Indonesian Journal of Electrical Engineering and Computer Science
Vol. 18, No. 1, April 2020, pp. 502~507
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v18.i1.pp502-507

Investigating dengue outbreak in Tamil Nadu, India

N. Komal Kumar[1], R. Lakshmi Tulasi[2], D. Vigneswari[3]
[1]Department of Computer Science and Engineering, St. Peter's Institute of Higher Education and Research, India
[2]Department of Computer Science and Engineering, RVR & JC College of Engineering, India
[3]Department of Information Technology, KCG College of Technology, Karapakkam, India

Article Info

Article history:

Received Jul 29, 2019
Revised Oct 13, 2019
Accepted Nov 2, 2019

Keywords:

Dengue
India
Outbreak
Preventive
Regression
Tamil Nadu

ABSTRACT

Dengue has been indigenous to India in last decade. There was a major outbreak in the state of Tamil Nadu in 2017. Here, we investigate the dengue outbreak in parts of Tamil Nadu, India. Dengue case data were obtained from the hospital records in the Chennai district of Tamil Nadu. The data were analyzed using statistical approaches such as correlation and regression. The result shows that the dengue outbreak in Tamil Nadu during 2017 was due to the population, water stagnation, and sewage, whereas the human activity weren't the cause of the dengue outbreak which caused 65 deaths. Male constitutes 54.71% whereas female accounted for 45.29% of dengue incidence in Tamil Nadu, majority deaths were children aged less than 10 years due to the outbreak of Dengue Hemorrhagic Fever (DHF). This investigation was evaluated using mathematical regressions, Geographically Weighted Regression (GWR) regression outperformed Ordinary Least Square (OLS) regression model in detecting dengue incidence. This investigation can be strengthened by implementing a surveillance system in parts of Tamil Nadu before an outbreak.

Corresponding Author:

N. Komal Kumar,
Department of Computer Science and Engineering,
St. Peter's Institute of Higher Education and Research, Avadi, Chennai, India.
Email: komalkumarnapa@gmail.com

1. INTRODUCTION

Dengue is a mosquito-borne disease which registers nearly 50 million cases annually [1]. This disease is spread by Aedes type mosquitoes. Dengue virus family contains three serotypes namely DENV 1 (Dengue Fever), DENV 2 (Dengue Hemorrhagic Fever), DENV 3 (Dengue Shock Syndrome). Dengue affects entire globe, nearly 4 billion people are at risk by DENV infections [2], and major affected regions are from the developing countries. It causes threat to the tropical regions where the temperature is abnormal. Rapid urbanization [3], global warming, environmental factors, human activities are the factors for its spread. Dengue virus fever is a symptomless, but it can lead from fever to shock syndrome [4]. Dengue is transmitted from the vectors namely Aedesaegypti and Aedesalbopictus. They are the black colored mosquitoes with white stripes and they are approximately 5 mm in size. Their feeding habits includes day biter, mainly feeds on human beings in domestic and peridomestic situations, with continuous biting. They rest on domestic and peridomestic areas, dark corners of room, curtains, unwashed clothes, umbrella etc., they breed in manmade containers or containers which contains even small quantities of water. These vectors can live without water even for a year [5]. The period of communicability of the affected person becomes infective to mosquitoes of 6 to 12 hours before the onset on the disease and it remains up to 3 to 5 days. Favored places of breeding includes Desert coolers, drums, jars, pots, buckets, flower vases, cisterns, bottles, tyre, tins, refrigerator dip pans, cemetery urns, bamboo stumps, coconut shells, tree holes, where rain water is stagnated. These Dengue Hemorrhagic Fever vector viruses are mainly found in the tropical regions of Asia, Africa, Middle East, Australia and Americas [6]. Other factor for these vector creating includes rainfall patterns and temperature change due to global warming. High temperature reduces the breading of such vectors [7]. Precipitation and rainfall also contribute to the growth to the larvae, but severe rainfall tends to kill big mosquitoes, causing depletion of dengue causing vectors. Tamil Nadu has experienced a severe dengue outbreak in 2012 [8] since then it experienced average number of deaths till 2016. Total dengue cases registered thought India in the year 2015 was 99913, causing 220 deaths which increased to 325 deaths in 2017, nearly 65 deaths in Tamil Nadu, mainly children in the age <10 years due to the outbreak of Dengue Hemorrhagic Fever. Recent work in [9] investigated the DENV serotypes and analyzed the patterns using machine learning classifier for detecting DENV serotypes; the hybrid algorithm MSO-MLP outperformed all classifiers under the analysis. Several works were carried out for investigating infectious disease such as hepatitis [10] and predicting non small cell lung cancer [11] using machine learning classifiers. Hybrid machine learning classifiers [12] can be more effective in classifying patients affected with infectious disease [13]. Geographical Information Systems (GIS) tools can be used to find the affected area which helps the researcher to identify its major factors involved in it. Several studies [14-18] reported the use of GIS for identifying heat maps of the affected areas. Spatial rings maps can be used to show the spatio-temporal diffusions in certain regions. Since 2012, Tamil Nadu has experienced a severe dengue outbreak in the year 2017. To our knowledge no other research work has been carried out to investigate the spatio-temporal pattern in parts of Tamil Nadu. Table 1 shows the number of dengue cases recorded in Tamil Nadu and throughout India, the scenario depicts that the cases registered follows a steady increase in years, Table 2 shows number of deaths due to dengue outbreak, 2017 accounted a terrible 325 deaths in India, which accounted 65 deaths only in Tamil Nadu. There seems to be a certain reason for this severe outbreak and deaths, this investigation focuses on the factors involved in dengue outbreak. This helps the health organizations and national agencies to provide prior indication about the cause and effect of dengue virus onto the population.

Table 1. Comparison of Tamil Nadu and India no of Dengue Cases Recorded from 2014 to 2017

Year/ Region	2014	2015	2016	2017
India	40571	99913	111896	188401
Tamil Nadu	2804	4535	2531	23294

From Table 1, it can be seen that there is a steady increase in number of dengue cases registered; number of cases registered in 2015 was 40% more than in 2014. In 2017 the cases registered were 188401 which is greater than the cases registered in 2016, the majority of the cases were children <10 years.

Table 2. No of Deaths Due to Dengue Outbreak from 2014 to 2017 in Tamil Nadu and India

Year/ Region	2014	2015	2016	2017
India	137	220	227	325
Tamil Nadu	3	12	5	65

From Table 2, it can be seen that the no of deaths in 2014 were recorded to be 7 in Tamil Nadu out of 137 deaths in India. The deaths slowly increased to 12 out of 220 in 2015. No of deaths increased drastically from 5 to 65 from 2016 to 2017 in Tamil Nadu. This paper focuses on the parameters or the activities which caused major outbreak in Tamil Nadu in 2017.

2. RESEARCH METHOD

This investigation was carried out in Tamil Nadu, India which covers a geographical area of 1,30,060 sq km. Statistical data such as demographics were retrieved from “National Health Profile” database [19] from Central Bureau of Health Intelligence (CBHI). Dengue case data has been retrieved from National Vector Borne Disease Control Program (NVBDCP) [20] and Government of India (GOI) websites. These data includes age, gender, and location of each dengue case. With these data, incidences were calculated and patterns of dengue diffusion [21] were plotted. Analysis of the spatio-temporal diffusion pattern was carried out by SaTScan. Maximum window size of population at risk was set to 50%. Monto Carlo simulation was used to get the p-value through the hypothesis testing. Geographically Weighted Regression (GWR) and Ordinary Least Square (OLS) [22-23] were the two methods used in this investigation. GWR model was applied on the population density, human activities, and sewage parameters. OLS has been used to find the relationship among dependent variables and independent variables on dengue occurrences. An Akaike Information Criterion was used to test the two model’s performance and also Variance inflation factor was used to test the co-linearity of GWR and OLS models. Table 3 represents the demographics of the dengue cases in Tamil Nadu. The most affected ones are the children in the age group of <10 years. Total number of cases registered in 2017 was 23294, out of which 12743 were male and 10551 were female. No of cases registered were categorized as age groups. 8421 cases were registered with age less than 10, 3260 with age 10 to 20, 2980 with age 21 to 30, 1691 with age 31 to 40, 245 with age 41 to 50, 563 with age 51 to 60, 281 with age 61 to 70 and 132 cases with age greated than 70. There were 5721 missing values, the patterns in the missing values were analyzed and imputed using normal distribution.

Table 3. Demographics of Dengue Cases in Tamil Nadu, 2017

Variables	Number of Cases	Percentage (%)
Gender		
Male	12743	54.71
Female	10551	45.29
Age (in years)		
<10	8421	36.15
10 to 20	3260	14.00
21 to 30	2980	12.79
31 to 40	1691	7.26
41 to 50	245	1.05
51 to 60	563	2.42
61 to 70	281	1.21
>70	132	0.57
Missing	5721	24.56
Total	23294	100

3. RESULTS AND ANALYSIS

In this investigation, a total of 23294 cases were reported in Tamil Nadu, India in 2017, the outbreak started in August, 2017, reached its peak in September, and decreased in December. The majority of the cases were <10 years due to the outbreak of Dengue Hemorrhagic Fever (DHF). The investigation of dengue outbreak falls into various steps, number of cases registered, deaths and various factors causing dengue were obtained and stored in a local file (dataset), obtained data is preprocessed using impute function for removing missing values. Later, the datset is subjected to Ordinary Least Square (OLS) and Geographically Weighted Regression (GWR) statistical approaches and the results obtained are tabulated in Table 4 and Table 5 respectively.

Table 4 shows the summary of the Ordinary Least Square (OLS) regression model, the estimates of various parameters causing dengue were recorded with its standard error. Depending on the p-value certain parameters can be removed. In this investigation, population falls below 0.05 (common alpha value), so it can be removed, other parameters are statistically significant and VIF value is used to check the multicollinearity among variables.

Table 4. Summary of the Ordinary Least Square (OLS) regression model

Parameters	Estimates	Stderr.	*p*-value	VIF
Intercept	16.25			
Population	0.06	0.03	0.0002	1.42
Human activities	-0.006	0.04	0.86	8.61
Water stagnation	0.04	0.08	0.75	8.72
Sewage	0.12	1.12	0.96	9.84
R-squared	0.48			
Akaike Information Criteria	896.21			

The regression framework is shown in Figure 1. Four risk factors were considered for OLS and GWR regression models. In case of OLS regression model population, water stagnation and sewage were positively associated with the dengue incidence, where as human activities are negatively associated, this proves that human activities doesn't cause more effect on the dengue outbreak. By comparing OLS with GWR regression model, the GWR performance was more in the predicting dengue out break (Akaike Information Criteria (AIC) 896.21 Vs. 800.98), the population showed high correlation [24] in dengue outbreak, whereas water stagnation, sewage, human activities were to be negatively correlated. The diffusion patterns of the dengue outbreak of India were shown in Figure 2. The diffusion pattern seems to be randomly distributed causing havoc. As per the health records from a National health agency, highest number of dengue cases registered next to Tamil Nadu was Kerala which shares its boundaries with Tamil Nadu, accounted 19994 cases and 37 deaths in 2017, which was less than 2016 where 6083 cases have been registered with only 8 deaths. Next to Kerala was the Karnataka state with 17844 dengue cases registered with 10 deaths. This shows even the cases registered in these two states were higher, the deaths encountered were less compared to Tamil Nadu, due to the preventive action taken by the officials during and after math. This shows the negligibility of health officials in Tamil Nadu, India.

Table 5 shows the summary of Geographically Weighted Regression model. Population, human activities, water stagnation, sewage and dengue incidence were provided as input features and explanatory variables.

Table 6 shows the number of dengue cases registered and number of deaths in 2017, in Tamil Nadu, Kerala and Karnataka. Number of cases and deaths registered in Tamil Nadu were 23294 and 65. Kerala accounted 19994 cases and 37 deaths and Karnataka accounted 17844 cases and 10 deaths. It can be seen that the cases registered and number of deaths due to dengue were greater in Tamil Nadu than the other southern states under comparison.

Table 5. Summary of the Geographically Weighted Regression (GWR) Regression Model

Parameters	Minimum	25% quartile	50% quartile	75% quartile	Maximum
Intercept	46.27	12.45	36.98	73.98	207.76
Population	0.01	0.02	0.04	0.06	0.09
Human activities	-0.15	-0.19	-0.08	-0.04	-0.02
Water stagnation	-0.43	-0.34	-0.26	-0.21	0.06
Sewage	-0.54	-0.32	0.03	0.07	0.1
Condition number[a]	16.26	18.19	20.98	30.98	39.98
R-squared	0.56				
Akaike Information Criteria	800.98				

[a]If the condition number is greater than 30, it represents the collinearatity

Table 6. Number of Dengue Cases and Deaths Registered in Tamil Nadu, Kerala and Karnataka, 2017

Affected State	2017	
	Cases	Deaths
Tamil Nadu	23294	65
Kerala	19994	37
Karnataka	17844	10

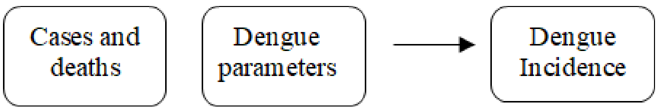

Figure 1. Regression framework

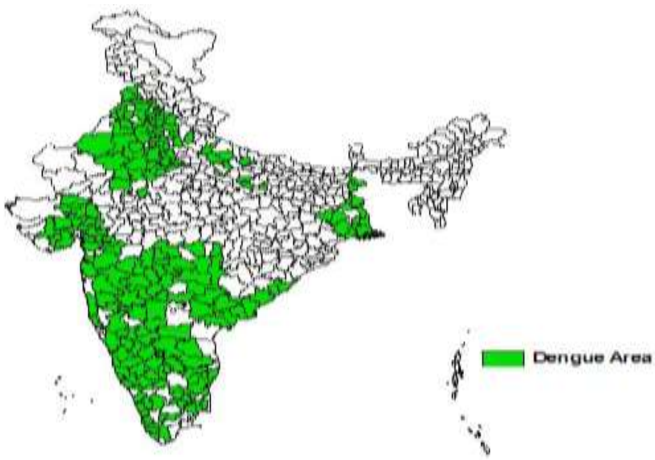

Figure 2. Diffusion pattern of Dengue in India, 2017

4. CONCLUSION

This investigation attempts to investigate the outbreak of dengue in 2017 in Tamil Nadu, India. This is the first study to analyze the dengue pattern in Tamil Nadu to understand the factors involved in the dengue outbreak. These investigations help the health organization to plan accordingly and take safety measures in case of severe outbreak of dengue virus. It is worth noting that in gender distribution, the dengue incidence was found to be more in male than female, this is due to the fact that female restrict their food habits unhealthy surroundings. By contrast male work from dawn to dusk and exposed to such surroundings than female. A previous study in [25] introduced the dengue incidence in the Swat district, Pakistan. The study looked to the space-time diffusion patterns of dengue outbreak by analyzing several factors [26] such as distant to river, population density, elevation. The study revealed that favorable factors [27] led to the outbreak in 2013. The dengue outbreak totally occurred in the coastal regions of India, especially in eastern and western ghats, which gets a heavy rainfall during monsoon season. Rainfall in 2015 was the exact reason for the huge increase in the number of cases registered in India. Considering the above investigation, there is no doubt that dengue is severe health problem that costs life. It is necessary for the health organization and health sectors to take preventive actions to reduce the deaths in near future [28].

ACKNOWLEDGEMENTS

The authors are grateful to the management of St. Peter's Institute of Higher Education and Research, Chennai, India. RVR & JC College of Engineering, Guntur, India and KCG College of Technology, Karapakkam, Chennai, India that greatly assists research and development.

REFERENCES

[1] Normile D, "Tropical medicine. Surprising new dengue virus throws a spanner in disease control efforts,"*Science*, vol. 342, pp. 415, 2013.
[2] Guzman MG, Kouri G, "Dengue: an update," *Lancet Infect Dis*, vol. 2, pp. 33-42, 2002.
[3] Wu PC, Lay JG, Guo HR, Lin CY, Lung SC, Su HJ, "Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan," *Sci Total Environ*, vol. 407, pp. 2224-33, 2009.
[4] Senaka Rajapakse, "Dengue Shock," *J Emerg Trauma Shock*, vol. 4, pp. 120-127, 2011.
[5] Bayoh MN, Walker ED, Kosgei J, Ombok M, Olang GB, Githeko AK, Killeen GF, Otieno P, Desai M, Lobo NF, et al, "Parasit Vectors," *Epub*, vol. 7, pp. 380, 2014.
[6] Gyawali N, Bradbury RS, Taylor-Robinson AW, "The epidemiology of dengue infection: harnessing past experience and current knowledge to support implementation of future control strategies," *J Vector Borne Dis*, vol. 53, pp. 293-304, 2016.
[7] Craig, Marlies H., R. W. Snow, and David le Sueur, "A climate-based distribution model of malaria transmission in sub-Saharan Africa," *Parasitology today*, vol. 15, pp. 105-111, 1999.
[8] N. Komal Kumar, D. Vigneswari, B A S Roopa Devi, "MSO-MLP diagnostic approach for detecting DENV serotypes," *International J. of Pure and Applied Mathematics*, vol. 118, pp. 1-6, 2018.
[9] National Health Profile 2017. Available from: http://www.indiaenvironmentportal.org.in/files/file/NHP_2017-1.pdf
[10] N Komal Kumar, D Vigneswari, "Hepatitis- Infectious Disease Prediction using Classification Algorithms," *Research J. Pharm. and Tech,* vol. 12, pp. 3720-3725, 2019.
[11] N. Komal Kumar, D. Vigneswari, M. Kavya, K. Ramya, T. Lakshmi Druthi, "Predicting Non-Small Cell Lung Cancer: A Machine Learning Paradigm," *Journal of Computational and Theoretical Nanoscience*, vol. 15, pp. 2055-2058, 2018.
[12] N. Komal Kumar, R. Lakshmi Tulasi, D. Vigneswari, "An ensemble multi-model technique for predicting chronic kidney disease," *International Journal of Electrical and Computer Engineering*, vol. 9, pp. 1321-1326, 2019.
[13] D. Vigneswari, N. Komal Kumar, V. Ganesh Raj, A. Gugan, S.R. Vikash, "*Machine learning tree classifiers in predicting diabetes mellitus,* "*International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India*, pp. 84-87, 2019.
[14] Alzahrani AG, Al Mazroa MA, Alrabeah AM, Ibrahim AM, Mokdad AH, Mem-ish ZA, "Geographical distribution and spatio-temporal patterns of dengue cases in Jeddah Governorate from 2006-2008," *Trans R Soc Trop Med Hyg*, vol. 107, pp. 23-29, 2013.
[15] Chen CC, Teng YC, Lin BC, Fan IC, Chan TC, "Online platform for applying space-time scan statistics for prospectively detecting emerging hot spots of dengue fever," *Int J Health Geogr*, vol. 15, pp. 43, 2016.
[16] Langford, Mitchel, Gary Higgs, and David J. Dallimore. "Investigating spatial variations in access to childcare provision using network-based Geographic Information System models." *Social Policy & Administration* 53.5: 661-677, 2019.
[17] Lee, Minhyun, and Taehoon Hong. "Hybrid agent-based modeling of rooftop solar photovoltaic adoption by integrating the geographic information system and data mining technique." *Energy conversion and management* 183:266-279, 2019.
[18] Lu, Bin, et al. "Geographic information system algorithms to locate prospective sites for pumped hydro energy storage." *Applied energy*, 300-312, 2018.

[19] www.cbhidghs.nic.in
[20] https://www.nhp.gov.in/national-vector-borne-disease-control-programme_pg
[21] Chan T-C, Wang C-M, Lee Y-M, "Looking at temporal changes-use this pythontool for creating ring maps," *ArcUser*, pp. 1-3, 2013.
[22] Hosmer DW, Lemesbow S, "Goodness of fit tests for the multiple logistic regression model. Communications in statistics," *Theory and Methods*, vol. 9, pp. 1043-69, 1980.
[23] Y. Peranginangin, A. Alamsyah, "Multiple regression to analyse social graph of brand awareness," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 15, pp. 336-340, 2017.
[24] A. Khamis, Y. Xu, A. Mohamed, "Comparative study in determining features extraction for islanding detection using data mining technique: Correlation and coefficient analysis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7. pp. 1112-1124, 2017.
[25] SulemanAtique, Ta-Chien Chan, Chien-Chou Chen, Chien-Yeh Hsu, SomiaIqtidar,Valérie R. Louis, Syed A. Shabbir, Ting-Wu Chuang, "Investigating spatio-temporal distribution and diffusion patterns of the dengue outbreak in Swat, Pakistan," *Journal of Infection and public health*, vol. 11, pp. 550-557, 2018.
[26] K. Lavangnananda, S. Chattanachot, "*Study of Discretization Methods in Classification*," in IEEE Press, Proc. International Conference on Knowledge and Smart Technology, pp.50-55, 2017.
[27] M. Othman, S. A. Mohamed, M. H. A. Abdullah, M. M. Yusof, and R. Mohamed, "A framework to cluster temporal data using personalised modelling approach," *Adv. Intell. Syst. Comput.*, vol. 700, pp. 181-190, 2018.
[28] N. Kasabov et al., "Evolving spiking neural networks for personalised modelling, classification and prediction of spatio-temporal patterns with a case study on stroke," *Neurocomputing*, vol. 134, pp. 269-279, 2014.

BIOGRAPHIES OF AUTHORS

N. Komal Kumar is currently working as an Assistant Professor in the Department of Computer Science and Engineering, St. Peter's Institute of Higher Education and Research, Avadi, Chennai, India. His research interests include Machine Learning, Data Mining, and Cloud Computing.

Dr. R. Lakshmi Tulasi is currently working as an Associate Professor in the Department of Computer Science and Engineering, RVR & JC College of Engineering, Guntur, Andhra Pradesh, India. Her research interests include Information Retrieval and Semantic web.

D. Vigneswari is currently working as an Assistant Professor in the Department of Information Technology, KCG College of Technology, Karapakkam, Chennai, India. Her research interests include Data Mining and Data Analytics.