Epidemiology and Health System Journal
doi:10.34172/ehsj.26611 2026;13(1):35-44 http://ehsj.skums.ac.ir

Original Article

# Modeling Dengue Dynamics of Sri Lanka Using Window-Wise Dynamic Mode Decomposition

**Tanmoy Roy Choudhury**[1*]

[1]Department of Basic Science and Humanities, MCKV Institute of Engineering, Liluah, Howrah-711204, West Bengal, India

**Abstract**

**Introduction:** Dengue fever poses a major global public health challenge, making early outbreak forecasting essential for effective intervention. While traditional time-series models provide baseline forecasts, they inadequately capture spatial coherence. Advanced approaches, including long short-term memory (LSTM) and spatial or Bayesian state-space models, better represent spatiotemporal dynamics but frequently require strong assumptions or high computational costs. However, dynamic mode decomposition (DMD) provides a low-rank, equation-free framework for extracting dominant dynamical patterns from complex datasets. Therefore, this study assessed the effectiveness of DMD in characterizing and forecasting dengue transmission dynamics across Sri Lanka.

**Methods:** Monthly dengue incidence data (2010-2020) from 26 cities were Z-score normalized and analyzed using standard DMD, with no exclusions. The dataset was partitioned into a pre-outbreak period (2010-2016) and an outbreak/post-outbreak period (2017-2020), with the final six months of each period reserved for testing. Then, forecast accuracy was evaluated using six-month-ahead root-mean-square error (RMSE) while not employing imputation, cross-validation, or rolling-window validation.

**Results:** DMD identified low-rank coherent modes corresponding to long-term trends and multi-year oscillations. In the pre-outbreak period, the stationary mode (~40%) was dominated by Colombo and Gampaha ($R \approx 0.898$), whereas during the outbreak, dominance shifted toward Jaffna (>12%). Moreover, six-month forecasts achieved high accuracy (RMSE: $0.04 \pm 0.03$ pre-outbreak; RMSE: $0.005 \pm 0.006$ post-outbreak), outperforming LSTM and autoregressive integrated moving average (ARIMA) models.

**Conclusion:** Overall, DMD effectively captures remarkable spatiotemporal dengue dynamics and enables accurate short-term forecasting. Its computational efficiency and minimal assumptions make it a promising tool for real-time dengue early warning and broader infectious disease surveillance.

**Keywords:** Dengue fever, Dynamic mode decomposition, Temporal spatiotemporal analysis, Disease forecasting, Data-driven modeling, Machine learning in epidemiology

***Corresponding Author:**
Tanmoy Roy Choudhury,
Email: tanmoy.roychoudhury@mckvie.edu.in

**Received:** October 9, 2025
**Revised:** February 18, 2026
**Accepted:** February 21, 2026
**ePublished:** June 29, 2026

## Introduction

Dengue fever is a major global public health threat, with over 11 million reported cases and more than 7,000 deaths worldwide since January 2024. It is caused by four dengue virus serotypes (DENV1-4) and transmitted by Aedes aegypti and Aedes albopictus mosquitoes. [1] In addition, its spread is strongly influenced by environmental factors, such as temperature, humidity, and rainfall. These factors affect mosquito survival, biting behavior, and transmission efficiency,[2] thereby making dengue dynamics complex and highly variable.

One approach is to use traditional compartmental models (e.g., susceptible infected recovered and its variants) to study host-vector interactions[3] and estimate key epidemiological parameters (e.g., the basic reproduction number) and analyze disease-free equilibrium points.[4,5] However, these nonlinear, coupled models are sensitive to initial conditions and may yield unstable long-term predictions. Some studies have applied empirical mode decomposition to coronavirus disease 2019 (COVID-19) data from Tokyo to separate low-frequency and high-frequency components associated with external influences and random fluctuations.[6] Therefore, data-driven approaches that directly analyze spatiotemporal disease patterns offer a promising alternative.

Dengue continues to impose a heavy burden in tropical regions, including Sri Lanka, which experienced over 200,000 cases during the 2017-2019 epidemic and more than 24,000 cases in 2024 alone. These public health challenges underscore the need for robust, generalizable analytical tools to support early warning and surveillance. Several statistical methods have been applied to Sri Lankan dengue data. For example, autoregressive integrated moving average (ARIMA)

models could successfully forecast weekly dengue incidence in Colombo but were limited to a single district and lacked spatial generalization.[7] Moreover, exponential smoothing methods (e.g., Holt–Winters) have shown reasonable predictive accuracy for monthly dengue cases in Colombo, but their robustness across other regions remains untested.[8] Additionally, machine-learning approaches (e.g., k-nearest neighbors, support-vector machines, and regression models) have been explored for dengue forecasting.[9, 10] While these methods can achieve acceptable predictive performance, they frequently lack interpretability and fail to explicitly capture spatiotemporal coherence across regions. The wavelet-based analyses of monthly dengue incidence in Vietnam revealed heterogeneous multiannual components across provinces and increased synchronization during high-incidence years.[11] While these approaches highlight the importance of spatial coherence, they primarily capture periodicity or static clustering while not thoroughly characterizing evolving spatiotemporal patterns. Few studies have explicitly examined spatiotemporal patterns in dengue to jointly capture temporal and spatial correlations within specific regions. Such analyses are crucial for developing region-specific control strategies and policies. Spectral techniques (e.g., fast Fourier transform and wavelet analysis) have been applied to dengue data from Colombo, Jakarta, and Thailand to reveal hidden spectral characteristics.[12]

This study demonstrates the capability of dynamic mode decomposition (DMD), a data-driven spectral technique, in order to extract dominant spatiotemporal patterns from infectious disease surveillance data.[13, 14] By decomposing multi-location dengue incidence into a small set of coherent modes, DMD provides interpretable insights into disease evolution and dynamic relationships between regions.[15-18] Furthermore, the method enables short-term forecasting directly from observed data, without requiring mechanistic assumptions. Using a multi-city dengue dataset from Sri Lanka as a proof of concept, DMD is highlighted as a powerful and generalizable approach for real-time epidemic monitoring and early warning applications.

## Materials and Methods

### *Algorithm*

This study relies on DMD's capability as a powerful data-driven approach for modeling and forecasting infectious disease dynamics directly from surveillance data. By applying DMD to monthly dengue data spanning 11 years (2010-2020) across 26 cities in Sri Lanka as a proof-of-concept, the study aims to demonstrate that the essential spatiotemporal structure of disease transmission can be captured using only a limited number of dominant modes. As a purely data-driven technique, DMD does not require knowledge of the governing epidemiological processes or their differential equations; therefore, even highly nonlinear systems can be modelled and analyzed without first deriving first-principles models.

Although originally developed to analyze spatial and temporal patterns in fluids,[13,14] DMD has since been applied to epidemiology to better analyze disease dynamics, including Google Flu, measles, Type-1 polio, and COVID-19.[15-18] To the best of our knowledge, however, no study has so far applied DMD to dengue surveillance data. Accordingly, this work sought to demonstrate DMD as a generalizable modeling framework, using dengue data from Sri Lanka as a proof-of-concept application. It should be noted that no specific inclusion or exclusion criteria were applied since the data were obtained from a publicly available dataset reporting monthly dengue cases from all cities in Sri Lanka.

DMD analyzes the correlation between pairs of data that are temporally spaced $\Delta t$ apart. The equation for discrete snapshots in time is as:

$$X_{k+1} = AX_k$$

Where $X_k \in \mathbb{R}^n$ represents the $n$-dimensional spatial data at a particular instant of time ($t = k$). In this work, $n = 26$ cities of Sri Lanka are considered at month $k$. Thus, if the $n$-dimensional spatial data from all the time snapshots are concatenated into a single matrix $X$ and a time-shifted matrix $X'$, then

$$X = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ X_1 & X_2 & \dots & \dots & X_k \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \text{ and } X' = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ X_2 & X_3 & \dots & \dots & X_{k+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Where $X, X' \in \mathbb{R}^{n \times k}$ The relation between $X$ and $X'$ can be approximated as

$$X' \approx AX$$

This equation can be solved using data alone, without knowing the exact physical principles governing the system's dynamics. In addition, $A \in \mathbb{R}^{n \times n}$ is the best fit solution for all pairs of

$$X_m, X_{m+1}, m = 12 \dots k$$

Ideally $A = X'X^{\dagger}$, where $X^{\dagger}$ is the pseudo-inverse of $X$. DMD aims to compute this using the singular value decomposition (SVD) of $X$. The truncated SVD of $X$ can be written as

$$X \approx U_r \Sigma_r V_r^*$$

Where $U_r \in \mathbb{R}^{n \times r}$, $\Sigma_r \in \mathbb{R}^{r \times r}$, $V_r^* \in \mathbb{R}^{r \times k}$, and $r$ denote the number of singular values remaining after truncation. Thus,

$$A \approx X'V_r \Sigma_r^{-1} U_r^*$$

An approximate operator for $A$ can be found as

$$\overline{A} = U_r^* X'V_r \Sigma_r^*$$

Where $\bar{A}$ is the reduced order matrix with the Eigen decomposition.

$$\bar{A}Y = Y\Lambda$$

Where $Y$ and $\Lambda$ are the reduced Eigen vectors and the Eigen values, respectively. The Eigen vectors for the complete $A$ matrix are expressed as

$$\phi = X'V_r\Sigma_r^{-1}Y$$

Where $\phi$ represents the various dynamic modes, and each DMD mode has a corresponding Eigen value $\lambda_m$, with $m = 1,2\ldots r$. The continuous time frequency for each mode may be computed as

$$f_m = img\left(\frac{\ln \lambda_m}{\Delta t}\right)/2\pi$$

### *Data Collection and Processing*

Monthly dengue case counts across 26 cities in Sri Lanka from January 2010 to May 2020 were collected from a publicly available dataset hosted on Kaggle ("SRI LANKA Dengue cases from 2010 to 2020" - https://www.kaggle.com/datasets/sadaruwan/sri-lanka-dengue-cases-2010-2020), which claims to have scraped the data from the Epidemiology Unit of Sri Lanka (https://www.epid.gov.lk/). The dataset, therefore, provides routine case-report surveillance covering 133 months.

Considering that the monthly dengue incidence can exhibit abrupt variability, cumulative case series were calculated to generate smoother temporal profiles better suited for DMD analysis. The cumulative values for each city were standardized using the following equation:

$$X_{st} = \frac{X - \langle X \rangle}{\sigma}$$

Where $X$ denotes the cumulative series. In addition, $\langle X \rangle$ and σ are the mean and standard deviation (SD), respectively. The resulting dataset constitutes a spatiotemporal matrix of size $26 \times 133$, with each row representing a city and each column a monthly time point. To examine changes in dynamical structure, the full study period was divided into two windows (Table 1).

For each window, snapshot matrices $X$ and $X'$ are constructed as:

$$X, X' \in \mathbb{R}^{26\times 84}\,(\text{Window1}),\ X, X' \in \mathbb{R}^{26\times 41}\,(\text{Window2})$$

To evaluate predictive performance, the final six months of each window were withheld as test data, while the remaining observations were used for model training.

### *Software*

All computations were performed in Python (version 3.13.5) using standard scientific libraries. In addition, SVD was implemented via numpy.linalg, and NumPy was utilized for numerical operations. Moreover, data handling relied on pandas, and visualization and graphical outputs were generated using Matplotlib version 3.10.0.

## Results

### *Eigenvalues and Dominant DMD Modes*

#### *Window 1: January 2010 – December 2016*

DMD was performed on the cumulative case matrix, $X' \in ,\mathbb{R}^{26\times 84}$, and the root-mean-square error (RMSE) was calculated only for future prediction of dengue cases for the subsequent 6 months. To determine the optimal model order, RMSE values between DMD predictions and observed values were computed for all 26 cities across truncation orders ranging from 2 to 20. Forecasting accuracy varies considerably depending on the retained rank. The lowest average RMSE is achieved with a truncation order of 8, yielding an average RMSE of 0.044 and a SD of 0.034. Ranks lower than 8 underfit the data, whereas higher truncation orders result in overfitting, characterized by rapidly increasing RMSE and variance, particularly beyond rank 12. Therefore, a truncation order of 8 was selected as the most appropriate balance between model complexity and forecasting performance for Window 1. Out of the 8 Eigen values, two were non-oscillatory, and the remaining six were in the form of conjugate pairs. The power of the $i^{th}$ DMD mode was computed as follows:

$$P_i = \lambda_i^k \, \phi_i$$

Where $k$ is obtained by an iterative process and is chosen as 10 in this work. It should be noted that a variation in $k$ beyond this value keeps the order of significance for various modes unaltered.

The most dominant of the eight modes was Mode 5, a stationary mode (with an Eigen value of 1) that represents constant flow over time, reflecting the persistent background growth in the cumulative trend. Modes 3 and 6 were the next most dominant, both demonstrating decaying behavior with oscillatory characteristics (Figure 1). Mode 3 was the faster oscillatory mode, with a repeatability of approximately 37 months, compared to Mode 6, which had a longer repeatability of approximately 98 months. Therefore, none of these modes can be attributed to the data's seasonality.

#### *Window 2: January 2017 – May 2020*

DMD was then applied to the period associated with the major dengue outbreak in Sri Lanka, using $X' \in \mathbb{R}^{26\times 41}$.

**Table 1.** Time Windows for Dengue Outbreak Analysis

| Window | Period | Duration | Interpretation |
|---|---|---|---|
| 1 | January 2010 – December 2016 | 84 months (7 years) | Pre-dengue outbreak dynamics |
| 2 | January 2017 – May 2020 | 41 months (4 years) | Outbreak and post-outbreak behavior |

To determine an appropriate truncation order for Window 2, RMSE was again computed for all 26 cities and truncation ranks ranging from 2 to 20. Forecast performance systematically improved as the retained rank increased from 2 up to 12. The lowest average RMSE was obtained for a truncation order of 12, with an average RMSE of 0.0048 and a SD of 0.0064. Further increasing the truncation order led to a noticeable deterioration in forecasting accuracy, with both the mean and the variance of RMSE rising, indicating overfitting and reduced generalizability. Thus, rank 12 was selected as the optimal truncation order for Window 2, enabling optimum short-term prediction while preserving the intrinsic dynamical structure of the outbreak phase. The dominant modes in this time window markedly differed from those in Window 1, indicating changes in the underlying system dynamics. In this window, Mode 8 was the strongest and most stationary, followed by Mode 9 (also stationary), Mode 10 (with a longer periodicity of ~45 months), and Mode 4 (with a shorter cycle of ~11 months). The temporal dynamics of Modes 10 and 4 are displayed in Figure 2. Therefore, the seasonal variation in the data can only be attributed to Mode 4.

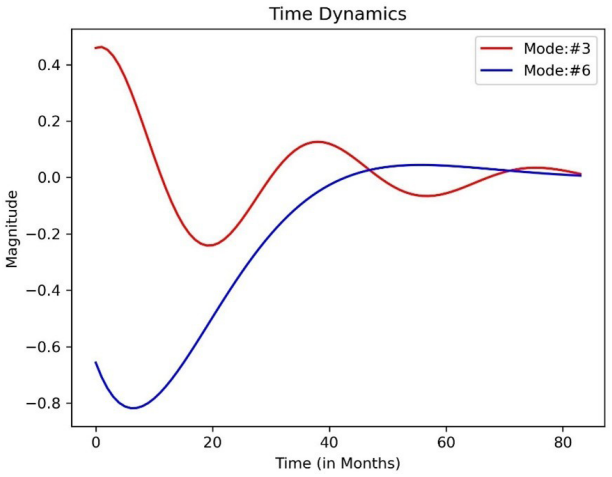


**Figure 1.** Temporal Dynamics of Selected Dominant DMD Modes in Window 1 (2010-2016). *Note*. DMD: Dynamic mode decomposition

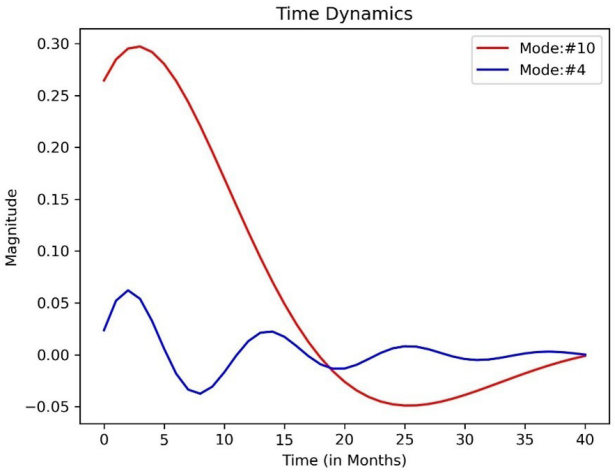


**Figure 2.** Temporal Dynamics of the Most Dominant Non-Stationary Modes in Window 2 (2017-2020)

### *Magnitude and Phase Analysis*

DMD provided a natural decomposition of disease dynamics into spatial modes (magnitude) and their temporal progression (phase). Moreover, the magnitude analysis of a particular mode identified the spatial locations that contributed most to it. In contrast, phase relationships revealed temporal synchrony or latency between regions, a key advantage of DMD in analyzing surveillance data.

#### *Window 1: January 2010 - December 2016*

Figure 3 illustrates how 26 cities contributed to the top three dominant modes. Mode 5 had the highest contribution from Colombo and Gampaha, both accounting for about 40% of its total magnitude. Mode 5 was non-oscillatory and had a magnitude close to 1, implying that cumulative dengue cases remained almost steady throughout the year. This suggests that a stable, non-changing component accounts for most disease growth during the pre-outbreak period.

Modes 3 and 6 were oscillatory with different time scales. Colombo, Gampaha, Kurunegala, and Rathnapura strongly contributed to Mode 3 (over 37%), indicating a group of cities with similar long-term trends. In addition, Mode 6 was dominated by Jaffna and Rathnapura (25%) and varied more slowly, suggesting a distinct subgroup with its own dynamics.

Some city pairs also shared similar mode patterns: Colombo and Gampaha ($R \approx 0.898$), Batticaloa and Jaffna ($R \approx 0.92$), and Kurunegala and Rathnapura ($R \approx 0.97$), where $R$ is Pearson's correlation coefficient calculated on the modal distribution patterns (Figure 3). In Mode 3, Gampaha was out of phase with Colombo, Jaffna, Kurunegala, and Rathnapura by about 18 months, while in Mode 6, Jaffna and Rathnapura were out of phase by nearly 48 months.

These phase differences revealed time lags in outbreak evolution and indicated that cities can behave similarly even if they are not geographically close, demonstrating DMD's ability to detect non-spatial transmission links.

#### *Window 2: January 2017 – May 2020*

Following the major 2017 outbreak in Sri Lanka, the modal structure changed noticeably (Figure 4), showing a shift in system behaviour that DMD can detect without any model assumptions. The dominant stationary component, Mode 9, was mostly driven by Jaffna (~12% of its total magnitude), representing a new dominant growth pattern centered in the north during this period. Jaffna and Kandy showed tightly correlated behavior in this mode ($R \approx 0.98$), indicating shared post-outbreak transmission dynamics.

Among the oscillatory modes, Mode 10 was influenced by Jaffna, Gampaha, Kandy, Batticaloa, and Trincomalee, which all accounted for over 34% of the mode, forming a wider connected group that stays in phase. Additionally, Mode 4 was mainly driven by Jaffna, Gampaha, Kandy, Colombo, and Batticaloa (over 53%), demonstrating

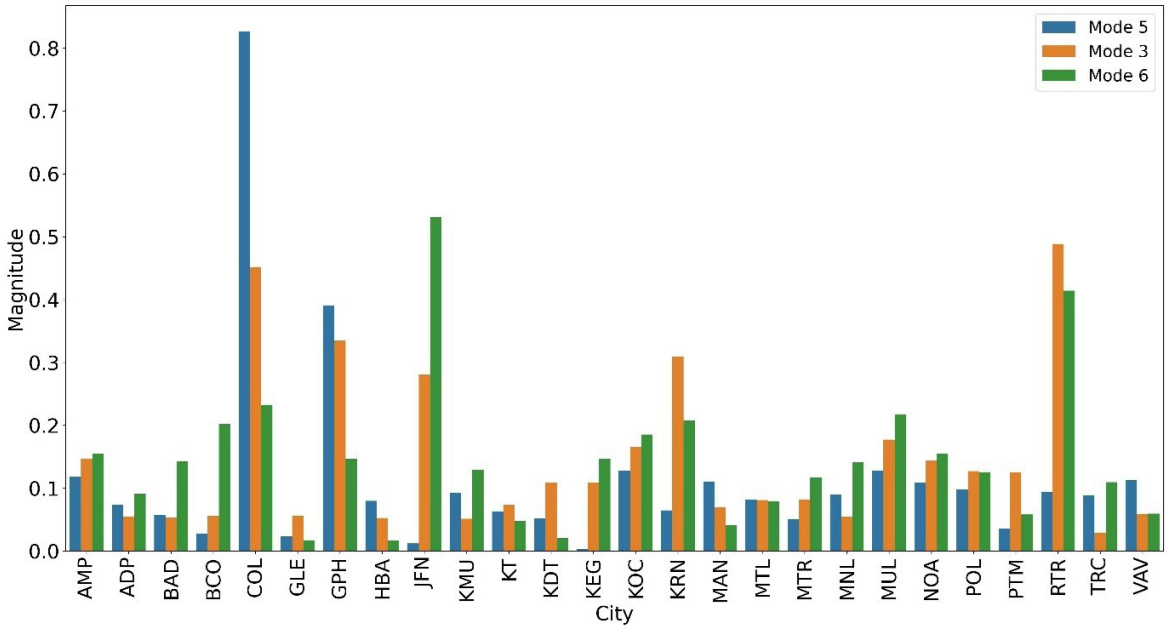


**Figure 3.** Modal Contribution of Different Cities in the Time Interval 2010-2016

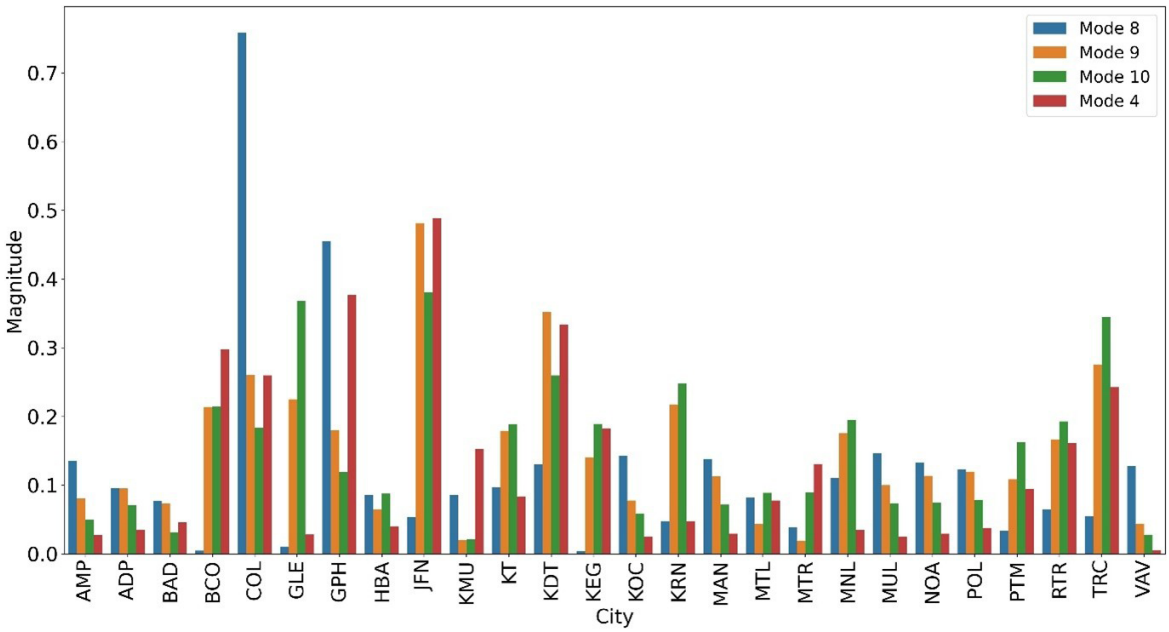


**Figure 4.** Modal Contribution of Different Cities in the Time Interval 2017-2020

enhanced synchronization across multiple high-burden regions during the outbreak and post-outbreak period.

### *Reconstruction and Short-Term Forecasting*

Figures 5 and 6 compare DMD-simulated reconstructions with observed cumulative case data for different cities of Sri Lanka for time Windows 1 and 2, respectively. Red dots denote the actual data points, and the DMD prediction is shown as a blue curve. The predicted values from DMD closely matched the actual values. Therefore, DMD accurately captured the primary dynamics despite being low-rank, highlighting an effective dimensionality reduction without compromising fidelity.

In addition to reconstruction, DMD was used to forecast dengue cases for six months beyond the final data point of each window. These six-month-ahead forecasts closely align with observed trajectories in both cases, indicating that the extracted modal dynamics generalize effectively over short prediction horizons.

A quantitative assessment of forecasting accuracy was performed using RMSE values calculated for each city over a six-month prediction horizon, excluding the period used for model training. The average RMSE across all 26 cities during the pre-outbreak period (2010-2016) was approximately $0.044 \pm 0.033$, with the highest errors observed in Galle (0.067) and Trincomalee (0.071). During the outbreak/post-outbreak period (2017-2020), the average RMSE reduced to approximately $0.005 \pm 0.006$, representing improved consistency in model performance at higher case magnitudes. Slightly elevated RMSE values

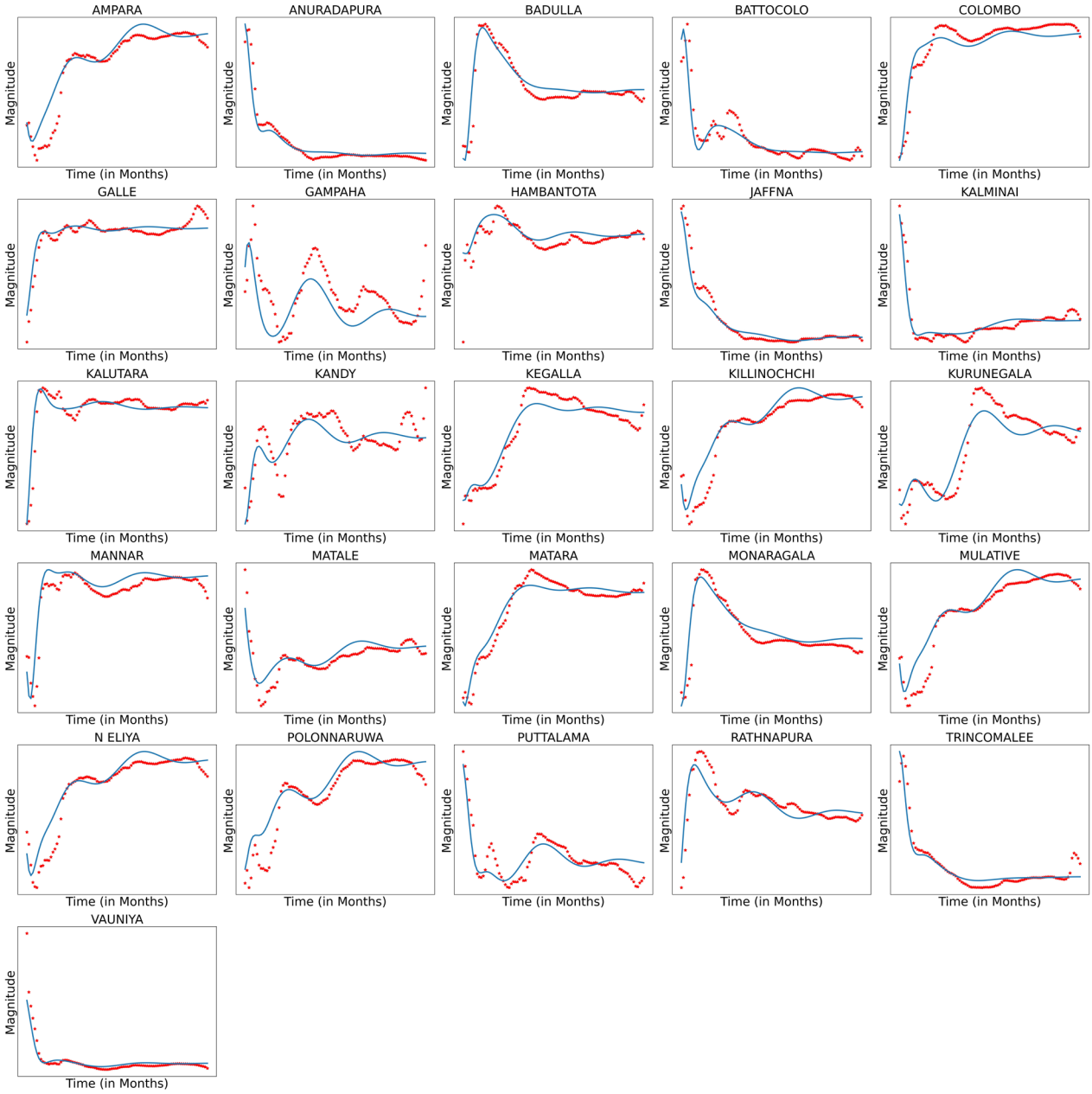


**Figure 5.** Comparison of Actual and Simulated Dengue Data Using DMD in Window 1. *Note*. DMD: Dynamic mode decomposition

were observed in Kandy (0.011) and Rathnapura (0.013). Overall, DMD achieved accurate reconstruction and short-term prediction performance across both study windows.

To benchmark the predictive capability of DMD, comparative forecasting models were evaluated under identical conditions. A long short-term memory (LSTM) neural network was implemented using two stacked LSTM layers with 64 units each, followed by a dense layer of 32 neurons and a single-neuron output layer. ARIMA models were also trained individually for each city as a conventional statistical time-series baseline. Moreover, ARIMA models were fitted individually for each city as a traditional statistical baseline, selecting optimal p (autoregressive order), d (integrated order), and *q* (moving average order) parameters using standard model-selection procedures. The LSTM network yielded RMSE values of $0.798 \pm 0.410$ for Window 1 and $0.1768 \pm 0.094$ for Window 2, whereas the ARIMA approach led to RMSE values of $0.0944 \pm 0.094$ and $0.0107 \pm 0.0104$ for the two windows, respectively (Table 2). In contrast, DMD consistently achieved the lowest prediction error in both windows, demonstrating that a reduced-order representation derived from dominant spatiotemporal structures can outperform both deep learning and classical time-series approaches for six-month dengue forecasting. These findings reinforce the use of DMD as a lightweight and interpretable early-warning framework for dengue surveillance.

Finally, to quantify how prediction error changes with increasing forecast horizon, the average RMSE was calculated across all 26 cities for each of the six future months predicted by the model. Based on the results (Table 3), RMSE gradually increased as the forecast

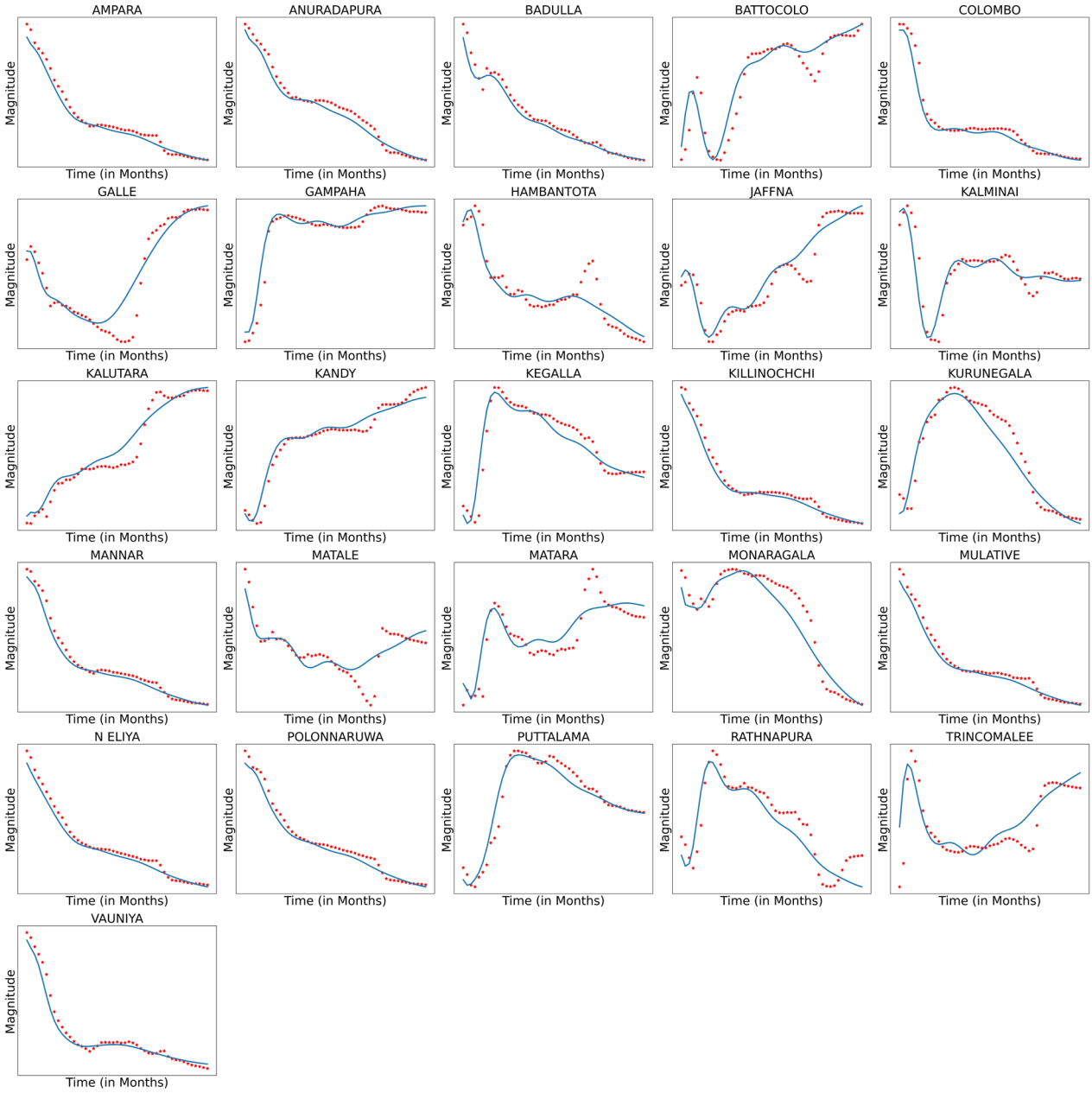


**Figure 6.** Comparison of Actual and Simulated Dengue Data Using DMD in Window 2. *Note*. DMD: Dynamic mode decomposition

**Table 2.** Comparison of the Six-Month Forecasting Accuracy Using DMD, LSTM, and ARIMA Models for Two Temporal Windows

| Window | Truncation Order Selected | DMD RMSE (AVG ± SD) | LSTM (AVG ± SD) | ARIMA (AVG ± SD) |
|---|---|---|---|---|
| Window-1 (2010 - 2016) | 8 | 0.044 ± 0.034 | 0.798 ± 0.410 | 0.0944 ± 0.094 |
| Window-2 (2017 - 2020) | 12 | 0.0048 ± 0.0064 | 0.1768 ± 0.094 | 0.0107 ± 0.0104 |

*Note*. DMD: Dynamic mode decomposition; LSTM: Long short-term memory; ARIMA: Autoregressive integrated moving average; AVG: Average; SD: Standard deviation.

**Table 3.** Month-Wise RMSE Evaluation of Six-Month DMD Forecasts Across Two Temporal Windows

| | Month 1 | Month 2 | Month 3 | Month 4 | Month 5 | Month 6 |
|---|---|---|---|---|---|---|
| Window 1 | 0.0474 | 0.0535 | 0.0558 | 0.0505 | 0.0475 | 0.0736 |
| Window 2 | 0.0040 | 0.0062 | 0.0076 | 0.0087 | 0.0097 | 0.0104 |

*Note*. DMD: Dynamic mode decomposition.

horizon extended into the future, due to the expected accumulation of uncertainty. Nevertheless, the magnitude of error remained low in both analyzed windows, indicating that DMD maintains reliable predictive skills over the short time frame most relevant to operational dengue preparedness and public health response planning.

## Discussion

The findings of this study revealed that DMD can effectively characterize and forecast dengue transmission using routinely collected surveillance data, capturing dynamics that traditional statistical and machine-learning methods often miss. It is noteworthy that the identification of a low-rank structure, in which a small number of modes reconstruct the observed temporal patterns, provides

a compact representation of dengue transmission processes. Similar low-dimensional dynamics have been reported for influenza, measles, and COVID-19 [15-18], and our results extend this evidence to dengue, highlighting the value of DMD for arboviral disease surveillance. The dominant modes for Sri Lanka reflect sustained growth trends and multi-year oscillations, consistent with known inter-epidemic cycles driven by immunity, climate variability, and viral serotype dynamics.[19]

The modal structure considerably differed between the pre-outbreak (2010-2016) and outbreak/post-outbreak (2017-2020) periods, indicating that DMD can detect structural shifts in transmission dynamics without requiring an explicit epidemiological model. Before the outbreak, persistent transmission was concentrated in Colombo and Gampaha, which is consistent with the prior identification of the Western Province as a major dengue hotspot.[7,8] Additional cities formed distinct modal clusters, such as Kurunegala Rathnapura (R~0.97), and Jaffna Batticaloa Trincomalee, suggesting shared climatic, socioeconomic, or mobility-related drivers. Moreover, phase analysis revealed delayed transmission in Gampaha (~18 months) and Jaffna and Rathnapura (~48 months), underscoring spatial heterogeneity and opportunities for early interventions based on leading regions.

During and after the 2017 epidemic, sustained transmission shifted northward toward Jaffna, although Colombo and Gampaha remained influential. Seasonal variability was captured by a separate mode involving Jaffna, Colombo, Gampaha, Kandy, and Batticaloa, with phase differences in line with climate-driven timing offsets. Additionally, DMD demonstrated coherent dynamics between non-adjacent city pairs, such as Batticaloa–Jaffna and Kurunegala–Rathnapura, supporting evidence that dengue spread is driven more by human mobility and socioeconomic connectivity than geographic proximity.[20] These insights can inform coordinated surveillance and vector-control strategies across functionally connected regions.

Despite its reduced-order formulation, DMD achieved consistently high reconstruction accuracy in this study, with RMSE as low as 0.0048 and robust six-month predictive performance, outperforming both ARIMA models and LSTM neural networks. Existing dengue forecasting studies commonly rely on ARIMA, SARIMA, and their hybrids, combined with machine-learning models (e.g., LSTM), reporting RMSE values of 108.27–102.59 and mean absolute percentage error of 1.554.[7,21,22] More advanced deep-learning approaches, including SSA-LSTM in Malaysia [23] and climate-driven LSTM models in India [24], achieved RMSE ranges of 2.91–4.55 and 0.345, respectively. Furthermore, artificial neural network-based models have been applied in Sri Lanka, reporting RMSE as low as 0.025.[25,26]

While these statistical and machine-learning models demonstrate strong predictive performance, they are typically applied to individual locations and, therefore, fail to capture broader spatial coherences critical to understanding transmission dynamics. Conversely, DMD's superior performance, along with its interpretability and computational efficiency, makes it well-suited for real-time early-warning systems, particularly in data-limited settings where rapid model updating is essential.

Several cities, including Ampara, Kilinochchi, Mullaitivu, and Nuwara Eliya, revealed stable modal contributions throughout the study period, indicating persistent structural characteristics and relatively consistent transmission pressures that can be addressed through routine surveillance and baseline seasonal control measures.

Although climatic and ecological factors (e.g., temperature, humidity, and precipitation) play a critical role in dengue transmission [27], this analysis relied solely on routinely reported case data because comparable environmental data were not publicly available at suitable spatial and temporal resolutions. Nevertheless, DMD is readily extensible to multivariate settings through the joint decomposition of coupled time series. Integrating environmental variables into a stacked DMD framework could enable direct, data-driven quantification of climatic influences on dengue dynamics without explicit mechanistic assumptions, potentially improving forecast robustness during climate-sensitive periods.

Overall, our findings demonstrated that DMD is an effective short-term forecasting tool and a powerful analytical framework for uncovering latent temporal structures and evolving spatial relationships relevant to dengue preparedness. By identifying dominant contributors to transmission modes and shifts in regional influence, DMD provides actionable insights to support targeted vector control, improved inter-district coordination, and more proactive public-health responses. While demonstrated using dengue data from Sri Lanka, the approach is broadly applicable to other infectious diseases and geographic settings, underscoring its value for real-time disease surveillance and decision support.

## Conclusion

The findings of this study confirmed that DMD is an effective and generalizable tool for characterizing and forecasting dengue transmission patterns using routinely collected surveillance data. Using a test case with multi-city data from Sri Lanka, DMD successfully extracted dominant low-dimensional structures governing the disease system, thereby revealing key cities that consistently exert a strong influence on broader infection trends. By identifying leading dynamic contributors, such as Colombo and Gampaha, in the pre-outbreak period and the later emergence of Jaffna as a key driver, DMD provided critical information on where strategic interventions could yield the greatest national impact. In addition, the method uncovered persistent relationships between cities that were not necessarily adjacent geographically but shared synchronized outbreak

behavior and quantified meaningful phase delays that reflected staggered epidemic timing.

Importantly, a major advantage of DMD was its ability to identify these coherent structures directly from observational data without prior knowledge of the mechanisms of disease spread, as it relies solely on time-series data. In addition to interpretability, DMD demonstrated reliable short-term predictive capability, maintaining low error values in six-month forecasts despite substantial changes in epidemic intensity between windows. This reliability underscores its utility as a rapid decision-support tool in public health, including targeted surveillance enhancement, optimized timing of vector-control operations, and informed allocation of clinical resources during periods of potential outbreak escalation.

As DMD can be continuously updated as new data become available, it offers an efficient and scalable approach for real-time epidemic intelligence, especially when mechanistic models are constrained by uncertainty in biological parameters. Accordingly, future extensions integrating environmental or mobility indicators may enhance interpretability and improve prediction in climatically sensitive or highly connected regions. Although in this study, DMD was applied to dengue in Sri Lanka, the methodology is broadly applicable to surveillance systems worldwide and has the potential to strengthen real-time epidemic intelligence and preparedness planning.

**Acknowledgments**

The author gratefully thanks the MCKV Institute of Engineering for providing the necessary environment and computer facilities, including installed software, for conducting the research work. In addition, special thanks go to the Kaggle dataset repository and the Epidemiology Unit of Sri-Lanka for providing the data.

**Authors' Contribution**

This study was solely conducted by Dr. Tanmoy Roy Choudhury.

**Competing Interests**

The author declares no conflict of interests regarding this manuscript.

**Ethical Approval**

No ethical approval was required as the research used publicly available data and contained no identifying details of individuals.

**Funding**

The author received no financial support for this research.

**References**

1. Maduranga S, Valencia BM, Sigera C, Adikari T, Weeratunga P, Fernando D, et al. Genomic Surveillance of Recent Dengue Outbreaks in Colombo, Sri Lanka. Viruses 2023;15(7):1408. doi:10.3390/v15071408
2. Mamenun, Koesmaryono Y, Sopaheluwakan A, Hidayati R, Dasanto BD, Aryati R. Spatiotemporal Characterization of Dengue Incidence and Its Correlation to Climate Parameters in Indonesia. Insects 2024;15(5):366. doi:10.3390/insects15050366
3. Huppert A, Katriel G. Mathematical modelling and prediction in infectious disease epidemiology. Clin Microbiol Infect 2013;19(11):999–1005. doi:10.1111/1469-0691.12308
4. Side S, Noorani M. A SIR model for spread of dengue fever disease (simulation for south sulawesi Indonesia and selangor Malaysia)2015. doi:10.13140/RG.2.1.5042.6721
5. El-shenawy A, El-Gamel M, Teba A. Simulation of the SIR dengue fever nonlinear model: A numerical approach. Partial Differential Equations in Applied Mathematics 2024;11:100891. doi:10.1016/j.padiff.2024.100891
6. Dong R, Ni S, Ikuno S. Nonlinear frequency analysis of COVID-19 spread in Tokyo using empirical mode decomposition. Sci Rep 2022;12(1):2175. doi:10.1038/s41598-022-06095-w
7. Karasinghe N, Peiris S, Jayathilaka R, Dharmasena T. Forecasting weekly dengue incidence in Sri Lanka: Modified Autoregressive Integrated Moving Average modeling approach. PLoS One 2024;19(3): e0299953. doi:10.1371/journal.pone.0299953
8. Attanayake AMCH, Perera SSN, Liyanage UP. Exponential smoothing for forecasting dengue cases in Colombo, Sri Lanka. JSc 2020;11(1):11-22. doi:10.4038/jsc.v11i1.24
9. Gupta G, Khan S, Guleria V, Almjally A, Alabduallah BI, Siddiqui T, et al. DDPM: A Dengue Disease Prediction and Diagnosis Model Using Sentiment Analysis and Machine Learning Algorithms. Diagnostics (Basel) 2023;13(6):1093. doi:10.3390/diagnostics13061093
10. Mishra VK, Tiwari N, Ajaymon SL, editors. Dengue Disease Spread Prediction Using Twofold Linear Regression. 2019 IEEE 9th International Conference on Advanced Computing (IACC); 2019 13–14 Dec. 2019. doi:10.1109/IACC48062.2019.8971567
11. Cuong HQ, Vu NT, Cazelles B, Boni MF, Thai KT, Rabaa MA, et al. Spatiotemporal dynamics of dengue epidemics, southern Vietnam. Emerg Infect Dis 2013;19(6):945–53. doi:10.3201/eid1906.121323
12. Wickramaarachchi WPTM, Erandi KKWH, Perera SSN. 9 - Mathematical tools and their applications in dengue epidemic data analytics. In: Dutta H, Hattaf K, editors. Advances in Epidemiological Modeling and Control of Viruses: Academic Press; 2023. p. 253–83. doi:10.1016/B978-0-32-399557-3.00014-4
13. Schmid PJ. Dynamic mode decomposition of numerical and experimental data. Journal of Fluid Mechanics 2010;656:5–28. doi:10.1017/S0022112010001217
14. Tu JH, Rowley CW, Luchtenburg DM, Brunton SL, Kutz JN. On dynamic mode decomposition: Theory and applications. Journal of Computational Dynamics 2014;1(2):391–421. doi:10.3934/jcd.2014.1.391
15. Proctor JL, Eckhoff PA. Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. Int Health 2015;7(2):139–45. doi:10.1093/inthealth/ihv009
16. Bistrian DA, Dimitriu G, Navon IM. Processing epidemiological data using dynamic mode decomposition method. AIP Conference Proceedings 2019;2164(1):080002. doi:10.1063/1.5130825
17. Fang D, Guo L, Hughes MC, Tan J. Dynamic Patterns and Modeling of Early COVID-19 Transmission by Dynamic Mode Decomposition. Prev Chronic Dis 2023;20: E95. doi:10.5888/pcd20.230089
18. Rana K, Kumari N. Application of dynamic mode decomposition and compatible window-wise dynamic mode decomposition in deciphering COVID-19 dynamics of India. Computational and Mathematical Biophysics 2023;11(1):20220152. doi:10.1515/cmb-2022-0152
19. Wagner CE, Hooshyar M, Baker RE, Yang W, Arinaminpathy N, Vecchi G, et al. Climatological, virological and sociological drivers of current and projected dengue fever outbreak dynamics in Sri Lanka. J R Soc Interface 2020;17(167):20200075. doi:10.1098/rsif.2020.0075
20. Harish V, Colón-González FJ, Moreira FRR, Gibb R, Kraemer MUG, Davis M, et al. Human movement and environmental barriers shape the emergence of dengue. Nat Commun 2024;15(1):4205. doi:10.1038/s41467-024-48465-0

21. Chen X, Moraga P. Assessing dengue forecasting methods: a comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil. Trop Med Health 2025;53(1):52. doi:10.1186/s41182-025-00723-7
22. Hasan P, Khan TD, Alam I, Haque ME. Dengue in Tomorrow: Predictive Insights From ARIMA and SARIMA Models in Bangladesh: A Time Series Analysis. Health Sci Rep 2024;7(12): e70276. doi:10.1002/hsr2.70276
23. Majeed MA, Shafri HZM, Zulkafli Z, Wayayok A. A Deep Learning Approach for Dengue Fever Prediction in Malaysia Using LSTM with Spatial Attention. Int J Environ Res Public Health 2023;20(5). doi:10.3390/ijerph20054130
24. Kakarla SG, Kondeti PK, Vavilala HP, Boddeda GSB, Mopuri R, Kumaraswamy S, et al. Weather integrated multiple machine learning models for prediction of dengue prevalence in India. Int J Biometeorol 2023;67(2):285–97. doi:10.1007/s00484-022-02405-z
25. Nandani E, Bandara A. Forecasting Dengue Outbreaks in Kalutara District using Hybrid Machine Learning Techniques. Journal of Science of the University of Kelaniya 2025;18:145–52. doi:10.4038/josuk.v18i2.8145
26. Kumarihamy PGDK. Analysis for Dengue tracking and prediction using Artificial Neural Network. Int Res J Mod Eng Technol Sci 2023;5(6):2935-2952. doi:https://www.doi.org/10.56726/IRJMETS42288
27. Alam KE, Ahmed MJ, Chalise R, Rahman MA, Mathin TT, Bhuiyan MIH, et al. Time series analysis of dengue incidence and its association with meteorological risk factors in Bangladesh. PLoS One 2025;20(8): e0323238. doi:10.1371/journal.pone.0323238