dev. Sustain., vol. 21, no. 6, pp. 2647–2671, 2019, doi: 10.1007/s10668-018-0154-7.

[25] T. Bonvoisin, L. Utyasheva, D. Knipe, D. Gunnell, and M. Eddleston, “Suicide by pesticide poisoning in India: a review of pesticide regulations and their impact on suicide trends,” pp. 1–16, 2020.

[26] L. A. Pardo et al., “Pesticide exposure and risk of aggressive prostate cancer among private pesticide applicators,” pp. 1–12, 2020.

[27] Reading the tea leaves Climate change and the British cuppa May 2021; Authors: Dr Katherine Kramer Joe Ware; christianaid.org.uk

[28] Seasonal abundance of tea mosquito bug, Helopeltis Antoni Signoret infesting name; Bobby J Kalloor, Dr. M Suganthy, Dr. A Balasubramanian, Dr. P Renukadevi and Dr. M Senthil Kumar; Journal of Entomology and Zoology Studies 2020; 8(6): 2006-2009