Prediction of Dengue Outbreaks in Mumbai Region Based on Disease Surveillance and Meteorological Factors using Big Data Approach

Asha Bharambe[1], Dhananjay Kalbande[2]
[1]*Department of Information Technology, Vivekanand Education Society's Institute of Technology, Mumbai, India*
[2]*Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, India*

Abstract **- One of the most prevalent vector-borne disease in India is Dengue. This study investigated the effects of various environmental/climatic factors on dengue incidence using time series analysis and machine learning approach using the big data environment. The aim of the current study is to understand environmental impact on vector-borne disease (Dengue) in a particular area from Mumbai and use it for prediction of the disease. Disease Incidence Data for dengue was collected from 2012 to 2019 from the study hospital in Mumbai while the climatic data was obtained from the web resources. Statistical time series analysis methods and machine learning techniques are used to make predictions. Our preliminary analysis has identified that there is a correlation between Dengue incidence rate and climatic conditions. Amongst machine learning model, random forest technique gave the results with RMSE of 47.08. Amongst time series analysis techniques, SARIMA model gave significant results with RMSE as 20.79 in univariate analysis as compared to ARIMA with RMSE of 61.27. The multivariate VAR model gave results with RMSE of 27.32. In both univariate and multivariate models, we concluded that when confounding factors are incorporated into forecasting model, it significantly improves AIC value (an AIC value of 5.89 was obtained for VAR model).**

Index Terms **- Big Data, Dengue, Machine Learning, Prediction, Time-Series.**

I.INTRODUCTION

Dengue transmission predominantly occurs in tropical and subtropical areas where mosquitoes can survive and multiply. Currently there are more than 100 countries that are endemic to dengue disease [1], [2] with about 390 million dengue virus infections per year [3] as shown in Figure-1 and 40% of world's population, about 3.9 billion people [2] are at the risk of suffering from the disease.

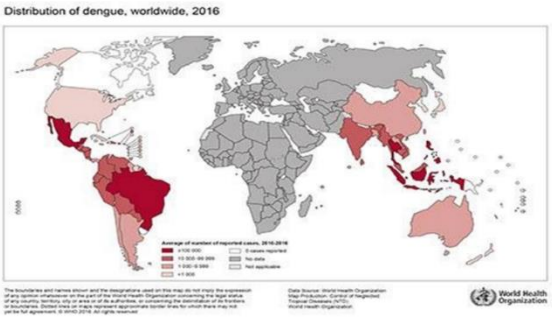

Figure 1: Prevalence of Dengue across world from 2010-2016. [3]

The Indian government has identified Dengue as important diseases to track from 12 diseases in laboratory confirmed form and from 22 diseases in presumptive form [4]. Hence a nationwide surveillance called Integrated Disease Surveillance Project (IDSP) [4] is conducted to identify these disease incidences.

Though nationwide health plans have succeeded in reducing fatality of few vector-borne disease to a certain extent, there is however a need for further improved and innovative dengue control measures [5] and efficient area-specific approach to predict the extent of these diseases. There are very few prediction systems developed so far to predict the spread of vector borne diseases in Mumbai city taking into considerations its environmental, social, and economic factors. Further there is no integrated data (clinical, disease and its environmental factors) available to model the disease.

Dengue is spread by the day-biting Aedes Aegypti mosquito that breeds in fresh water stored in

containers or collected in and around people's homes in urban and peri-urban areas. In India, the incidence and severity of dengue has increased rapidly in recent years.

Since there is no specific treatment for dengue fever [1], prevention is the best model that can be adopted. The Government and public health sector can be better prepared if they can predict the incidence rate and can take measures to reduce this rate. Further if the spread pattern and rate can be identified, they can be successful in controlling the spread of disease well in time and thus would be able to control the epidemic outbreak.

II. LITERATURE REVIEW

This section presents the study of existing approaches taken for disease forecasting and the effect and use of confounding factors for forecasting. Traditional time series approach and modern machine learning approaches have been conducted to predict the outbreak by use of climatic conditions [6]-[8]. Many studies have reported changing spatial patterns for dengue transmission. The reasons for such changes are related to several factors, ranging from the globalization of travel and trade, which favours the propagation of pathogen and vectors, to climatic changes or modified human behaviour [6]. Temperature and humidity are important climatic factors in mosquito population and disease transmission dynamics. Temperature influences the developmental rates, mortality, and reproductive behaviour of mosquitoes. Humidity provides the water that serves as a habitat for larvae and pupae [7].

A seasonal autoregression integrated moving average model SARIMA(0,0,1)(0,1,1)12 was developed in [9] for Rajasthan region. This model demonstrated the seasonal pattern in the data and used it to predict the outbreaks of Dengue. Univariate SARIMA (0,1,1)(0,1,1)52 model was used in [10] for modelling and predicting dengue cases in Guadeloupe, French West Indies. The study carried in [10] also suggested statistically significant results when external independent variables like temperature is included in model building. Association Rule Mining (ARM) methodology was suggested in [11] to describe relationship among different attributes (symptoms) of Dengue. The authors used fuzzy representation of the attributes and defined fuzzy association rules based on which a classifier model was built which gave the predicted dengue incidence as HIGH or LOW. An big data framework was proposed in [12], for rolling window time series prediction for large scale data. A Non-linear auto-regressive model (NARX) was built in [12] using Support Vector Machine. The proposed framework in [12] was tested on synthetic data set and gave significant improvement over the benchmark framework.

Significance of climatic parameters along with movement patterns and spatial heterogeneity was studied in [13] for 50 districts in Thailand. The study concluded that these factors were influential.

A study carried out in Singapore [14] developed a weather-based forecast model and were able to predict the cases up to 16 weeks in advance. Another study in Singapore [15] developed an early warning system based on meteorological data, vector surveillance data and population statistics.

Time series regression models for dengue forecasting were developed in [16] using monthly rainfall data, rainy days, temperature, humidity, wind and incidence case data.

A dynamic, ensemble learning approach was taken in [17] using weather and population data to predict dengue cases in Brazil.

It is suggested that climate-based disease forecasting models in India should be refined and tailored for different climatic zones, instead of use of a standard model across the country [8]. It is also suggested that dengue incidence be studied based on certain cohort attributes like community[13], geographic locations[13][14], and age-based [13].

A. Findings and Gaps

Traditional time series models are effective for forecasting the values but cannot be generalized. The model parameters can changes depending on the region

The survey identified climatic conditions to be one of the major confounding factors and suggested that the existing forecasting model needs to be defined based on the area and should be custom-made for different climate zones in India. Thus, identifying the confounding factors for the given study area needs to be carried out.

Incorporating these factors into the forecasting model and verifying the effect of these on forecast accuracy needs to be carried out.

Further, a big data approach for time series forecasting is not tested for real-time multi-variate time-series data.

By identifying the gaps and addressing the issues, the current study identifies the association between weather and dengue for the given region and incorporates it to forecast future incidence cases using traditional time series and machine learning approaches.

III. DATA MODEL

Real time data was collected from the study hospital in Mumbai region of Maharashtra state in India. The data of Dengue incidence cases for the duration of 8 years from 2012 till 2019 was collected through the study hospital. The study uses daily incidence cases for the past 8 years. No personal information was collected. The data attributes for incidence case consists of the age, gender, date of admission, test result, and the ward(area). The area parameters help us to distribute the data which would further enable us to create a model for each area. Confounding factors (climatic data) like temperature, humidity, wind and pressure was collected through web source [14] for the same duration of 2012 to 2019. The frequency of these time series were mismatched and combining them was done based on its representation described in [15] using TemporalRDD. A join operation was performed on these time series.

A. Study Area

As a pilot study, the research is restricted to only Mumbai and its suburban area. Mumbai is capital of Maharashtra state in India and is located at 19.0760° N, 72.8777° E. For ease of administration the city is divided into 24 wards[16].

B. Data Preprocessing

Pre-processing of data helps to prepare the data for the task of building the forecast model. The real-time data collected will be processed using these techniques – data cleaning, data transformation and data reduction. Data cleaning will handle missing values and remove any noisy data. For handling missing values for attribute “ward(area)”, a default value of “others” was filled in whereas for the missing value of the attribute “Date of admission”, the record was deleted. A simple analysis of minimum and maximum value for “Date of admission” attribute helped in identification of outliers. There were cases where the date was mistyped, and a manual analysis was done to replace the value with appropriate values. Data was converted for a single time series to multiple time series based on the “ward(area)” attribute.

The data was reduced by using numerosity reduction technique. The data was aggregated based on the “Date of admission” attribute.

C. Data Analysis

The disease incidence data for Dengue is collected from the study hospital in Mumbai. As shown in Figure-2, Mumbai region has shown fluctuating patterns in the disease. Available data indicates that there is an increasing number of cases for dengue every year till 2016 and then it has decreased.

This caused a dilemma to establish a trend and a pattern of disease. In this context it was felt that the time series analysis would be a reliable tool for finding the trend and patterns and forecasting future incidences.

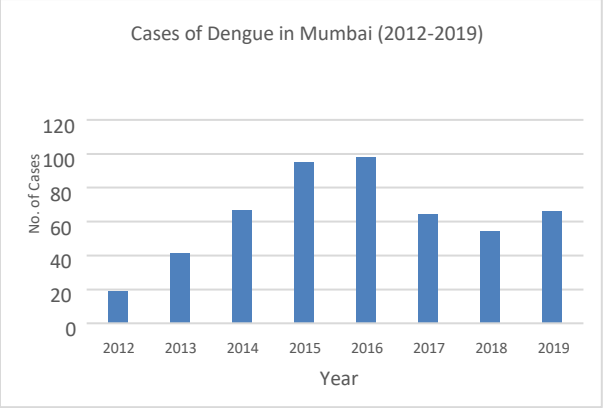

Figure 2: Dengue Status in Mumbai from 2012 to 2019 - data extracted from the study Hospital.

Statistical analysis of the study focused on identifying the following relationships.

1. Relationship between the meteorological factors and the incidence data.
2. Relationship between the current incidence cases with the previous incidence cases in the given area.

For identifying the relation between meteorological factors and incidence data, the correlation was found between them and is shown in Table-I. The preliminary analysis shows that there is a correlation between the climatic data and the incidence data. Factors like humidity shows a positive correlation with the incidence cases.

Table I: Correlation matrix of Incidence cases and Climatic Factors

	Cases	Average Temperature	Average Humidity	Average Wind	Average Pressure
Cases	1.000	0.027	0.3	-0.051	-0.11
Average Temperature	0.027	1.000	0.18	0.03	-0.32
Average Humidity	0.3	0.18	1.000	0.4	-0.77
Average Wind	-0.051	0.03	0.4	1.000	-0.55
Average Pressure	-0.11	-0.32	-0.77	-0.55	1.000

For identifying the relationship between the current incidence cases with the previous incidence cases in the given area, time series models were developed, which is explained in next section. Also, the research proposes a forecasting model by including the past data and climatic factors for early prediction of dengue.

IV. METHODS

Time series modelling is one of the scientific ways to learn about the future or predict the future. Time series captures data at a particular interval. In our study, we have a time series data for representing incidence cases and climatic factors. Every time series has three components namely: trend, seasonality and randomness or error. Trend component describes the overall changes that are taking place and the direction in which it is moving (upwards or downwards). Seasonal component describes how the changes are occurring in the given duration of time i.e. yearly or monthly etc. Random component describes the activities that are not explained by the trend and seasonal component. The seasonal decomposition for the Dengue incidence cases is shown in Figure 3. The trend shows an initial rise and then a fall, the seasonal pattern is regular.

There are many traditional methods used for prediction using timeseries such as auto-regressive (AR), auto-regressive moving average (ARMA), autoregressive integrated moving average (ARIMA), seasonal forecast model (SARIMA). These models can forecast the values based on the previous values, for eg. Based on previous data of dengue incidence, these models can predict the future values of dengue incidence. There are certain models which take help of other influencing factors and predict the future values. For eg, we can predict dengue incidence cases based not only on previous dengue cases but certain confounding factors like temperature, rainfall etc. These models are known as multivariate models.

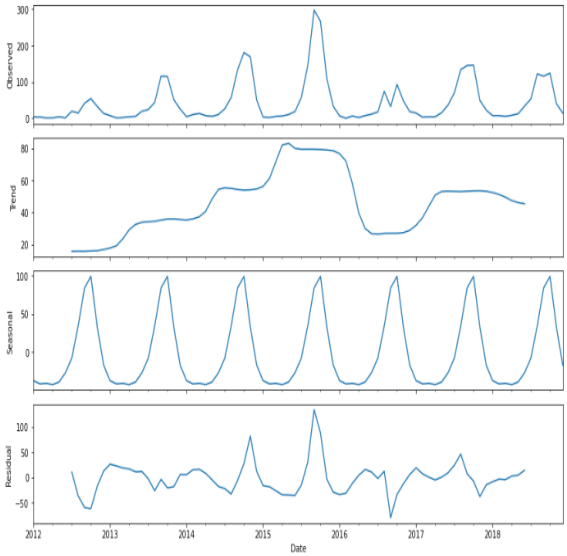

Figure 3: Decomposition of Dengue Data into trend, seasonal and random noise.

ARIMA(p,d,q) model consists of three components indicating Auto-Regression(p), Integration(d) and Moving Average(q) parameters. It can be given using a standard form as shown in eq(1)

$$y_t = \xi + \phi_1 y_{t-1} + \phi_2 y_{t-2} \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} \ldots - \theta_q e_{t-q} \quad (1)$$

The ‘d’ parameter will tell us the order of differencing so that the time series can be made stationary. The ‘p’ and ‘q’ parameters can be determined by using the auto-correlation function (ACF) and partial auto-correlation function (PACF) plots.

The addition of seasonality to the above model gives us the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. A SARIMA (p,d,q) x (P,D,Q,s) requires the following parameters to be identified:

- p and Seasonal P, which gives the number of lags in the time series i.e AR term
- d and Seasonal D, which gives the differencing required for the series to be made stationary
- q and Seasonal Q, which gives the number of lags of forecast errors i.e the MA terms
- s, which gives the length of seasonal pattern

An analysis of ACF and PACF plots will help us identify the parameter values. The above methods are univariate methods which makes use of the past data

only to make predictions about the future. Multivariate methods make use of correlated or influencing factors along with the past data to make predictions. Examples of multivariate methods are SARIMA for exogenous variables (SARIMAX) and vector autoregression (VAR) model.

The VAR model works by identifying relations between the current values of the observed data with previous values of itself and previous values of other variables.

Machine learning techniques can be used to capture the patterns in the data and use it to predict the future values. Techniques like regressor decision tree(DT), random forest (RF), and Gradient boosting tree(GBT) can be used.

These models were built in the big data SPARK environment, which allows execution of data in a distributed and parallel manner. For present study, a model was developed and validated by dividing the data into two subsets i.e. from 2012 to 2018 was used for training the model while data from 2019 was used for testing and validating the fitted model.

A. Methodology Used

In the methodology adopted consists of the following steps in building the predictive model:

1. The data is represented in the format specified in [15] and is analyzed to check for stationarity using the Dickey-Fullers Test
2. The model parameters are calculated using the ACF and PACF plots or by extensive search.
3. The model coefficients are calculated using the multivariate modified VAR model using the distributed approach.
4. The model is used to forecast the cases
5. Accuracy of the model is calculated.

After the data is pre-processed, it is converted into the TemporalRDD[15] format to be used in the big data environment. The stationarity of the data series is obtained and checked for stationarity. The parameter is identified by analyzing the ACF and PACF plots or by performing a extensive search on all parameter combination. Once the model parameters are identified we can determine the coefficients using the distributed approach.

In this study, we proposed a multivariate modified vector autoregression model and compared its results with three types of models - a univariate model, a multivariate model and machine learning model. The above methodology is used for multivariate model. Whereas for univariate and machine learning models, built in function of python and Pyspark are used. The multivariate model is implemented on the big data environment.

B. Model Comparison and Evaluation

Results of various models were compared using metrics like root mean squared error (RMSE), and Akaike Information Criteria (AIC). Using these metrics, best fit model can be identified for making predictions.

Root mean squared error is one of the most commonly used metrics for selection of model and is given by equation-2.

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}(y_j - \bar{y}_j)^2}{n}} \qquad (2)$$

Broadly, the RMSE provide the standard deviation of the prediction errors and offers us with a good evaluation of how accurately the model predicts the values.

The Akaike information can be obtained from the maximum-likelihood of the model and the number of k parameters used to reach the likelihood and is given by equation-3.

$$\text{AIC} = 2k - n(\text{log-likelihood}) \qquad (3)$$

A lower score of AIC is desired to select the better model as a lower score model would minimize the information loss.

Suitability of best-fit model parameters was selected by examining the auto-correlation function (ACF) and partial auto-correlation function (PACF) plots.

V. RESULTS AND DISCUSSION

The distribution of Dengue in the study area from 2012 to 2018 is emphasized in Figure-4. The data from 2012 to 2018 was used for training purpose, while the data of year 2019 was used for testing and validation.

The order of differencing in the series can be identified by applying the Augmented Dickey-Fuller test. The test gave the p-value lesser than 0.05 specifying that we can reject the null hypothesis and the series is stationary, thus giving us the d value as 0.

Next step is to identify the ‘p’ and ‘q’ values. An analysis of the auto-correlation function (ACF) and partial auto-correlation function (PACF) plots can help us identify these values.

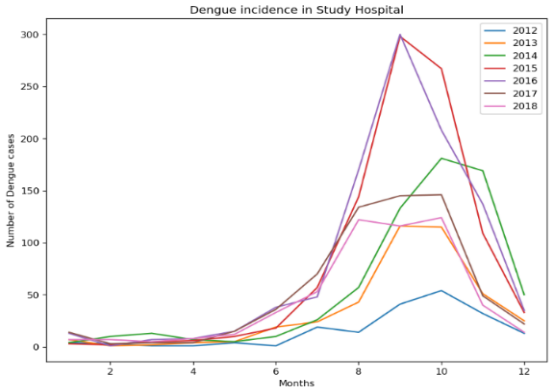

Figure 4: Plot of Dengue from 2012 to 2018 in the study hospital.

The ACF and PACF plots for the data is shown in Figure-5. The assessment of the PACF plot allows us to determine the number of AR terms (p) that are needed to explain the autocorrelation pattern in a time series. The partial autocorrelation is observed to be significant at lag 2 and there are no significant values at any higher order lags. Hence the value of p can be considered as 2. The analysis of ACF plot would allow us to identify the MA terms(q). The ACF plot shows multiple positive and negative peaks even after a differencing is performed. Hence, we can take q value to be 0

It is observed that there are significant spikes in the series at regular intervals, indicating seasonality in the data. The spikes repeat after every 12 intervals indicating a yearly pattern or an 's-value' of 12.

Another approach to identify these parameters values is to use a grid approach, which performs an extensive search by having multiple combinations of the parameter values. The best combinations is the one which will give us the lowest Akaike Information Criterion (AIC).

Based on the above analysis we tested ARIMA(1,0,1), ARIMA(2,0,1), ARIMA(3,2,1) models. The seasonal model SARIMA(2,0,0)x(0,1,1,12) is chosen based on the analysis. The plot of the best model amongst the ARIMA is given in figure-6(a), whereas the plot of SARIMA model is given in figure 6-(b).

VI. MODEL SELECTION, DIAGNOSTICS AND FORECAST ERRORS

During the procedure of selecting the best model for forecasting, we applied and tested in total nine models, of which five are traditional time series forecasting models and four are machine learning models. The models were tested and validated on the test data as mentioned. The accuracy measures that were used are root mean square error (RMSE), Akaike Information Criterion (AIC)

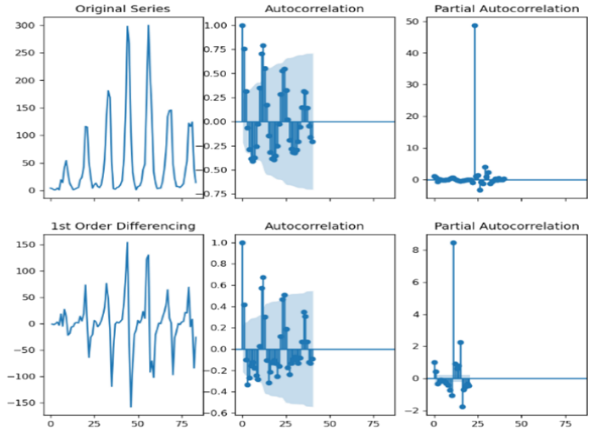

Figure 5: ACF and PACF plots of Dengue Data

Traditional time series analysis models were executed for forecasting of dengue cases, results of which are listed in Table-II. The SARIMAX model gives us the best results for univariate analysis with a root mean square error of 20.79. Incorporating the climatic factor, VAR model gave the least AIC value with a RMSE of 27.32.

Table II: Comparison of Prediction models

Model	AIC	RMSE
ARIMA(1,0, 1)	864.58	61.34
ARIMA(2, 0, 1)	861.35	61.34
ARIMA(3, 2, 1)	853.62	61.27
SARIMAX (2,0,0)x(0,1,[1],12)	715.38	20.79
VAR(2)	5.89	27.32

Various big data machine learning algorithms like decision tree(DT), random forest (RF), and Gradient boosting tree(GBT) were used to build the model using only the incidence data and then using the climatic data.

The results of metrics for machine learning algorithms are shown in Table-III. The results show that the random forest regression gives the least RMSE value forming the best machine learning model for the given data.

Table III: Comparison of Prediction models

Algorithm	RMSE
Decision tree	78.20
Random forest regression	53.11
Gradient-boosted tree	78.19

To identify the effects of the climatic parameter each of the model was trained by incorporating individual features initially and then combining all the features.

The results are tabulated in Table-IV. Individual parameters were used to build the model and the results are shown in columns b-e of Table-IV. Incorporation of humidity and temperature parameters yielded us with lower RMSE values indicating that they have a better effect. The results show that incorporation of all climatic parameters (as shown in column-a of Table-IV) improves the prediction in case of Random Forest model and shows no improvement in other models.

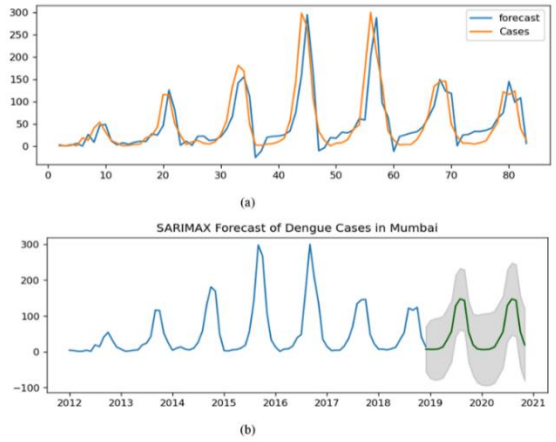

Figure 6: Prediction plot using a) ARIMA(3,2,1) b) Seasonal ARIMA

Table IV: Root mean squared error for machine learning models considering the climatic parameters

Algorithms	a	b	c	d	e
	RMSE_ All	RMSE_ Temp	RMSE_ Humidity	RMSE_ Wind	RMSE_ Pressure
DT	47.08	45.66	43.70	46.31	45.49
RF	43.03	45.65	43.61	46.27	45.63
GBT	46.87	45.66	43.69	46.37	45.44

VII. CONCLUSIONS

The cross-sectional study about the epidemiological analysis of Dengue infection based on lab-data was conducted in Mumbai over a span of 8 years from 2012 to 2019. The number of dengue cases increased every year till 2016 and then decreased. The decreasing patterns was not observed continuously, and this variation was modeled into the current study.

The current study identified the association between weather and dengue for the given region and incorporates it to forecast future incidence cases using traditional time series and machine learning approaches. Of the total around 4000 cases, strong correlation was found between incidence cases and climatic parameter.

This study confirms the previous studied relations with climatic factors like temperature, pressure, and humidity. The study demonstrated a relation between temperature, humidity, pressure, and dengue cases.

One of the major strengths of our study is that it is a comprehensive regional analysis which included climatic factors and studied its effects on the incidence of Dengue. These factors were incorporated for building a forecasting model. The time series analysis techniques SARIMA model gave significant results with RMSE as 20.79 in univariate analysis while VAR model gave results with RMSE of 27.32. This confirms that confounding factors like temperature, humidity significantly influence the prediction of dengue incidences. They also showed that these factors when incorporated into forecasting model significantly improves AIC values to 5.89 resulting in a better model for prediction.

VIII.ACKNOWLEDGMENT

Authors are grateful to the ethics committee of KEM hospital for providing the data necessary for the study under the protocol EC/OA-42/2018

REFERENCE

[1] https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue accessed on 21st April 2021

[2] https://www.cdc.gov/dengue/index.html accessed on 21st April 2021

[3] https://www.who.int/denguecontrol/epidemiology/en/ accessed on 21st April 2021

[4] https://idsp.nic.in/index1.php?lang=1&level=1&sublinkid=5985&lid=3925 accessed on 21st April 2021

[5] Shepard, Donald S et al. "Economic and disease burden of dengue illness in India." The American journal of tropical medicine and hygiene vol. 91,6 (2014): 1235-1242. doi:10.4269/ajtmh.14-0002

[6] Hopp MJ and Foley JA. Global-scale relationships between climate and the dengue fever vector, Aedesaegypti. Clim Chang 2001;48:441–463

[7] Arcari P, Tapper N, Pfueller S. Regional variability in relationships between climate and dengue/DHF in Indonesia. Singapore. J Trop Geo 2007; 28: 251–72.

[8] Mutheneni SR, Morse AP, Caminade C, Upadhyayula SM. Dengue burden in India: recent trends and importance of climatic parameters. Emerging Microbes & Infections. 2017;6(8):e70-. doi:10.1038/emi.2017.57.

[9] Bhatnagar S, Lal V, Gupta SD, Gupta OP. Forecasting incidence of dengue in Rajasthan, using time series analyses. Indian J Public Health. 2012 Oct-Dec;56(4):281-5. doi: 10.4103/0019-557X.106415. PMID: 23354138.

[10] Gharbi M, Quenel P, Gustave J, et al. Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. BMC Infect Dis. 2011;11:166. Published 2011 Jun 9. doi:10.1186/1471-2334-11-166

[11] Buczak, A.L., Koshute, P.T., Babin, S.M. et al. A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. BMC Med Inform Decis Mak 12, 124 (2012). https://doi.org/10.1186/1472-6947-12-124

[12] Lei Li, Farzad Noorian, Duncan J. M. Moss, Philip Heng Wai Leong, Rolling window time series prediction using MapReduce, IEEE International Conference on Information Reuse and Integration (IRI) IRI 2014: 757-764

[13] Jain, R., Sontisirikit, S., Iamsirithaworn, S. et al. Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data. BMC Infect Dis 19, 272 (2019). https://doi.org/10.1186/s12879-019-3874-x

[14] Hii YL, Zhu H, Ng N, Ng LC, Rocklöv J. Forecast of dengue incidence using temperature and rainfall. PLoS Negl Trop Dis. 2012;6(11):e1908. doi: 10.1371/journal.pntd.0001908. Epub 2012 Nov 29. PMID: 23209852; PMCID: PMC3510154.

[15] Shi Y, Liu X, Kok SY, Rajarethinam J, Liang S, Yap G, Chong CS, Lee KS, Tan SS, Chin CK, Lo A, Kong W, Ng LC, Cook AR. Three-Month Real-Time Dengue Forecast Models: An Early Warning System for Outbreak Alerts and Policy Decision Support in Singapore. Environ Health Perspect. 2016 Sep;124(9):1369-75. doi: 10.1289/ehp.1509981. Epub 2015 Dec 11. PMID: 26662617; PMCID: PMC5010413.

[16] Withanage GP, Viswakula SD, Nilmini Silva Gunawardena YI, Hapugoda MD. A forecasting model for dengue incidence in the District of Gampaha, Sri Lanka. Parasit Vectors. 2018 Apr 24;11(1):262. doi: 10.1186/s13071-018-2828-2. PMID: 29690906; PMCID: PMC5916713.

[17] McGough SF, Clemente L, Kutz JN, Santillana M. 2021 A dynamic, ensemble learning approach to forecast dengue fever epidemic years in Brazil using weather and population susceptibility cycles. J. R. Soc. Interface 18: 20201006. https://doi.org/10.1098/rsif.2020.1006

[18] Ganeshkumar P, Murhekar MV, Poornima V, Saravanakumar V, Sukumaran K, Anandaselvasankar A, et al. (2018) Dengue infection in India: A systematic review and meta-analysis. PLoS Negl Trop Dis 12(7): e0006618. https://doi.org/10.1371/journal.pntd.0006618

[19] https://dm.mcgm.gov.in/ward-directory

[20] Bharambe A., Kalbande D. (2022) Self-organizing Data Processing for Time Series Using SPARK. In: Shakya S., Bestak R., Palanisamy R., Kamel K.A. (eds) Mobile Computing and Sustainable Informatics. Lecture Notes on Data Engineering and Communications Technologies, vol 68. Springer, Singapore. https://doi.org/10.1007/978-981-16-1866-6_17

[21] www.worldweatheronline.com through API