**Introduction**

Dengue is a vector-borne viral disease which is one of the major causes of public health problem in India, and its control is often the major challenges of municipal bodies in the country, especially in West Bengal. The previous outbreaks of the disease can be used to forecast the future occurrence and burden, so that authorities may optimize the available resources in order to contain and minimize the impact. **Materials and Methods:** Weekly disease outbreak data were extracted from Integrated Disease Surveillance Programme website and arranged as monthly data. Mann-Kendall test was used to determine the significance of the disease trends in various districts of Gangetic West Bengal. Time series analysis was done by using Seasonal ARIMA method to predict the number of Dengue outbreak cases for the year 2020. **Results:** Murshidabad was the only district of Gangetic West Bengal that had a significant upward Dengue cases outbreak trend. Nadia had a downward trend but it was not statistically significant. Model SARIMA (1,0,0) (1,0,0)_{12} was chosen to forecast the Dengue outbreak cases which showed that the cases might start from the month of June, peak in August and wane off in October 2020. However, this prediction was not significant. **Conclusion:** Gangetic West Bengal might experience similar dengue cases as the previous year, but their numbers would be low. Only the district of Murshidabad would have upward trend. Knowledge in advance about periods of disease occurrence may enable health authorities to initiate control measures during the start of the outbreak season.

**Keywords:** Dengue outbreaks, MAKESENS application, seasonal ARIMA, time series analysis, trend analysis

---

**Background and Aims:** Dengue is a vector-borne viral disease which is one of the major causes of public health problem in India, and its control is often the major challenges of municipal bodies in the country, especially in West Bengal. The previous outbreaks of the disease can be used to forecast the future occurrence and burden, so that authorities may optimize the available resources in order to contain and minimize the impact. **Materials and Methods:** Weekly disease outbreak data were extracted from Integrated Disease Surveillance Programme website and arranged as monthly data. Mann-Kendall test was used to determine the significance of the disease trends in various districts of Gangetic West Bengal. Time series analysis was done by using Seasonal ARIMA method to predict the number of Dengue outbreak cases for the year 2020. **Results:** Murshidabad was the only district of Gangetic West Bengal that had a significant upward Dengue cases outbreak trend. Nadia had a downward trend but it was not statistically significant. Model SARIMA (1,0,0) (1,0,0)_{12} was chosen to forecast the Dengue outbreak cases which showed that the cases might start from the month of June, peak in August and wane off in October 2020. However, this prediction was not significant. **Conclusion:** Gangetic West Bengal might experience similar dengue cases as the previous year, but their numbers would be low. Only the district of Murshidabad would have upward trend. Knowledge in advance about periods of disease occurrence may enable health authorities to initiate control measures during the start of the outbreak season.

**Keywords:** Dengue outbreaks, MAKESENS application, seasonal ARIMA, time series analysis, trend analysis

---

**Address for correspondence:** Dr. Jitendra Majhi, Department of CM and FM, AIIMS, Kalyani, West Bengal, India. B-15/5, Mangal Apartment, Lake Road, Kalyani, Nadia - 741 235, West Bengal, India. E-mail: drjitendrascb@gmail.com

Received: 08-05-2020
Accepted: 02-07-2020
Revised: 11-06-2020
Published: 30-11-2020

**Access this article online**

Quick Response Code:  
Website: www.jfmpc.com  
DOI: 10.4103/jfmpc.jfmpc_800_20

**How to cite this article:** Majhi J, Singh R, Yadav V, Garg V, Sengupta P, Atul PK, et al. Dynamics of dengue outbreaks in gangetic West Bengal: A trend and time series analysis. J Family Med Prim Care 2020;9:5622-8.

---

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.
Dengue which were reported to the National Health Authority of India. In 2019, till November, 136422 cases and 132 deaths were reported to the Ministry of Health and Family Welfare, Government of India.[10] This is an underestimate of all the actual cases in the community.[11]

In India, the National Vector Borne Disease Control Programme, under the aegis of the National Health Mission, is primarily responsible for overall vector and Dengue control. Integrated Disease Surveillance Programme is the source of information for us about the dengue numbers in India. Currently, the dengue surveillance in India is performed by passive, sentinel, and hospital based information.[12] Dengue is one of the major public health problems in West Bengal.[13] Almost all districts of the state are dealing with this mosquito borne disease. Primary care physicians see patients of fever. Knowing the local epidemiology of dengue will help the physician take appropriate steps in managing probable dengue cases.

Past numbers of the dengue outbreaks and cases are very important and useful source for predicting the future burden. The outbreaks of dengue can be predicted by suitable epidemiological models.[14] In the present study, we have developed a ARIMA-based prediction model for dengue using time series data over the past decade of state of West Bengal in order to forecast incidences of Dengue outbreak and cases for the year 2020.

Materials and Methods

Data collection

A disease outbreak is the occurrence of excess number of cases than expected in a given area or among a specific group of people over a particular period of time.[15] Integrated Disease Surveillance Project (IDSP) receives disease outbreak reports from all the States and Union Territories on a weekly basis through its IDSP portal. All disease outbreaks reported from the states/UTs are compiled in the form of a Weekly Outbreak Report and is uploaded on its website for availability to scholars, authorities and public. Weekly disease outbreak data available in public domain, spanning for a period of 10 years from 2010-2019 were downloaded from (IDSP) web portal.[16] Obtained data were cleaned and arranged (disease, state and district wise) in MS excel spread sheet. Subsequently, all forms of weekly outbreaks of dengue occurring in West Bengal were extracted. The number of total outbreaks reported from 2010-2019 was 151. Of these, 56 were excluded which were follow up of reports of the same outbreaks reported in previous weeks. Furthermore, 10 outbreaks occurring in districts in the Sub Himalayan West Bengal were also excluded. Finally, the remaining 85 dengue outbreaks occurring in the districts of Gangetic West Bengal were considered for descriptive analysis, trend analysis and building time series for Dengue forecast modelling as depicted in the flowchart in Figure 1.

Study area

The State of West Bengal is meteorically divided into two parts: Northern Sub Himalayan and Southern Gangetic.[17] In the present study the following 12 districts in the Gangetic West Bengal were considered: Bankura, Birbhum, Burdwan, Hooghly, Howrah, ...

| Table 1: Profile of Incidences of Dengue Outbreaks reported from West Bengal (2010 to 2019) |
|----------------------------------|----------------------------------|
| Profile of Dengue Outbreaks     | Characteristics of Outbreak Reporting |
| (1) Dengue outbreaks            | (1) Timely reported 20 (23.5%)   |
| (2) Suspected Dengue Outbreaks  | (2) Follow up reporting 39 (45.8%) |
| (3) Mixed Infection Outbreaks   | (3) Late reported 26 (30.6%)     |
| Total Outbreaks                 | Total Outbreaks 85 (100%)        |

![](image)

**Table 2: Dengue outbreaks in districts of Gangetic West Bengal**
Majhi, et al.: Dengue dynamics in gangetic West Bengal

Building time series of cases of Dengue outbreaks
Weekly Dengue outbreak data including suspected Dengue outbreak and mixed infection (Dengue and Chikungunya) reported from Gangetic West Bengal districts were arranged according to their dates of the onset of outbreak. Weekly data was transformed into monthly data and arranged in chronological order in Excel spread sheet for descriptive analysis. Trend analysis was done by Mann-Kendall statistics using MAKESENS application.[13] Time series analysis including seasonal ARIMA modelling was done by IBM SPSS® version 23 software.

In accordance with IDSP website’s accessibility statement and copyright policy, the Central Surveillance Unit team at IDSP New Delhi had been informed about the utilisation of displayed weekly disease outbreak data for the generation of secondary data through analysis with the objective of forecasting of Dengue outbreak cases during the year 2020 in West Bengal.

Results
Of the 85 outbreaks incidences 72 (84.7%) were purely Dengue outbreaks, 7 (8.2%) were suspected Dengue outbreaks and 6 (7%) were mixed infection (Dengue and Chikungunya) outbreaks. 20 (23.5%) outbreaks were timely reported, 39 (45.8%) outbreaks were reported as follow up outbreak incidences, whereas

Figure 2: Trends of Dengue Outbreak Cases in Districts of Gangetic West Bengal from 2010 to 2019
26 (30.6%) outbreaks were reported late in subsequent weeks as shown in Table 1.

From the weekly outbreak data from 2010 to 2019 it was observed that although maximum numbers of outbreaks (26 outbreaks, 30.6%) were reported from the district Nadia, the corresponding maximum Dengue outbreak cases (9291 cases, 35.6%) were reported from the district of North 24 Parganas in the last decade (2010-19). The highest numbers of outbreaks (15) were recorded in the year 2012. Highest number of cases (7573) were recorded in 2016. District North 24 Parganas recorded the highest number of cases (4607) in a single year, 2017 [Tables 2 and 3].

Trends of Dengue outbreak cases in districts of Gangetic West Bengal (2010 to 2019)

MAKESENS application trends statistics program was used to deploy Mann-Kendall Statistical Test to infer the annual trends of dengue outbreak cases in the districts of Gangetic West Bengal. In computation of this statistical test MAKESENS exploits both S-statistics (for time series less than 10 values) and Z-statistics (for time series equal or more than 10 values). The 12 different time series comprising of annual Dengue outbreak cases of the 12 districts were computed and fed into the MAKESENS application. Since the values of all the districts were less than 10, S-statistics was utilized to calculate the significance of trends for each district.
It was observed that the districts Bankura, Birbhum, Burdwan, Hooghly, North 24 Parganas, Paschim Medinipur, Purba Medinipur and South 24 Parganas displayed an upward trend at value of $\alpha > 0.1$ denoting the trends were non-significant. District Murshidabad also displayed upward trend at a significance level equal to 0.1 ($\alpha = 0.1$) which was significant. District Nadia displayed downward trend for Dengue outbreak cases at significance level more than 0.1 ($\alpha > 0.1$), which was non-significant. No trends could be computed for the districts of Howrah and Purulia as they have only one year of reported dengue outbreaks. The trend analysis statistics are shown in Table 4 and visually depicted in Figure 2.

**Forecasting of Dengue outbreak cases in Gangetic West Bengal**

**Building forecasting model**

SARIMA forecasting modelling was used to predict outcomes from independent variables that have seasonal behaviour or trends. The dengue outbreak cases from 2010-2018 were used as the estimation period and cases from 2019 were used as validation period. The cases predicted by the forecasting model for the year 2019 were validated with the actual number of observed cases for the same year. Various simulations were conducted using SPSS Seasonal ARIMA forecasting method. Some of the forecasting models are described in Table 5.

All models under consideration had a non-significant Ljung-Box statistical test, implying the residuals left over after model building process did not have any correlation between them and no model was missed out in the final model building process. All models gave an overestimate of the predicted values against the observed values. It was observed that the model F: $(2,1,1)(0,1,1)_12$ had the maximum Stationary R squared value (0.564) close to 1. On a background of seasonal trend this model could utilize the most of the stationary component of the model to project as a simple mean model.

Model I, $(0,1,1) (0,1,1)_12$ had the least Mean Average Percentage Error (MAPE) value (122.680), showing the variance of dependent (dengue outbreak cases) series with the model predicted level. Model A, $(1,0,0) (1,0,0)_12$ displayed the least Normalized Bayesian Information Criterion (BIC) value (12.964) implying that this model was least penalized for over parameterization. This model also shows the closest predicted value (1314.3) against the observed value (971) which was an overestimate of about 35.36% of the observed value. Based on the above findings and keeping the model parsimonious, the criteria for selecting the forecasting model was least BIC value, model A: $(1,0,0) (1,0,0)_12$ was chosen as the baseline model to forecast dengue outbreak cases for the year 2020. No transformation of inputs were performed in the modelling process. Constant have been included in the model to account for under/over differencing. Auto regression (p) has been performed in Lag 1 Seasonal auto regression (P) have been performed at Lag 1. The model parameters are displayed in Figure 3 and Table 6.

Furthermore, the analysis of the residuals that remained after model building process revealed that the residuals were not correlated as depicted by the Auto Correlation Function and Partial Auto Correlation Function. The correlation picture portrayed by the residuals were just “white noise” as depicted in Figure 4. Also, the Ljung-Box Test was also applied for identifying any significance of the residuals that turned out to be non-significant, i.e., the final model building process did not miss out any model that could have been possible with the residuals.

On comparing the forecast values of dengue outbreak cases with the observed cases that occurred in 2019, it was seen that the model projects non-significant upper critical limit (UCL) and

| Variable                  | Estimate | Standard Error | T statistics | Sig. |
|---------------------------|----------|----------------|--------------|------|
| Constant                  | 228.917  | 119.444        | 1.917        | 0.058|
| AR Lag 1                  | 0.307    | 0.093          | 3.302        | 0.001|
| AR, Seasonal Lag 1        | 0.328    | 0.107          | 3.060        | 0.003|

Figure 3: SARIMA (1,0,0)(1,0,0)12 Model Estimation period (2010 to 2018) & Validation period (2019): Incidence of Observed and Forecast of Dengue Outbreak Cases

| Year  | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Total |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------|
|       | 0   | 0   | 0   | 0   | 0   | 0   | 971 | 0   | 0   | 0   | 0   | 0   | 971   |
| Predicted | 106.5 | 139.3 | 149.3 | 152.4 | 153.3 | 153.6 | 367.5 | 1314.3 | 311.0 | 153.7 | 153.7 | 153.7 | 1314.3 |
| LCL   | 1337.2 | 1409.5 | 1420.6 | 1434.1 | 1445.2 | 1455.5 | 1639.5 | 2566.1 | 1582.9 | 1435.6 | 1425.6 | 1425.6 | 1314.3 |
| UCL   | -1104.0 | -1126.9 | -1122.0 | -1119.4 | -1118.5 | -1118.2 | -904.4 | 42.5 | -960.8 | -1118.1 | -1118.1 | -1118.1 | -1118.1 |

Table 7: Validation of selected SARIMA model

Predicted values having both their Upper Critical Limit (UCL) & Lower Critical Limit (LCL) on either side of zero are significant (August 2019 value).

LCL & UCL values (remaining months) on both sides of zero are treated as non-significant (NS).
lower critical limit (LCL) values for all months of 2019 except August with 1314.3 cases. In reality there are 971 reported dengue outbreak cases only in the month of August 2019. There is a difference of 343.3 cases which is an over-estimate of 35.4% of the actual value [Table 7].

Forecasting Dengue outbreak cases

On application of the model A: (1,0,0) (1,0,0)_{12} to predict forecast of dengue outbreak cases in the year 2020, it was observed that momentum of Dengue cases occurring in 2020 would start somewhere in June, peak around in August and wane off toward October [Figure 5].

However, on the basis of Upper/Lower Confidence Limits (UCL/LCL) of the predicted values for each months of 2020, it was found that they were not significant for any months.

Discussion

There is a current mechanism of weekly reporting of diseases for its causative aetiology from all Government recognized institutes and facilities in the Indian health system. The reporting is done for cases which are not only laboratory diagnosed but also suspected and physician approved symptomatic cases. The reports of all such facilities are compiled at a central level and any spurt of cases (in specific health facility purview) which are suggestive of any outbreak are relayed back and published in a weekly disease outbreak format consisting of the number of outbreaks reported, actions taken and the current status of the outbreak.\textsuperscript{14} This chain of information sharing is satisfactory as long there is no delay in information sharing and taking appropriate measures to contain the outbreak. Our observation for West Bengal was that there was delay of reporting of more than 30% of outbreaks. Previous studies have divided West Bengal meteorologically into two parts, the northern Sub-Himalayan and southern Gangetic West Bengal.\textsuperscript{15} The 10-year data extracted for West Bengal from IDSP showed that only 10 outbreaks were reported from Sub-Himalayan while 85 outbreaks from the Gangetic plains. It suggests that dengue outbreaks are more common and frequent in the Gangetic West Bengal with proportionately higher numbers of cases than the Sub-Himalayan region. This phenomenon may be attributed to be more density of population as well as greater density of responsible vectors due to more favourable mosquito genic conditions in the Gangetic plains than the Himalayan foot hills and also added with other factors such as more industrialization, urbanization and unplanned settlements.

The highest number of outbreaks in the last decade was reported from the district of Nadia and highest number of Dengue cases from the district of North 24 Parganas. The only district with a significant upward rising trend was that of Murshidabad. Cases of disease outbreaks were seen in all the districts of Gangetic West Bengal except Jhargram. The only district that displayed a downward trend was Nadia, which had the maximum number of Dengue outbreaks in the past decade (2010-2019). However, the trend was non-significant.

The forecasting model was chosen with information criterion in consideration like previous studies which have used Akiake Information Criteria (AIC). In this study more strict information criterion than the AIC is used in the form of Bayesian Information Criterion (BIC).\textsuperscript{16}

No data transformation procedure was performed during the model building process. The SARIMA (1,0,0) (1,0,0)_{12} model had no differencing like in other prediction model (2,0,0) (1,0,0) used by Luz et al. for predicting Dengue in Rio de Janeiro, Brazil.\textsuperscript{17} However, other studies of Dengue forecasting in Brazil, Bangladesh and India have incorporated differencing in their models.\textsuperscript{7,18,19}

It was observed in the study that the predicted validation values have substantial difference with respect to the observed values. There are many reasons that might be attributed to this difference. It could be due to inclusion of only cases reported from outbreaks and not including sporadically occurring cases in the modelling process and also there is a sharp dip in the number of reported outbreak
cases in 2019 than the preceding years. Further, the accuracy of the model could be achieved up to greater extent if other climatic parameters such as rainfall, temperature and humidity were also incorporated into the modelling process. Scope of improving modelling by including the diseases like Chikungunya transmitted by the same vector (Aedes mosquito) can also be explored.

Forecasting of a disease equips health administrators to have prepared themselves to deal better with the impending health problem. Having a prior knowledge about the estimates of outbreaks, cases, location and period, the limited resources can be optimized to tide over the situation. Preventive measures such as planning and organizing awareness campaigns, health activities and cleaning of potential vector breeding sites can be accelerated around the period of forecast.

Acknowledgements
The authors are grateful to the Central Surveillance Unit (CSU) team at IDSP New Delhi for research support.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References
1. World Health Organization. Dengue and Severe Dengue. Factsheet, revised March 2020. Geneva: World Health Organization; 2020. Available from: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue. [Last accessed on 2020 Mar 16].
2. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature 2013;496: 504‑7.
3. Dengue/DHF situation in India. National Vector Borne Disease Control Programme, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India. Available from: https://nvbdcp.gov.in/index4.php?lang=1&level=0&linkid=431&lid=3715. [Last accessed on 2020 Mar 16].
4. Bagchi S. Dengue surveillance poor in India. Lancet 2015;386:1228.
5. Piloti E, Nittas V, Murthy GVS. The organization, implementation, and functioning of Dengue surveillance in India-A systematic scoping review. Int J Environ Res Public Health 2019;16:661.
6. Hati AK. Studies on dengue and dengue haemorrhagic fever (DHF) in West Bengal State, India. J Commun Dis 2006;38:124‑9.
7. Bhatnagar S, Lal V, Gupta SD, Gupta OP. Forecasting incidence of dengue in Rajasthan, using time series analyses. Indian J Public Health 2012;56:281‑5.
8. Liu D, Guo S, Zou M, Chen C, Deng F, Xie Z, et al. A dengue fever predicting model based on Baidu search index data and climate data in South China. PLoS One 2019;14:e0226841.
9. Jayarat VJ, Avoi R, Gopalakrishnan N, Raja DB, Umasa Y. Developing a dengue prediction model based on climate in Tawau, Malaysia. Acta Trop 2019;197:105055.
10. Centers for Disease Control and Prevention (CDC). Lesson Six: Investigating an outbreak. Principles of epidemiology in public health practice: an introduction to applied epidemiology and biostatistics. 3rd Edition. October 2006. Page 6‑11.
11. Integrated Disease Surveillance Programme, NCDC, DGHS. Weekly outbreaks [Internet]. Available from https://idsp.nic.in/index4.php?lang=1&level=0&linkid=406&lid=3689.
12. Ghosh KG. Spatial and temporal appraisal of drought jeopardy over the Gangetic West Bengal, eastern India. Geoenvironmental Disasters. 2019 Dec 1;6(1):1.
13. Salmi T, Määttä A, Anttila P, Ruoho-Airola T, Amnell T, Maatta A. Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates: The Excel template application MAKSENS. Helsinki: Finnish Meteorological Institute; August 2002.
14. Raut D, Bhola A. Integrated disease surveillance in India: Way forward. Global Journal of Medicine and Public Health. 2014;3(4):1‑10.
15. Dastidar AG, Ghosh S, De UK, Ghosh SK. Statistical analysis of monsoon rainfall distribution over West Bengal, India. Mausam 2010;61:487‑98.
16. Clement EP. Using normalized bayesian information criterion (BIC) to improve box-jenkins model building. Am J Math Stat 2014;4:214‑21.
17. Luz PM, Mendes BV, Cedeño CT, Struchiner CJ, Galvani AP. Time series analysis of dengue incidence in Rio de Janeiro, Brazil. Am J Trop Med Hyg 2008;79:933‑9.
18. Martinez EZ, Silva EA. Predicting the number of cases of dengue infection in Ribeirão Preto, São Paulo State, Brazil, using a SARIMA model. Cad Saude Publica 2011;27:1809‑18.
19. Jain R, Sontisirikit S, Iamsirithaworn S, Prendinger H. Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data. BMC Infect Dis 2019;19:272.