Potential Use of Artificial Neural Networks in Prediction of Dengue Abundance in Agra District

Dr. Tatavarty Guru Sant¹*, Dr. Neh Satsangi²
¹Assistant Professor, IBM GLA University, Mathura
²Knowledge Consultant, Finovista
*tguru.sant@gla.ac.in

Abstract. India being a sub-tropical country faces several diseases spread through mosquito vectors, one of which being Dengue. Dengue is classified as mosquito-borne viral infection causing severe illness and mortality in children and adults in Asian subcontinent. There are an estimated 390 million infections annually and to reduce the impact of dengue among population in sub-tropical regions, a novel disease prediction model is required. The aim of this study is to estimate the dengue cases with clinical and environmental variables like rainfall, relative humidity and temperature from 2005-2018 using data available on the India meteorological department for the geographical location of Agra district, Uttar Pradesh, India. Datasets were represented as average minimum and maximum temperature, average annual rainfall in cm, relative humidity and number of reported dengue cases. In this paper, we examined the effectiveness of the artificial neural network (ANN) for dengue prediction for four major sites in Agra (Dayalbagh, Sikandra, Agra cantonment and Kamla Nagar) and prediction model was designed for the period of 2019 and compared with the observed values. The results show variation in the prediction model with respect to clinical and environmental variables corresponding to four different sites. The average error of the prediction model ranged between 27.3% and 84% respectively. Until now several approaches have been discovered for controlling mosquito vectors but their effectiveness is still questioned. Therefore developing prediction model for mosquito prevalent areas in India will help in better encounter with the disease, causing less mortality.

Keywords: Environmental variables; Artificial Neural Network; Prediction model

1. Introduction
Mosquito borne diseases like malaria, dengue, chikungunya, etc. are potential threats to universal health, food security and growth today leading to prolonged hospitalization, increased medical costs and high mortality [1]. Several techniques including mechanistic models have been employed to predict dengue abundances worldwide. Out of the several diseases spread by mosquitoes dengue is classified as mosquito-borne viral infection causing severe illness and mortality in Asian sub-continent. Many techniques, protocol and models having different ways for prediction were developed to predict and control in endemic countries [2]. Generally, conventional methods to control mosquito population include use of
insecticides and larvicides but were not found to be much effective due to the development of resistant. Prediction models are a new approach and could be found to be superior over the conservative procedures because these methods use recent data on environmental conditions to predict dengue for a certain period of time [3].

Few researchers have reported different models which were based on characteristics like rainfall and temperature variations [4; 5; 6]. Some reported approaches for disease prediction models are either based on clinical conditions showing symptoms and other being clinical conditions without any symptoms. The latest approach which employ environmental factors for disease prediction had not been investigated much to provide appropriate clinical treatments. Among the popular techniques in the fields of bioinformatics and environmental studies the artificial neural networks (ANNs) were proven worthy to develop good predicting model [7; 8]. The cases where patient do not show any symptoms do not require any treatment, but they are identified as possible reservoirs for the parasite causing the disease [9]. There is high rate of prediction in cases of dengue which are untreated and asymptomatic leading to extensive growth of the dengue virus due to neglected clinical conditions [10]. In Asian continent, India being the country where the reported cases of dengue are very high every year. In India, the mosquito control programs has initiated in 1950s. There are an estimated 390 million infections every year out of which India has reported 800-900 deaths with dengue since 2015 [11]. Initially dengue is considered as one of the major diseases of the southern states of India but recently it has got the spreading sequences in all states of India. In the present paper dengue abundances in the geographical location of the Agra district area which is situated in the state of Uttar Pradesh on the bank of the Yamuna river with latitude 27.18°N and longitude 78.02°E were conducted. The related dengue abundances were monitored at the sites of (a) Dayalbagh, (b) Sikandra, (c) Agra cantonment, and (d) Kamla Nagar. The geographical location of individual site is shown in Fig. 1. The district of Agra has semi-arid climate with extreme hot summer (May-June), cold winters (December-February) and adequate rainfalls in monsoons (July-September) making favorable conditions for the breeding of dengue vectors.

Fig. 1 City map of Agra district showing the sites selected for study (all the four sites are marked in green boxes)

2. Methods

2.1 Data collection

Related datasets of environmental conditions like rainfall, relative humidity and temperature of the four sites from 2005-2018 were collected from the meteorological survey of India portal. Whereas the clinical data for dengue cases from 2005-2018 were gathered from the primary health clinics of Agra
district. The staff members at all the health centers were efficiently trained in dengue detection and treatment. The datasets are represented as the average for each year for four different sites as shown in Figure 2-5.

**Fig. 2** Average rainfall (cm); humidity (%); minimum and maximum temperature (°C) and number of dengue cases from 2005-2018 of Dayalbagh region

**Fig. 3** Average rainfall (cm); humidity (%); minimum and maximum temperature (°C) and number of dengue cases from 2005-2018 of Sikandra region
2.2 Artificial neural network

Artificial neural networks (ANNs) have wide range of applications and research attractions like image processing, data mining, stock prediction, weather forecasting, etc. [12]. One of the major key features of ANNs model is its accuracy with which it not only captures linear but also nonlinear data thus producing high end results [13]. Recently, there were some papers which reported prediction model for mosquito abundances by applying simple back propagation method [14]. In some papers feed forward neural network model with a single hidden layer having three types of nodes (i) input, (ii) hidden and (iii) output

Fig. 4 Average rainfall (cm); humidity (%); minimum and maximum temperature (°C) and number of dengue cases from 2005-2018 of Kamla Nagar region

Fig. 5 Average rainfall (cm); humidity (%); minimum and maximum temperature (°C) and number of dengue cases from 2005-2018 of Agra Cantonment region
nodes aid in time series modelling which is useful in forecasting [15]. In the present paper a simple prediction model was developed by using Deep learning studios to forecast the values for every month.

2.3 Measurements

In order to predict dengue abundances, we consider environmental variables (rainfall, relative humidity and temperature) as input. In this model, data collection was represented as average minimum and maximum temperature, average annual rainfall in cm, relative humidity and number of reported dengue cases. The datasets were collected, averaged and arranged separately for all the four sites from 2005-2018. Near about, 10-20% of the values for environmental data and 15-20% of clinical data were missing from the records. The dataset for environmental factors and clinical cases ranged from Jan 2005 to Dec 2018 at four different geographical areas as shown in Figure 2-5. In order to set best prediction model, the combination of environmental factors are first averages and tuned whereas clinical cases were also categorized depending on the symptoms of the patients. The data so obtained is referred as training data that is used for building more accurate model. The training of data and working strategy of the model was achieved using “Deep cognition” software with functions like an auto correlation (ACF), and partial auto correlation (PACF). Another feature known as the root mean square percentage error (RMSPE) calculates the experimental error which has occurred between the actual output and predicted output. The following equation is applied to measure RMSPE

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2} \times 100
\]

Where \(x_i\) is the observed value of the area and \(y_i\) is the predicted value and \(n\) be the number of predictions or observations in the corresponding geographical area [16]. Several cycles of predictions were run for individual study site to get accurate results. For each individual study site root mean square percentage error is estimated. Finally, the observed cases were compared with predicted cases.

3. Results

Prediction of dengue using ANN is a novel approach to correlate various environmental factors with the prevalence of the disease. This study stated that the maximum average rainfall was recorded in Dayalbagh region (site 1) as a result of which total no. of dengue cases reported from 2005-2018 was also highest in this particular site. The minimum average rainfall was recorded in Sikandra (site 3) therefore the reported dengue cases were also less as compared to site 1 stating direct correlation of rainfall with reported dengue cases. The results also showed no significant difference in relative humidity, maximum and minimum temperature in all the four study sites but variation in rainfall could be seen due to dense vegetation in Dayalbagh and Agra cantonment sites (site 4). Although Figure 1 shows that all the four sites are close by but vegetation greatly affects the rainfall and dengue prevalence. The RMSPE was measured and found that the geographical location of Dayalbagh (site 1) has the maximum error rate at 84% and Sikandra (site 2) has the minimum error rate as 27.3%. The data series of each geographical location is represented in Figure 2-5. In this study, we observe that site 1 has the most number of cases, while site 3 has the least number of reported dengue cases. The possible reason for high dengue prevalence in site 1 is the presence of dense vegetation and agricultural land and farms which act as breeding grounds for mosquito even during dry seasons. We also observed that site 4 although being densely populated have least number of dengue cases compared with other sites due to lack of breeding habitats for the dengue vector. The geographical area of site 1 contains maximum relative humidity whereas minimum relative humidity was
recorded in site 3 showing the significance of environmental parameters in dengue prediction. The predicted and observed dengue cases for 1 year from Jan 2019 to 31 Dec 2019 for 4 different sites are depicted in figure 6. The percentage error calculated was maximum for site 1 (84%) and minimum for site 4 (27.3%). This variation might be due to high vegetation in site 1 and low vegetation in site 4, thus affecting mean percentage error as well as no. of dengue cases.

![Graph showing predicted vs observed dengue cases from Jan 2019 to 31 Dec 2019 in four different sites of Agra district](image)

**Fig. 6** Predicted vs observed dengue cases from Jan 2019 to 31 Dec 2019 in four different sites of Agra district

4. Discussion

The aim of this study is to examine different variables across four different localities in Agra district with different vegetations and explored that environmental factors are essential to achieve precise predictive power. Hitherto, this is the first report where ANN is applied for dengue prediction in four localities of Agra district. The results showed that environmental factors such as rainfall, temperature and relative humidity were identified as vital parameter for achieving results in dengue prediction. The result varies from site to site where the geographical location of Dayalbagh has highest cases with high error rate as 84%. Our results are in disparity with the results reported by Thakur and Dharavath [2] where the locality of Khammam has lowest recorded malaria cases with highest error rate as 117%. There are various parameters which aid in calculating effect and may also affect the results [17]. July to September was reported to be the most suitable month for Dengue transmission and large number of clinical cases of dengue was also observed in these months especially in August [18]. Dengue vector *Aedes aegypti* is adapted to lay its eggs in artificial water containers in order to live in close contact to humans; hence monsoon will amplify mosquito breeding sites resulting in higher transmission of dengue [19]. The present investigation highlights that maximum dengue cases were noted between July and August in the geographical location of Dayalbagh because of the extensive paddy farming in the specific area that acts as natural breeding sited for *Aedes aegypti*. The results showed that minimum prevalence of dengue was reported in Kamla Nagar. The possible reason for less transmission of dengue might be because of dense population and housing leading to unfavorable condition for *Aedes aegypti* to breed [20]. Our results are in
conformity with the findings of Gupta and Reddy [21] demonstrating that relative humidity and temperature plays vital role in dengue transmission, when the humidity was recorded between 70-80% and temperature between 28-30°C there is a rise in the population of *Aedes aegypti* thus increasing the probability of dengue infection. Benelli and Mehlhorn [22] studied the impact of temperature on mosquito breeding and explained that temperature more than 30 °C and lesser than 16 °C results in decrease of artificial and natural breeding sites. From the results, we observe that site 3 (Kamla Nagar) and Site 2 (Sikandra) have less reported cases of dengue as compared to other sites because these two sites are densely population and Site 2 being leading industrial area in Agra have less vegetation index making unfavorable conditions for dengue vector to breed. Climatic factors along with various other factors such as type of treatment, immunity of individual and drug resistance in parasite play important role in disease transmission [23]. With the dengue prediction, prior action plans by primary health facility could be initiated and better treatment and diagnoses could be made available for patient at the time of disease transmission. Moreover, primary health centers and hospitals who actively took part in dengue control programs could educate the people living in endemic places about preventive methods that would greatly affect the transmission rate of the disease. Dengue prediction using ANN is a novel approach leading to estimation of number of cases by monitoring the environmental factors. This technique could not only be helpful in reducing the death rate associated with dengue but can also be effective in making primary decisions in those areas and localities where medical facility is confined.

5. Conclusion

The aim of this study is to estimate the dengue cases with clinical and environmental variables like rainfall, relative humidity and temperature from 2005-2018 using data available on the India meteorological department for the geographical location of Agra district, Uttar Pradesh, India. In this paper, we explored the usefulness of the artificial neural network (ANN) for dengue prediction in 4 major sites of Agra (Dayalbagh, Sikandra, Agra cantonment and Kamla Nagar) and prediction model was designed for the period of 2019 using the approach of neural network and compared with the observed values. The results show variation in the prediction model with respect to clinical and environmental variables corresponding to four different sites. Until now several approaches have been discovered for controlling mosquito vectors but their effectiveness is still questioned. Therefore developing prediction model for mosquito prevalent areas in India will help in better encounter with the disease causing less mortality. However, extensive research is still required in prediction of dengue using modern techniques of artificial intelligence for precise results in real practice.

References

[1] World Health Organization (2018) www.who.int/westernpacific/emergencies/surveillance/dengue
[2] Thakur S, Dharavath R (2019) Artificial neural network based prediction of malaria abundances using big data: A knowledge capturing approach. Clin Epidemiol Global Heal 7(1): 121-126
[3] Schaller JG (2000) Drugs for children: the world situation. Drug Informat J 34(1): 197-201
[4] Ahumada JA, Laoointe D, Samuel MD (2004) Modeling the population dynamics of *Culex quinquefasciatus* (Diptera: Culicidae), along an elevational gradient in Hawaii. J Medical Entomol 41(6): 1157-70
[5] Barrett MA, Humblet O, Hiatt RA, Adler NE (2013) Big data and disease prevention: from quantified self to quantified communities. Big data 1(3): 168-175
[6] Shone SM, Curriero FC, Lesser CR, Glass GE (2014) Characterizing population dynamics of *Aedes sollicitans* (Diptera: Culicidae) using meteorological data. J Medical Entomol 43(2): 393-402
[7] Dopazo J, Wang H, Carazo JM (1997) A new type of unsupervised growing neural network for biological sequence classification that adopts the topology of a phylogenetic tree. In International Work-Conference on Artificial Neural Networks (pp. 932-941). Springer, Berlin, Heidelberg

[8] Sahoo GB, Schladow SG, Reuter JE (2009) Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models. J Hydrology 378(3-4): 325-342

[9] Das NG, Dhiman S, Talukdar PK, Goswami D, Rabha B, Baruah I, Veer V (2015) Role of asymptomatic carriers and weather variables in persistent transmission of malaria in an endemic district of Assam, India. Infec Ecol Epidemiol 5(1): 25442

[10] Cecilia D, Kakade MB, Bhagat AB, Vallentyne J, Singh A, Patil JA, Shah PS (2011) Detection of dengue-4 virus in Pune, Western India after an absence of 30 years: its association with two severe cases. Virol Jour 8(1): 46

[11] Weeraratne TC, Perera MDB, Mohamed Mansoor MAC, Parakrama Karunaratne SHP (2013) Prevalence and breeding habitats of the dengue vectors Aedes aegypti and Aedes albopictus (Diptera: Culicidae) in the semi-urban areas of two different climatic zones in Sri Lanka. Int J Trop Insect Sci 33: 216–226

[12] Boonkiatpong K, Sinthupinyo S (2011) Applying multiple neural networks on large scale data. In Proceedings of International Conference on Information and Electronics Engineering (ICIEE 2011).

[13] Chen AS, Leung MT, Daouk H (2003) Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. Comp Operat Res 30(6): 901-923

[14] Lee KY, Chung N, Hwang S (2016) Application of an artificial neural network (ANN) model for predicting mosquito abundances in urban areas. Ecolog Informatics 36: 172-80

[15] Zhang G, Patuwo BE, Hu MY (1998) Forecasting with artificial neural networks: The state of the art. Internat J of forecasting 14(1): 35-62

[16] Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscient Model Develop 7(3): 1247-1250

[17] Kiang R, Adimi F, Soika V, Nigro J, Singhasivanon P, Sirichaisinthop J, Looareesuwan S (2006) Meteorological, environmental remote sensing and neural network analysis of the epidemiology of malaria transmission in Thailand. Geospatial health 71-84

[18] Pant, C (1987) Vector-Borne Diseases of man and their Socio-Economic impact. Int J Trop Insect Sci 8: 655–664

[19] Rigau-Pérez, JG, Clark GG, Gubler DJ, Reiter P, Sanders EJ, Vorndam, AV (1998) Dengue and dengue haemorrhagic fever. The Lancet 352(9132): 971-977

[20] Guzman MG, Halstead SB, Artsob H, Bucyi P, Farrar J, Gubler DJ, Hunsperger E, Kroege A, Margolis HS, Martinez E, Nathan MB, Pelegro JL, Simmons C, Yoksan S, Peeling RW (2010) Dengue: a continuing global threat. Nature Rev Microbiol 8(12): S7-S16

[21] Gupta B, Reddy BN (2013) Fight against dengue in India: progresses and challenges. Parasitol Res 112(4): 1367-1378

[22] Benelli G, Mehlhorn H (2016) Declining malaria, rising of dengue and Zika virus: insights for mosquito vector control. Parasitol. Res 115(5): 1747-1754

[23] Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GR, Simmons CP, Scott TW, Farrar JJ, Hay SI (2013) The global distribution and burden of dengue. Nature 7446: 504-507