PREDICTION OF TRANSMISSION RATES OF DENGUE IN NATIONAL CAPITAL TERRITORY DELHI USING MACHINE LEARNING MODELS

Vipasha Sharma
Civil Engineering Department
Indian Institute of Technology Roorkee
Roorkee, India
v_sharma@ce.iitr.ac.in
ORCID: 0000-0002-3812-4151

Sanjay Kumar Ghosh
Civil Engineering Department
Indian Institute of Technology Roorkee
Roorkee, India
sanjay.ghosh@ce.iitr.ac.in
ORCID: 0000-0001-7849-9313

Siddhartha Khare
Civil Engineering Department
Indian Institute of Technology Roorkee
Roorkee, India
siddhartha.khare@ce.iitr.ac.in
ORCID: 0000-0002-6996-4583

***Abstract*— The use of Machine Learning (ML) algorithms for predictive modeling to monitor transmission rates of dengue has gained significant attention worldwide. Earlier research has focused on specific weather variables and algorithms, there is a significant demand for models incorporating a wide range of variables and algorithms for superior performance. This study aims to predict transmission rates of dengue at the ward level in National Capital Territory (NCT) Delhi, India using different ML models. Data regarding incidence of dengue along with population and meteorological data as predictors during the period 2015-2022 have been used. The incidence data of dengue and population data have been collected from the Municipal Corporation Delhi (MCD) and Census of India, respectively. Meteorological data consisting of 20 parameters has been downloaded from NASA POWER. Five ML algorithms, including an ensemble approach, have been trained and validated. Comparative assessments using the Receiver Operating Characteristic (ROC) with the Area Under the Curve (AUC), accuracy, and F1 score have been carried out. It has been found that the accuracy of ensemble ML methods, such as Gradient Boosting, and Random Forest outperformed other models, such as Logistic Regression, Decision Tree, and Support Vector Machine. A correlation coefficient (r) of 0.40 has been used to evaluate the influence of meteorological variables on dengue transmission. Variables that exceed this threshold are considered to make a significant contribution to the transmission of dengue. Variable importance analysis shows significant contribution of Surface Soil Wetness (r= 0.48), Relative Humidity at 2 Meters (r= 0.45), Precipitation (r= 0.42), Wind Direction at 2 Meters (r= -0.54), Wind Direction at 10 Meters (r= -0.53), Temperature at 2 Meters (r= -0.47), and Earth Skin Temperature (r= -0.46). This study provides an excellent basis for future research, notably on dengue transmission modeling in dense urban environments and early warning systems using predictive models. This study provides insight on how ML algorithms may forecast dengue transmission rates and emphasizes the need to examine a variety of variables for model success.**

Keywords— Machine Learning, Dengue, Predictive Modeling, Remote Sensing, NASA POWER, Ensemble Approach

I. INTRODUCTION

In recent years, the utilization of Machine Learning (ML) algorithms has emerged as a pivotal tool in public health, particularly in the predictive modelling of infectious diseases[1], [2]. Dengue is one of the mosquito-borne viral infectious disease which poses a significant threat to global health[3], [4]. Thus, researchers are encouraged to monitor its various factors related to it occurrence and transmission rates. Occurrence of dengue and its transmission is mostly governed by the meteorological factors which provides favourable conditions for development and growth of dengue vectors[5]–[7]. While high temperatures and humidity may provide favourable conditions for the growth and survival of mosquito eggs, heavy rains provide new mosquito breeding locations[8]. Temperature and precipitation have been found to be positively linked with dengue occurrences in the majority of investigations[9]–[12]. Patil and Pandya [13] found that mean minimum temperature and rainfall are moderately significant, while mean wind speed is the least significant component and is marginally negatively correlated with dengue occurrences[13]. Earlier studies have predominantly focused on a limited set of meteorological variables, specifically temperature, precipitation, and humidity. It is observed that these variables are correlated amongst themselves, so the transmission of dengue is influenced by a combination of environmental, ecological, and human factors. Further, other factors such as urbanization, population density, and public health interventions also contribute significantly to the spread of dengue. Thus, there is a need to study the role of the additional factors in relation to the occurrence of dengue.

Another important observation is that the current trend of predictive modelling techniques in dengue involves the utilization of predominantly single ML-based models for dengue modelling[1], [12], [14]. As large variety of models are applied in dengue-related studies, many researchers felt that a single modelling technique might not be reliable. Thus, several studies compared the performance of different models to predict dengue outbreaks.

A study in Manila by Carvajal et al., [15] compared ML models using 5-year dengue data and climatic variables as predictor. Four different ML techniques, namely Generalized Additive Model (GAM), Seasonal Auto Regressive Integrated Moving Average with eXogenous factors (SARIMAX), Random Forest (RF) and Gradient Boosting Model (GBM) have been used. It was found that RF shows best performance with Mean Absolute Error (MAE), of (0.23, followed by GBM (0.24). Similar performance of RF was found and assessed by Zhao et al., [16] who compared RF and Artificial Neural

This research is funded by Prime Minister Research Fellowship (PMRF).

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00171

Network (ANN) to predict dengue burden, in Colombia, at national and local scales. It was concluded that using Mean Absolute Error (MAE), RF (0.86) performed slightly better than ANN (0.95) and provided better result at the national scale. Patil and Pandya [13] compared RF, Decision Tree (DT), Support Vector Regression (SVR), and Logistic Regression for 9 cities in Maharashtra, India. The study concluded that the RF regression is the best-fit regression model working on five out of nine cities[13]. Thus, in order to advance the field of predictive modelling in dengue studies, it is essential to conduct a thorough comparative assessment of model performance of different models. This may lead to development of a more comprehensive framework for models that encompass a broader spectrum of variables, aiming for superior predictive capabilities.

The present study aims to contribute to the predictive modelling by forecasting transmission rates of dengue at the ward level in the NCT Delhi, India. The geographical specificity of this study adds granularity towards understanding of disease dynamics at fine scale. Additionally, the study aims to identify the least optimal scale or administrative level that would enhance prediction accuracy. This study also employs various ML models, recognizing the diversity and complexity inherent in dengue transmission.

II. Study Area

The National Capital Territory (NCT) of Delhi, is situated within the extent 28° 52' 59.88"N, 76° 50' 16.08" E and 28° 24' 15.84" N, 77° 17' 23.64" E covering an area of 1,484 km^2 (Fig 1). It has diverse urban landscape, consisting of 11 administrative districts and 282 wards, each characterized by unique features and varying population densities[17]. The infrastructure of NCT encompasses a mix of high-rise buildings, slum settlements, and expansive residential colonies, reflecting its diversified urban character. The climate of the region is hot and humid, with temperatures ranging from 25°C to 45°C during the summer months[18], [19]. The monsoon season lasts from June to September with an annual average rainfall of 670.7 mm[19]. Such conditions may lead to flooding and stagnant water, providing ideal breeding conditions for mosquitoes.

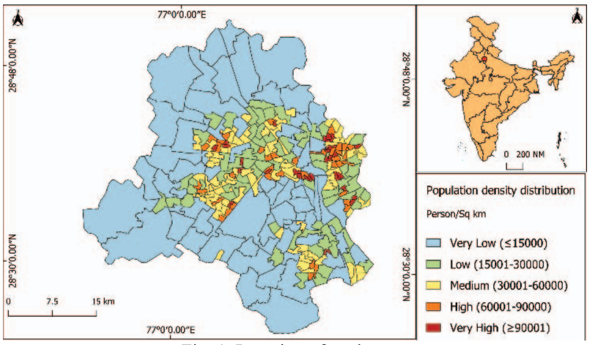

Fig. 1. Location of study area

III. Data and Methods

Here a comprehensive analysis of meteorological parameters and their effect on dengue transmission during 2015-2022 has been done. The details of data used, and methodology adopted are explained in the following sub-sections.

A. Details of dataset used

The primary data consists of daily hospitalized and laboratory confirmed dengue cases in Delhi during 2015-2022. The distribution of cases with respect to month during this period is shown in Fig 2.

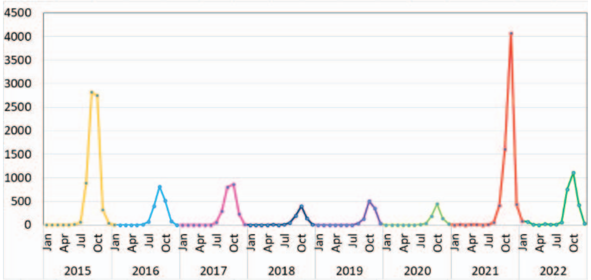

Fig. 2. Monthly distribution of dengue cases during 2015-2022

The administrative ward boundary of NCT Delhi has been downloaded from the website of Survey of India[20].

Population data for each ward in Delhi has been taken from Census of India 2011 (Census, 2011) along with Sample Registration System (SRS) statistical report for each year during study period[21]. The base population in the year 2011 projected for annual population determination during the study period using the Cohort Component method. Meteorological data consisting of 19 parameters during the study period are listed in Table 1 have been downloaded from NASA POWER data access viewer[22]. The listed parameters have been derived from the NASA's GMAO MERRA-2 assimilation model and GEOS 5.12.4 FP-IT.This data is available on daily scale and are available from NASA POWER data acesss viewer. The horizontal resolution of the dataset is ½° x ⅝° grid.

TABLE 1. Details of parameters used in study

T2M	Temperature at 2 Meters (ºC)
TS	Earth surface temperature (ºC)
T2M_RANGE	Temperature at 2 Meters Range (ºC)
T2M_MAX	Temperature at 2 Meters Maximum (ºC)
T2M_MIN	Temperature at 2 Meters Minimum (ºC)
RH2M	Relative Humidity at 2 Meters (%)
QV2M	Specific Humidity at 2 Meters (g/kg)
PRECTOTCORR	Precipitation Corrected (mm/day)
PS	Surface Pressure (kPA)
WS2M	Wind Speed at 2 Meters (m/s)
WS2M_MAX	Wind Speed at 2 Meters Maximum (m/s)
WS2M_MIN	Wind Speed at 2 Meters Minimum (m/s)
WD2M	Wind Direction at 2 Meters (Degrees)
WS2M_RANGE	Wind Speed at 2 Meters Range (m/s)
WS10M	Wind Speed at 10 Meters (m/s)
WS10M_MAX	Wind Speed at 10 Meters Maximum (m/s)
WS10M_MIN	Wind Speed at 10 Meters Minimum (m/s)
WS10M_RANGE	Wind Speed at 10 Meters Range (m/s)
WD10M	Wind Direction at 10 Meters (Degrees)
GWETTOP	Soil wetness
PD	Population density (person/sq km)

B. Methodology adopted

To model dengue cases with variation of meteorological parameters, five different machine learning models have been used. The flowchart of methodology adopted is shown in Fig 3. The first step is to collect and download data from various sources. As data comes from different sources, preprocessing is a very important step to bring all data in a common format and spatial resolution. This is done using pre-processing of the data. Pre-processing step includes segregation of columns, formatting of dates and accounting for the missing values in the dataset.

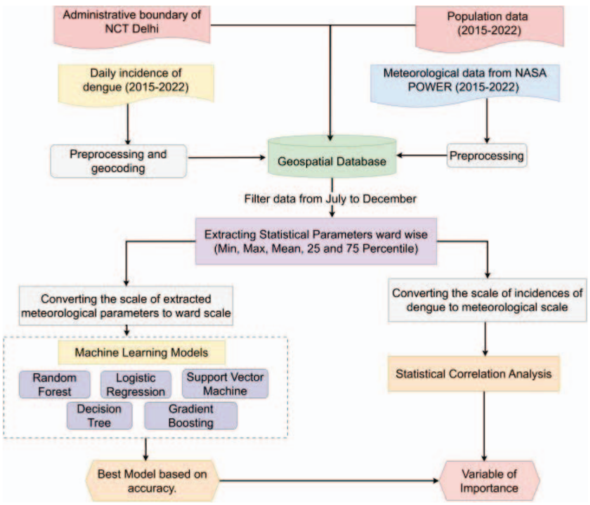

Fig. 3. Flowchart of methodology

As the vulnerable period of dengue in NCT Delhi is generally from July to December, therefore data only for this period has been extracted in order to avoid biasness. Statistical parameters are then computed for all 20 parameters such as minimum, maximum, mean, 25 percentiles and 75 percentiles. Therefore, two approaches have been adopted. The first approach uses statistical correlation analysis to understand the relationship between different variables and identify potential patterns or trends. For this, the cases of dengue are rescaled to a coarser resolution of meteorological parameters, i.e. at ½° x ⅝° grid. Statistical correlation analysis between 20 parameters and cases of dengue has been carried out.

In the second approach, for modelling, the meteorological parameters are re-scaled at ward level to model the transmission of dengue with respect to the selected parameters. The derived statistical parameters are associated to ward scale by linking the centroid of the ward to the nearest data point. To model the transmission of dengue at ward scale, DT, SVR, LoR, RF and GBM have been used. For modelling, data has been split as training and testing dataset in ratio of 80:20 at random state 42.

Each ML model used in the study have been configured with specific parameter values to model the incidences of dengue. For DT model, a complexity parameter (ccp_alpha) of 0.0, indicating no preference for simpler trees, a criterion (criterion) of 'gini' to measure impurity, and default settings for class weight, minimum samples per leaf, minimum samples for splitting, and the splitter strategy has been used. RF, on the other hand, used bootstrap sampling (bootstrap=True), a 'gini' criterion for splitting, and selected the square root of the number of features (max_features='sqrt'). For rest of the parameters, such as minimum samples per leaf, minimum samples for splitting, and the number of estimators, default settings have been used. Parameters for LoR included a regularization strength (C) of 1.0, the 'lbfgs' solver, and default settings for other parameters such as fit intercept, maximum iterations, and penalty type. For SVR, a penalty parameter (C) of 1.0, an 'rbf' kernel, have been used. For other parameters such as cache size, maximum iterations, and kernel coefficient default settings have been used. Lastly, for GB model, a learning rate (learning_rate) of 0.1, a 'friedman_mse' criterion for splitting, have been used. Other parameters including the number of estimators, maximum depth, and early stopping criteria have been taken as default. The configurations used for the models have been selected based on trial and error, and the efficacy in optimizing model performance.

Each technique is further evaluated based on key metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and R-squared (R^2)[1], [2]. MAE, MSE, and RMSE focus on the accuracy of predictions and closeness to the actual values[23]. Lower values for these metrics indicate better accuracy[2], [24]. R^2 provides an overall measure of the variance in the data. Higher R^2 values (greater than 0.5) may suggest a better fit, but it is essential to consider simultaneously other metrics to get a comprehensive view of the model performance[2].

IV. RESULTS AND DISCUSSION

A. Statisitcal correlation analysis

Due to differences in resolution of meteorological data and cases of dengue, second approach of analysis has been performed by converting and rescaling cases of dengue data at ½° x ⅝° grid. This leads to reduction in data points in gridded dataset which comprise of 48 grids. The statistical correlation analysis of number of dengue cases with statistical parameters of meteorological variables is shown in Fig 4. The Correlation Coefficient (CC) indicates the strength and direction of linear relationship between each meteorological variable and the occurrence of dengue cases.

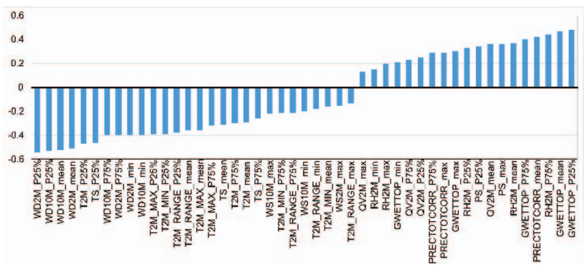

Fig. 4. Result of statistical correlation analysis

Analysis of statistical parameters at 25% and 75% percentile shows a negative correlation with WD2M_P25%, WD10M_P25%, T2M_P25%, TS_P25% , thus suggesting a negative association between these meteorological variables and the number of dengue cases. For example, a higher negative correlation for WD2M_P25% (-0.54) implies that for 25% times at 2m height wind direction is strongly associated to higher number of dengue cases. On the other hand, positive correlation with QV2M_min, T2M_max, PS_min, RH2M_max, GWETTOP_mean suggests a positive association between these meteorological variables and the number of dengue cases.

At higher wind speeds the mosquitoes, may make it difficult to fly. As strong winds may disperse mosquitoes, making it challenging for them to locate and feed on hosts, including humans. This may temporarily reduce the risk of dengue transmission[15]. Further, wind direction may influence the movement of mosquitoes. For example, if the wind is blowing towards populated areas, mosquitoes may be carried into urban or suburban regions, increasing the risk of dengue transmission in those areas[15].

Temperature plays a significant role in the transmission dynamics of dengue as it affects the growth of both mosquito vector and dengue virus[11]. *Aedes* mosquitoes, particularly *Aedes aegypti*, thrive in warmer climates[5]. The development and activity of *Aedes Aegypti* are strongly influenced by temperature. Higher temperatures generally lead to shorter incubation periods for mosquito eggs, faster development of larvae, and increased activity of adult mosquitoes. Also, high humidity levels may enhance mosquito survival and activity[25]. Soil moisture content on the other hand, may influence the availability of water for mosquito breeding sites. Adequate moisture is essential for the development of mosquito larvae[26]. The top-layer soil moisture content may affect the presence of suitable breeding sites for *Aedes* mosquitoes, thereby influencing mosquito population dynamics and dengue transmission.

B. Modelling incidence of dengue at ward scale

The study focusses on the implementation of a predictive modeling approach to estimate the incidence of dengue within ward level as an administrative unit. This section presents the outcomes of the analysis which is primarily aimed at gaining insights into the spatial distribution and factors influencing dengue occurrence at this localized level. Table 2 provides a summary of the performance metrics for different machine learning techniques applied for modelling incidences of dengue. These metrics offer insights regarding accuracy, precision and overall goodness of fit of each model.

TABLE 2. Summary of model performance

Model	**Performance metrics**				**Inference**
	MAE	***MSE***	***RMSE***	***R^2***	
DT	0.49	0.59	0.76	-0.34	Poor agreement
SVR	0.50	0.48	0.33	0.12	Weak agreement
RF	0.47	0.39	0.12	0.54	Moderate agreement
LoR	0.49	0.34	0.22	0.32	Weak agreement
GBM	0.46	0.31	0.30	0.56	Moderate agreement

It is observed that DT performs poorly as compared to other ML models based on MAE, MSE and R^2 whereas, GBM shows the lowest values for MAE (0.46), MSE (0.31), and RMSE (0.56) among the listed models, indicating superior predictive performance. The positive and higher value of R^2 further supports the model's effectiveness in capturing variance in the data. GBM is followed by RF in terms of predictive performance metrics.

C. Variable of importance

It is observed that GBM is the best model for modelling cases of dengue. Dengue is highly sensitive to meteorological conditions. The vectors, their growth, and the development of virus in vectors and hosts occur at specific meteorological conditions. The variable of importance for this model is shown in Fig 5. It is observed that at ward level, population density (0.58) is major factor governing transmission of dengue. In general, high population density is associated with urbanization. Urban areas tend to have a higher concentration of human-made structures such as open water storage containers, flower pots, discarded tires that may serve as mosquito breeding sites [27]. Further, in densely populated areas, inadequate waste management and sanitation infrastructure may contribute to the accumulation of water-holding containers, creating favorable conditions for mosquito breeding[28].

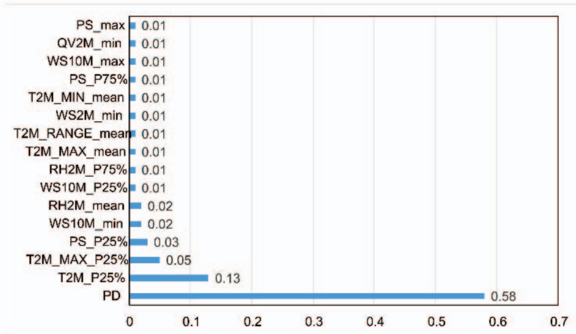

Fig. 5. Variable of importance

The next variable of importance is T2M_P25% (0.13), followed by T2M_MAX_P25% (0.05), PS_P25% (0.03), WS10m_min (0.02), RH2m_mean (0.02). Similar results have been reported by different studies where temperature, rainfall, humidity [29]–[32] have been found to be the governing factors. An increase in temperature (26˚C-28˚C to 30˚C) increases the dengue risk by providing favorable conditions as it enhances the rate of mosquito development, reducing virus incubation time, increasing the rate of transmission [18]. A similar condition is observed in the case of rainfall as heavy rainfall may lead to the washing away of eggs, larvae and pupae from breeding grounds. At the same time, the residual rainwater may create breeding habitats [33]. The speed of wind may affect their ability of mosquitoes to fly, feed, or rest[15]. Therefore, wind speed may also be considered as an important factor governing the spread of dengue.

It may be noted that resolution of input data plays a important role as it may impact the accuracy of modelling approaches[34]. Higher resolution models capture fine-scale details and variations in the data leading to better representation of complex processes and phenomena. In climate modeling, higher spatial resolution may yield better simulation of local weather patterns and capture short-term fluctuations in data patterns[35]. Certain variables may be critical at smaller scales but may not be apparent at coarser resolutions and vice versa. At lower spatial resolutions, certain variables may dominate the model due to their impact on larger-scale patterns. However, this may overlook variables that are crucial at smaller scales.

V. CONCLUSION

This study gives valuable insights into the complex dynamics of dengue transmission using different meteorological variables, demographic factors, and the incidence of dengue at the ward level. GBM and RF are found to be the best-performing models for NCT Delhi. A threshold value of 0.40 for correlation coefficient has been found suitable to evaluate the effect of meteorological variables on dengue transmission. The statistical correlation analysis reveals that the parameters close to the ground have a positive correlation to dengue incidences and negative correlation for parameters which are a higher height from ground. Thus it may be inferred that wind direction and wind speed at 10m above ground level, show an inverse association with incidences of dengue cases. Overall, population density is a major factor contributing to dengue followed by temperature at 2 m and surface pressure. The findings of this study also highlights the importance of combining data at same resolution and selecting appropriate modeling techniques to achieve accurate predictions at the localized ward level by comparing different techniques.

ACKNOWLEDGMENT

Authors extend their sincere appreciation to Public health department, Government of Delhi, India for providing dengue data. Also the first author would like express her gratitude to the Ministry of Education for the research grant [PM-31-22-670-414].

REFERENCES

[1] K. Roster, C. Connaughton, and F. A. Rodrigues, 'Machine-Learning-Based Forecasting of Dengue Fever in Brazilian Cities Using Epidemiologic and Meteorological Variables', *Am. J. Epidemiol.*, vol. 191, no. 10, pp. 1803–1812, 2022, doi: 10.1093/aje/kwac090.

[2] P. Guo *et al.*, 'Developing a dengue forecast model using machine learning: A case study in China', *PLoS Negl. Trop. Dis.*, vol. 11, no. 10, pp. 1–22, 2017, doi: 10.1371/journal.pntd.0005973.

[3] Z. Zeng, J. Zhan, L. Chen, H. Chen, and S. Cheng, 'Global, regional, and national dengue burden from 1990 to 2017: A systematic analysis based on the global burden of disease study 2017', *EClinicalMedicine*, vol. 32, p. 100712, 2021, doi: 10.1016/j.eclinm.2020.100712.

[4] WHO, 'Dengue and severe dengue', 2021. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue (accessed Jun. 19, 2021).

[5] P. C. Wu, H. R. Guo, S. C. Lung, C. Y. Lin, and H. J. Su, 'Weather as an effective predictor for occurrence of dengue fever in Taiwan', *Acta Trop.*, vol. 103, no. 1, pp. 50–57, 2007, doi: 10.1016/j.actatropica.2007.05.014.

[6] P. Ceccato, B. Ramirez, T. Manyangadze, P. Gwakisa, and M. C. Thomson, 'Data and tools to integrate climate and environmental information into public health', *Infect. Dis. Poverty*, vol. 7, no. 126, pp. 1–11, 2018, doi: 10.1186/s40249-018-0501-9.

[7] Y. L. Hii, H. Zhu, N. Ng, L. C. Ng, and J. Rocklöv, 'Forecast of Dengue Incidence Using Temperature and Rainfall', *PLoS Negl. Trop. Dis.*, vol. 6, no. 11, p. e1908, Nov. 2012, doi: 10.1371/JOURNAL.PNTD.0001908.

[8] S. Mala and M. K. Jat, 'Implications of meteorological and physiographical parameters on dengue fever occurrences in Delhi', *Sci. Total Environ.*, vol. 650, pp. 2267–2283, 2019, doi: 10.1016/j.scitotenv.2018.09.357.

[9] C. Davis *et al.*, 'A regional suitable conditions index to forecast the impact of climate change on dengue vectorial capacity', *Environ. Res.*, vol. 195, no. October 2020, 2021, doi: 10.1016/j.envres.2021.110849.

[10] N. Gyawali, B. J. Johnson, S. M. Dixit, and G. J. Devine, 'Patterns of dengue in Nepal from 2010–2019 in relation to elevation and climate', *Trans. R. Soc. Trop. Med. Hyg.*, no. November, 2020, doi: 10.1093/trstmh/traa131.

[11] J. Xiao *et al.*, 'Weather variables and the El Niño Southern Oscillation may drive the epidemics of dengue in Guangdong Province, China', *Sci. Total Environ.*, vol. 624, pp. 926–934, 2018, doi: 10.1016/j.scitotenv.2017.12.200.

[12] N. Nuraini, I. S. Fauzi, M. Fakhruddin, A. Sopaheluwakan, and E. Soewono, 'Climate-based dengue model in Semarang, Indonesia: Predictions and descriptive analysis', *Infect. Dis. Model.*, vol. 6, pp. 598–611, 2021, doi: 10.1016/j.idm.2021.03.005.

[13] S. Patil and S. Pandya, 'Forecasting Dengue Hotspots Associated With Variation in Meteorological Parameters Using Regression and Time Series Models', *Front. Public Heal.*, vol. 9, no. November, pp. 1–16, 2021, doi: 10.3389/fpubh.2021.798034.

[14] N. Iqbal, R. Scholar, and M. Islam, 'Machine learning for dengue outbreak prediction: An outlook', *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 1, pp. 93–102, 2017, [Online]. Available: www.ijarcs.info.

[15] 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 Infect. Dis.*, vol. 18, no. 1, pp. 1–15, 2018, doi: 10.1186/s12879-

018-3066-0.

[16] N. Zhao *et al.*, ‘Machine learning and dengue forecasting: Comparing random forests and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia’, *PLoS Negl. Trop. Dis.*, vol. 14, no. 9, pp. 1–16, 2020, doi: 10.1371/journal.pntd.0008056.

[17] O. Telle *et al.*, ‘Social and environmental risk factors for dengue in delhi city: A retrospective study’, *PLoS Negl. Trop. Dis.*, vol. 15, no. 2, pp. 1–17, 2021, doi: 10.1371/journal.pntd.0009024.

[18] V. G. Ramachandran, P. Roy, S. Das, N. S. Mogha, and A. K. Bansal, ‘Empirical model for estimating dengue incidence using temperature, rainfall, and relative humidity: a 19-year retrospective analysis in East Delhi’, *Epidemiol. Health*, vol. 38, p. e2016052, 2016, doi: 10.4178/epih.e2016052.

[19] P. Guhathakurta, N. Narkhede, P. Menon, A. K. Prasad, and S. T. Sable, ‘Climate Research and Services Observed Rainfall Variability and Changes over Assam State’, 2020. doi: ESSO/IMD/HS/Rainfall Variability/06(2020)/30.

[20] SOI, ‘Survey of India’, 2023. https://surveyofindia.gov.in/ (accessed Jul. 06, 2023).

[21] 2011 Census, ‘Census’, *Office of the Registrar General & Census Commissioner, India Affairs, Ministry of Home India, Government of India*, 2011. https://censusindia.gov.in/2011census/population_enumeration.html (accessed Sep. 29, 2021).

[22] ‘NASA Data Access Viewer’, 2023. https://power.larc.nasa.gov/data-access-viewer/ (accessed Dec. 02, 2023).

[23] V. J. Jayaraj, R. Avoi, N. Gopalakrishnan, D. B. Raja, and Y. Umasa, ‘Developing a dengue prediction model based on climate in Tawau, Malaysia’, *Acta Trop.*, vol. 197, no. February, p. 105055, 2019, doi: 10.1016/j.actatropica.2019.105055.

[24] H. Sun *et al.*, ‘Spatio-temporal analysis of the main dengue vector populations in Singapore’, *Parasites and Vectors*, vol. 14, no. 1, pp. 1–11, 2021, doi: 10.1186/s13071-020-04554-9.

[25] Y. Chen *et al.*, ‘Effects of natural and socioeconomic factors on dengue transmission in two cities of China from 2006 to 2017’, *Sci. Total Environ.*, vol. 724, p. 138200, 2020, doi: 10.1016/j.scitotenv.2020.138200.

[26] O. Mudele, A. C. Frery, L. F. R. Zanandrez, A. E. Eiras, and P. Gamba, ‘Modeling dengue vector population with earth observation data and a generalized linear model’, *Acta Trop.*, vol. 215, no. December 2019, p. 105809, 2021, doi: 10.1016/j.actatropica.2020.105809.

[27] B. R. Egid *et al.*, ‘Review of the ecology and behaviour of Aedes aegypti and Aedes albopictus in Western Africa and implications for vector control’, *Curr. Res. Parasitol. Vector-Borne Dis.*, vol. 2, no. October 2021, p. 100074, 2022, doi: 10.1016/j.crpvbd.2021.100074.

[28] B. Bisht, R. Kumari, B. N. Nagpal, H. Singh, and S. Kumar, ‘Influence of environmental factors on dengue fever in Delhi’, *Int. J. Mosq. Res.*, vol. 6, no. 2, pp. 11–18, 2019.

[29] J. Ashby, M. Moreno-Madriñán, C. T. Yiannoutsos, and A. Stanforth, ‘Niche Modeling of Dengue Fever Using Remotely Sensed Environmental Factors and Boosted Regression Trees’, *Remote Sens.*, vol. 9, no. 328, p. 15, 2017, doi: 10.3390/rs9040328.

[30] Q. T. Bui, Q. H. Nguyen, V. M. Pham, M. H. Pham, and A. T. Tran, ‘Understanding spatial variations of malaria in Vietnam using remotely sensed data integrated into GIS and machine learning classifiers’, *Geocarto Int.*, vol. 34, no. 12, pp. 1300–1314, 2019, doi: 10.1080/10106049.2018.1478890.

[31] X. Y. Ji, L. Y. Huang, J. Song, C. N. Fei, J. Liu, and H. Liu, ‘Short-term effects of meteorological factors, air pollution, and sunspot on childhood hand, foot, and mouth disease in Tianjin, China: a new time series regression, 2014–2018’, *Environmental Science and Pollution Research*, vol. 27, no. 29. pp. 37022–37035, 2020, doi: 10.1007/s11356-020-09794-x.

[32] N. Sumanasinghe, A. Mikler, C. Tiwari, and J. Muthukudage, ‘Geo-statistical dengue risk model using GIS techniques to identify the risk prone areas by linking rainfall and population density factors in Sri Lanka Geo-statistical Dengue Risk Model using GIS techniques to identify the risk prone areas by linking rain’, *Ceylon J. Sci.*, vol. 45, no. 3, pp. 39–46, 2016, doi: 10.4038/cjs.v45i3.7399.

[33] S. Bal and S. Sodoudi, ‘Modeling and prediction of dengue occurrences in Kolkata, India, based on climate factors’, *Int. J. Biometeorol.*, vol. 64, no. 8, pp. 1379–1391, 2020, doi: 10.1007/s00484-020-01918-9.

[34] V. Sharma, S. K. Ghosh, and S. Khare, ‘a Proposed Framework for Surveillance of Dengue Disease and Prediction’, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 48, no. M-1–2023, pp. 317–323, 2023, doi: 10.5194/isprs-archives-XLVIII-M-1-2023-317-2023.

[35] L. M. Stolerman, P. D. Maia, and J. Nathan Kutz, ‘Forecasting dengue fever in Brazil: An assessment of climate conditions’, *PLoS One*, vol. 14, no. 8, pp. 1–16, 2019, doi: 10.1371/journal.pone.0220106.