Sri Lankan Journal of Technology
Research

SLJoT
2026 Issue 01 57-66
ISSN-2773-6970

# Long Short-Term Memory-Based Multivariate Forecasting Model for Dengue: A Case Study in Sri Lanka

S. Vijitharan[1] and A. R. F.Shafana[2*]

[1]Department of Bio-science, University of Vavuniya, Sri Lanka

[2*]Department of Information and Communication Technology, South Eastern University of Sri Lanka, Sri Lanka

*Corresponding Author: arfshafana@seu.ac.lk || ORCID: 0000-0001-6926-9196

Received: 12-03-2026. * Accepted: 22-06-2026 * Published Online: 30-06-2026

**Abstract-** Dengue fever has been a significant health concern in Sri Lanka since the 1960s. The number of cases has drastically increased over the last decade. Different statistical and classical machine learning models have been proposed to forecast dengue in order to mitigate the disease from reaching its high transmission rate. More recently, the advent of neural networks has improved prognosis in an efficient manner through time forecasting of dengue data using multiple predictor variables. Understanding the strength of long short-term memory (LSTM) models in the past, this specific research has evaluated three variants of LSTM models: unidirectional LSTM, bidirectional LSTM (BiLSTM), and encoder-decoder LSTM, with the aim of predicting dengue occurrences in Sri Lanka. Weather data, including rainfall and mean temperature, were used as predictors while the efficiency of the models was assessed using RMSE. While all three models exhibited relatively good performance, the BiLSTM model significantly outperformed the other two models. This study affirms the use of LSTM for predicting dengue-like vector-borne diseases that are characterized by complex relationships with the predictors.

***Keywords**:* BiLSTM, Dengue fever, LSTM, Keras, Recurrent Neural Network

**Recommended APA Citation**

Vijitharan, S., & Shafana, A. R. F. (2026). Long short-term memory-based multivariate forecasting model for dengue: A case study in Sri Lanka. *Sri Lankan Journal of Technology, 7*(1), 57–66.

## Introduction

Dengue fever is a viral disease transmitted by female mosquitoes of the *Aedes* genus. During recent decades, this mosquito-borne disease has dramatically increased and affected mainly the developing nations. Dengue incidence has grown dramatically where WHO reported cases to have increased from 505,430 in 2000 to 5.2 million in 2019 (World Health Organization, 2019). It is also important to note that a vast majority of cases are unreported and WHO suggest that the actual number of dengue infections might be between 50 and 500 million annually (World Health Organization, 2023). Therefore, the global health sector attention has been drawn towards this tropical viral fever due to its virulence and became a public health problem. Sri Lanka is not an exceptional case for this widely spreading dengue. Notably, the highest total number of cases (105,049) have been recorded in the year 2019. Recent reports from Sri Lanka's Ministry of Health indicate that over half of the dengue cases are concentrated in Colombo, Gampaha, Kandy, Kalutara, and Jaffna (National Dengue Control Unit, 2025). In specific, the severely affected districts were Colombo and Gampaha when compared with other districts.

Presently, there are no best-fit models accepted to forecast the dengue outbreaks in Sri Lanka at the national level (Edussuriya et al., 2021). As a consequence, if a most appropriate prediction model is developed, the impacts could be minimized, and suitable preventive measures can be implemented before it emerges as an endemic situation. To predict the dengue incidences, as a basis, the relationship between influencing factors with this fever should be established. In general, transmission of dengue occurs year-round; however, a prominent transmission cycle is in connection with the rainy season, which itself varies across geographical regions. Previous studies have concluded that a strong relationship exists between dengue incidence and climatic factors (Chuang et al., 2017; Choi et al., 2016).

Numerous studies have provided strong evidence that weather factors like temperature and rainfall significantly influence dengue transmission patterns (Jain et al., 2019). While dengue transmission is continuous throughout the year in tropical regions, Ramadona et al. (2016) identified a pronounced cyclical pattern linked to the rainy season. In fact, many studies have predicted the dengue incidences using temperature and rainfall and in the context of Sri Lanka. For instance, Jayasani et al. (2021) have shown that there is a strong correlation of these two predictor variables on dengue cases. Sri Lankan rainfall patterns are influenced by Northeast and Southwest monsoonal seasons. Monsoonal seasons provide favorable breeding habitats for dengue transmitting mosquitoes (Götz et al., 2017; Morin et al., 2013; Sirisena & Noordeen, 2014). Considerably, the countries like Sri Lanka have limited in resources and technological advancements to effectively implement the control and prevention strategies of dengue. Therefore, the identification and implementation of machine learning forecasting models could reduce the burden on health and finance sectors to reduce the mortality and disease transmission.

In such a context, this study focuses on developing dengue forecasting models using long short-term memory (LSTM) which has proven efficiency in the past for forecasting time-series data with complex relationships with the predictor variables. The study employs three variants of long short-term memory (LSTM) such as unidirectional LSTM, bi-directional LSTM and encoder-decoder LSTM methods with the weather factors such as rainfall and temperature for forecasting dengue outbreak. The best-fit model of the study was identified by comparing their root mean squared error (RMSE). This paper contributes to the knowledge on the applicability of LSTM for predicting diseases like dengue.

## Related Works

Machine learning has become a promising directive in predicting dengue and its close monitoring (Sylvestre et al., 2022). Different types of predictor variables have been employed in these models depending on the availability of data types (Hoyos et al., 2021). The use of climatic, meteorological, and environmental data has shown an accuracy of 70% using Support Vector Machine (SVM) (Salim et al., 2021). SVM outperformed other machine learning models such as Decision Trees (DT), Bayesian network and Artificial Neural Network (ANN). A similar epidemic approach was undertaken by Rossi et al. (2018) where rainfall and temperature have been accounted as predictor variables. The study used boosted regression trees for predicting dengue outbreaks. Studies pertaining to dengue prognosis have prevalently used ANN Multilayer Perceptron (ANN (MLP)) as well (Davi et al., 2019; Gambhir et al., 2017,; Huang et al., 2020) . While these studies effectively predict dengue occurrence using supervised learning, further research is needed to explore methods that can forecast outbreak severity or peak timing.

Despite the fact that the prognosis is helpful, time series forecasting of dengue can better help to reinforce many counter measures to combat and control the disease from attaining its high transmission season (Mussumeci & Coelho, 2020). Time series forecasting could be defined as the process of predicting the future values based on previously observed values (Serafeim Loukas, 2020). For an infectious disease like dengue, the modeling based on a long-term time-dependent data is crucial for an accurate prediction (Xu et al., 2020). The intricate relationship between climatic factors and dengue incidence poses a challenge for traditional statistical methods, limiting their accuracy in prediction (Jiang et al., 2019; Sun et al., 2019; Zemouri et al., 2019). With the advent of neural networks, the LSTM model has proven records with less predictive errors in this kind of complex predictions (Mussumeci & Coelho, 2020; Xu et al., 2020). The models have in general used the weather factors as predictor variables and forecasted the pattern of dengue outbreaks for a specified period in future.

## Materials and Methods

### Study area

Sri Lanka is typically divided into three bioclimatic zones based on the amount of rainfall received annually. The climatic zones are wet zone (rainfall > 2,500mm), intermediate, and dry zone. The Colombo and Gampaha are particularly chosen for this study since more than 25% of dengue cases reported in these districts (National Dengue Control Unit, 2025). Over the decades, the highest dengue cases were reported in 2017 which is around 186,101. Furthermore, several endemic outbreaks have also been experienced (*Epidemiology Unit*, 2020). The mean temperature was ranged between 25-30° C which is ideal for the transmission of dengue as this temperature is within an optimal range of 20 - 35° C (Yu et al., 2011).

### Data collection

A ten-year dataset (January 2011- December 2020) on monthly dengue cases in Sri Lanka was extracted from the Epidemiology Unit of the Ministry of Health. In addition, monthly weather data, including rainfall and mean temperature for the Colombo and Gampaha districts were extracted from the statistical handbook published by the *Department of Census and Statistics (2020)* for the same period. Prior to the analysis, the dataset was cleaned and preprocessed by removing redundant records, addressing missing values, and standardizing all measurement units across variables.

## LSTM Modeling

LSTMs are a powerful type of neural network that can learn from time-series data, capturing both long-term patterns and short-term variations (Xu et al., 2020). It is an improved version of recurrent neural networks. The LSTM model employed in this study consisted of two input parameters namely the monthly mean temperature and the monthly rainfall data for a particular month. Data was first normalized using min-max scaling to the range of [0,1] in order boost the performance of the model. A one-month lag was incorporated into the input data, which were later reshaped into a time-step format with dengue cases represented as a one-dimensional output variable.

Realizing the outstanding performance of LSTM in performing time-series forecasting, this study investigates three different variants of LSTM model for predicting dengue cases namely uni-directional LSTM, bi-directional LSTM (BiLSTM), and Encoder-Decoder LSTM. A uni-directional LSTM is also known as vector output model which learns from the past data provided for training whereas BiLSTM learns the input sequence both forward and backward resulting in concatenated interpretations. Encoder-Decoder model as the name suggests, decoder utilizes the output from the encoder as an input (Abhishek, 2021).

To build the LSTM models, four hidden layers were used. The model employed the Adam variant of stochastic gradient descent for optimization and utilized the mean squared error function to evaluate training performance. For model development, eight months of data from each year were designated for training, while the remaining four months served as the testing set for evaluating model performance. The models were built using Keras, a high-level neural network library.

## Results and Discussion

This section presents the descriptive analysis of climatic variables and dengue incidence in the Colombo and Gampaha districts over the specified period of study, followed by performance evaluation of the models.

Mean temperature varied between $25^{o}$ to $35^{o}$ C, over the study period (See Figure 1). Trend of mean temperature shows a similar pattern between 2011 and 2020. High temperature was observed from the months of April to July.

Figure 2 illustrates the cumulative monthly rainfall trend of years from 2011 to 2020. High peaks of rainfall received between April and September and this was coincided with the Southwest monsoon. Given that Colombo and Gampaha lie in Sri Lanka's wet zone, they mostly obtain higher rainfall during this monsoonal season. However, comparatively less amount of rainfall is received during the remaining months of the year (Northeast monsoon).

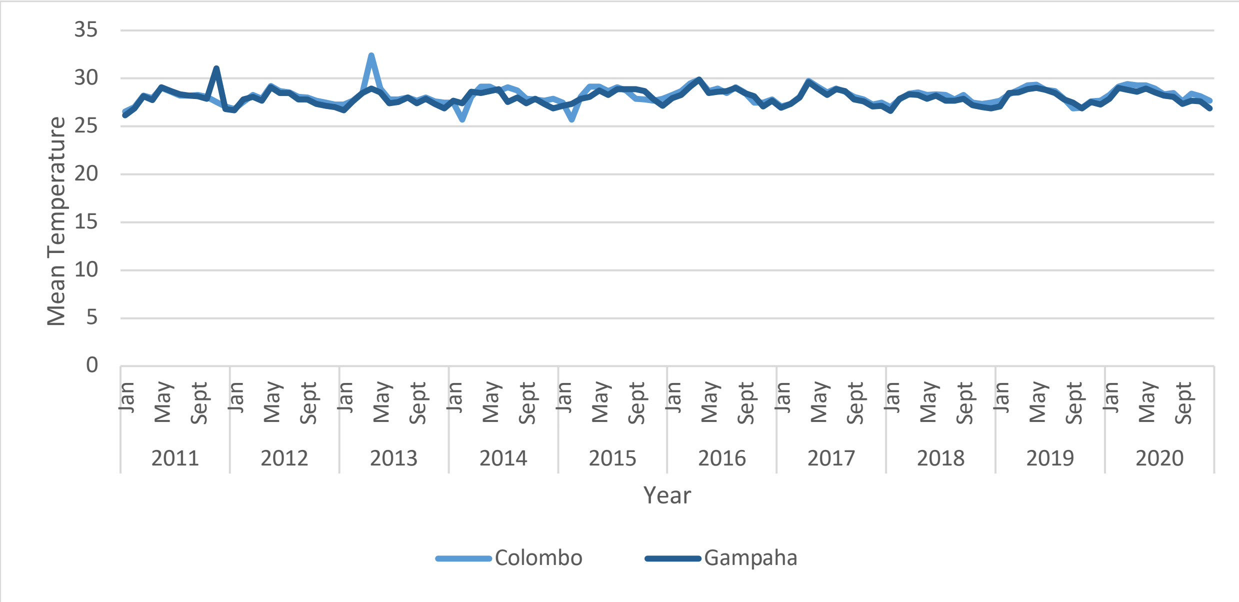


**Figure 1:** *Mean temperature variation with year*

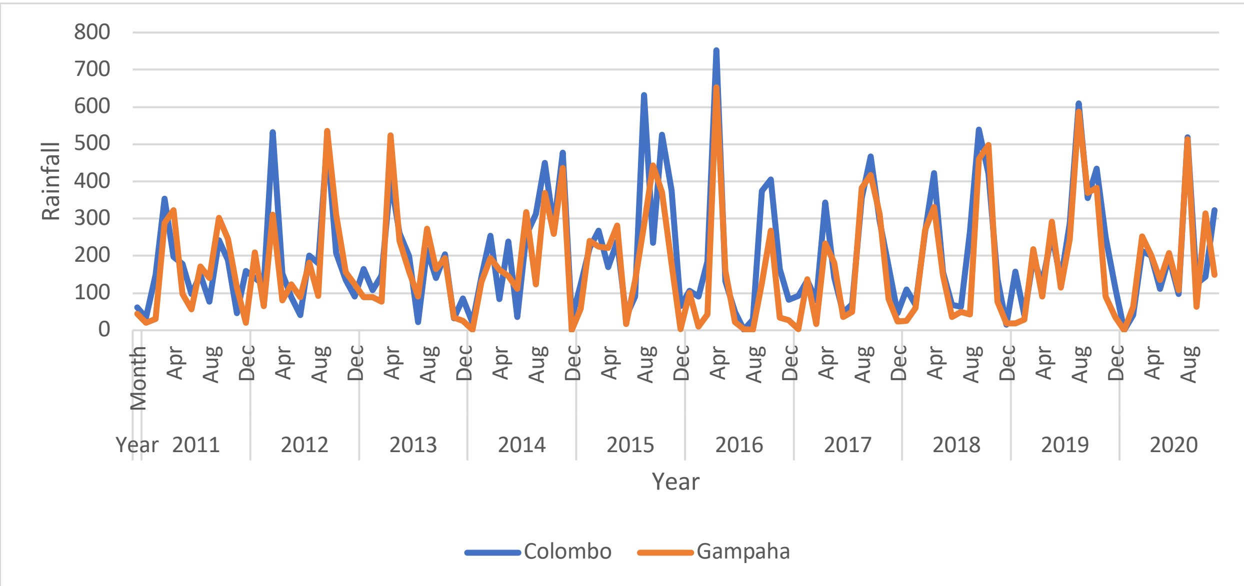


**Figure 2:** *Annual cumulative monthly rainfall*

The dengue outbreaks are depicted in Figure 3 where the high incidents are observed in the year 2017 and 2019 when compared to other years. Time series of historical dengue incidences clearly explained a greater number of cases recorded during June – July and December – January. The highest dengue case was in July 2017 (7,471 cases – Colombo; 9,039 cases - Gampaha), while lowest cases was in April 2020 (30 cases - Colombo; 36 cases - Gampaha). Interestingly, the more dengue cases observed between two peaks of rainfall. This pattern is similar to the one reported in the Kuala Lumpur city, Malaysia between 2012 and 2016 (Adnan et al., 2020). The study found that fluctuations in rainfall and temperature were associated with corresponding rises and falls in dengue cases.

Temperature is the factor that determines the vectorial capacity of Aedes mosquito. Warmer temperatures, however, could lead to a heightened risk of dengue transmission (Jing et al., 2014). High density of population, improper waste management, and lack of developed urban infrastructures could result numerous outdoor natural and artificial breeding sites. Dengue case statistics reveal these districts to have the highest burden compared to others, suggesting they may act as driver cities for transmission in the region.

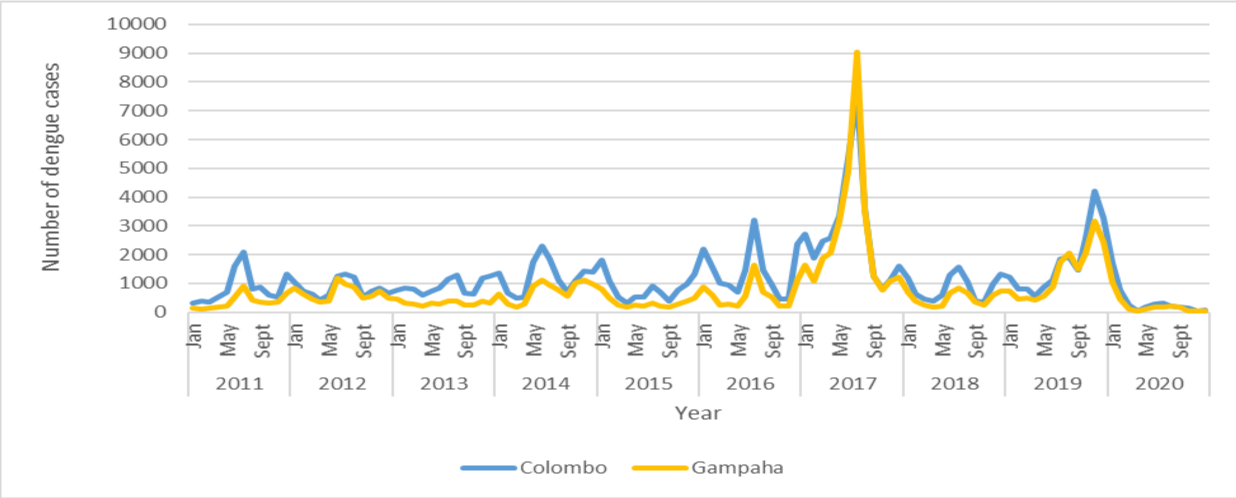


**Figure 3:** *Number of dengue cases reported from 2011 to 2020*

The result from the study suggests that the use of LSTM for the prediction of dengue cases is appropriate where the models resemble the pattern of the actual dengue cases. The standard deviation of the model was higher than the RMSE values of each of the models employed. The models studied have significantly reflected the unexpected rise and falls of the series as well. However, BiLSTM outperforms uni-directional and encoder-decoder model with the least RMSE. This is reflected in the time series graph presented in Figure 4 as well. By comparing different models, the analysis suggests that BiLSTM offers the most accurate predictions due to its effectiveness in modeling complex relationships between climate factors and dengue outbreaks.

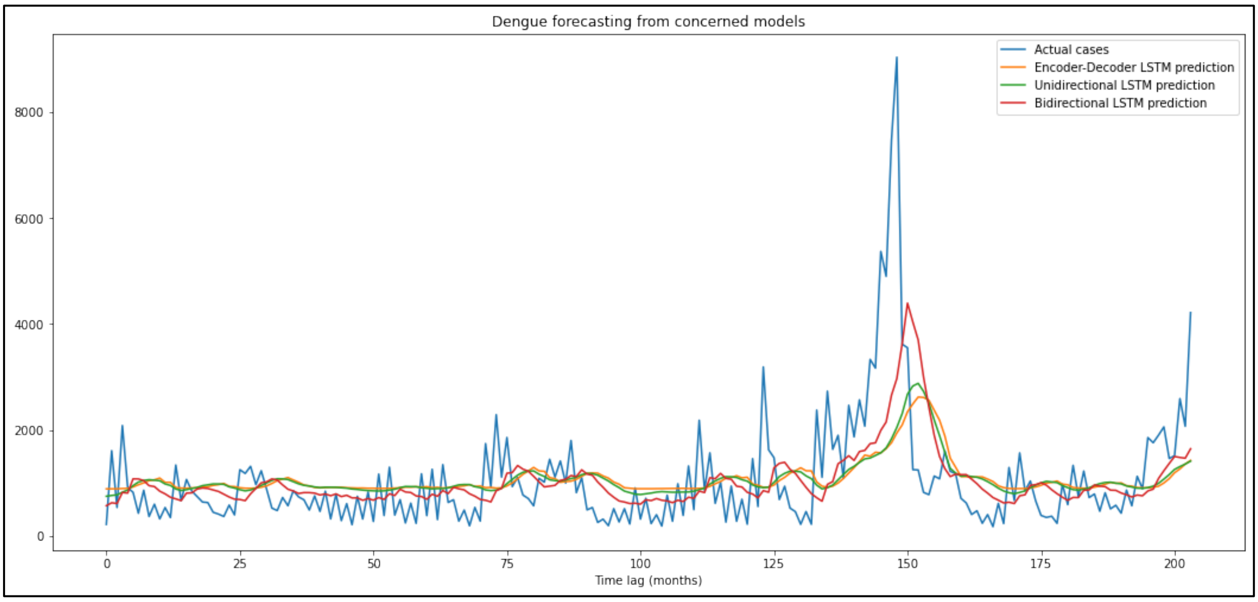


**Figure 4:** *Comparative analysis of the candidate models*

## Conclusion

This study compared three LSTM-based models for forecasting dengue cases in Sri Lanka using weather factors and dengue cases in Colombo and Gampaha districts, where the dengue cases have been highly reported in the last decade. RMSE served as the performance indicator for the LSTM models, measuring the prediction error. The models were trained on monthly data spanning 10 years, where BiLSTM outperformed other candidate models in terms of RMSE. The model has been trained only on two predictor variables, namely the mean temperature and rainfall due to the lack of available data on other weather factors, and this would be considered as a limitation of this study. To explore the possibility of even more accurate predictions, future studies could incorporate a wider range of weather factors, including wind speed, humidity, atmospheric pressure, and both minimum and maximum temperatures. The future work will also extend by incorporating Seasonal Autoregressive Integrated Moving Average (SARIMA) models to improve the predictive performance for dengue forecasts.

## References

Abhishek, J. (2021). *Multivariate LSTM* [Kaggle notebook]. Kaggle. https://www.kaggle.com/code/preatcher/multivariate-lstm/notebook

Adnan, R. A., Ramli, M. F., Othman, H. F., Asha'ri, Z. H., Ismail, S. N. S., Zaudi, M. A., Hamidon, A. K., Umar, D. A., & Samsudin, M. S. (2020). Implication of climatic factors on dengue fever in urban areas: Case study in 2012–2016. *EnvironmentAsia, 13*(3), 89–102. https://doi.org/10.14456/EA.2020.45

Choi, Y., Tang, C. S., McIver, L., Hashizume, M., Chan, V., Abeyasinghe, R. R., Iddings, S., & Huy, R. (2016). Effects of weather factors on dengue fever incidence and implications for interventions in Cambodia. *BMC Public Health, 16*(1), 1–7. https://doi.org/10.1186/s12889-016-2923-2

Chuang, T.-W., Chaves, L. F., & Chen, P.-J. (2017). Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan. *PLOS ONE, 12*(6), e0178698. https://doi.org/10.1371/journal.pone.0178698

Davi, C., Pastor, A., Oliveira, T., Neto, F. B. D. L., Braga-Neto, U., Bigham, A. W., Bamshad, M., Marques, E. T. A., & Acioli-Santos, B. (2019). Severe dengue prognosis using human genome data and machine learning. *IEEE Transactions on Biomedical Engineering, 66*(10), 2861–2868. https://doi.org/10.1109/TBME.2019.2897285

Department of Census and Statistics. (2020). *Statistical handbook*. http://www.statistics.gov.lk/

Edussuriya, C., Deegalla, S., & Gawarammana, I. (2021). An accurate mathematical model predicting the number of dengue cases in the tropics. *PLOS Neglected Tropical Diseases, 15*(11), e0009756. https://doi.org/10.1371/journal.pntd.0009756

Epidemiology Unit. (2020). *Epidemiology Unit, Ministry of Health, Sri Lanka*. http://www.epid.gov.lk/web/index.php?lang=en

Gambhir, S., Malik, S. K., & Kumar, Y. (2017). PSO-ANN-based diagnostic model for the early detection of dengue disease. *New Horizons in Translational Medicine, 4*(1–4), 1–8. https://doi.org/10.1016/j.nhtm.2017.10.001

Götz, T., Altmeier, N., Bock, W., Rockenfeller, R., Sutimin, & Wijaya, K. P. (2017). Modeling dengue data from Semarang, Indonesia. *Ecological Complexity, 30*, 57–62. https://doi.org/10.1016/j.ecocom.2016.12.010

Hoyos, W., Aguilar, J., & Toro, M. (2021). Dengue models based on machine learning techniques: A systematic literature review. *Artificial Intelligence in Medicine, 119*, 102157. https://doi.org/10.1016/j.artmed.2021.102157

Huang, S.-W., Tsai, H.-P., Hung, S.-J., Ko, W.-C., & Wang, J.-R. (2020). Assessing the risk of dengue severity using demographic information and laboratory test results with machine learning. *PLOS Neglected Tropical Diseases, 14*(12), e0008960. https://doi.org/10.1371/journal.pntd.0008960

Jain, R., Sontisirikit, S., Iamsirithaworn, S., & Prendinger, H. (2019). Prediction of dengue outbreaks based on disease surveillance, meteorological, and socio-economic data. *BMC Infectious Diseases, 19*(1), Article 272. https://doi.org/10.1186/s12879-019-3874-x

Jayasani, C., Ekanayake, P., Dammalage, P., Herath, V., Sarathchandra, S., Ekanayake, J., Godaliyadda, R., & Dharmaratne, S. (2021). *Limited data forecasting for dengue propagation*. In *Proceedings of the 2021 IEEE International Conference on Industrial and Information Systems*. https://doi.org/10.1109/ICIAfS52090.2021.9606032

Jiang, S., Xiao, R., Wang, L., Luo, X., Huang, C., Wang, J.-H., Chin, K.-S., & Nie, X. (2019). Combining deep neural networks and classical time series regression models for forecasting patient flows in Hong Kong. *IEEE Access, 7*, 154865–154879. https://doi.org/10.1109/ACCESS.2019.2936550

Jing, L.-H., Hans, S., Annelies, W.-S., & Joacim, R. (2014). Vectorial capacity of *Aedes aegypti*: Effects of temperature and implications for global dengue epidemic potential. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0089783

Morin, C. W., Comrie, A. C., & Ernst, K. (2013). Climate and dengue transmission: Evidence and implications. *Environmental Health Perspectives, 121*(11–12), 1264–1272. https://doi.org/10.1289/ehp.1306556

Mussumeci, E., & Coelho, F. C. (2020). Machine-learning forecasting for dengue epidemics: Comparing LSTM, Random Forest, and Lasso regression. *medRxiv*. https://doi.org/10.1101/2020.01.23.20018556

National Dengue Control Unit. (2025). *National Dengue Control Unit*. https://www.dengue.health.gov.lk/

Ramadona, A. L., Lazuardi, L., Hii, Y.-L., Holmner, Å., Kusnanto, H., & Rocklöv, J. (2016). Prediction of dengue outbreaks based on disease surveillance and meteorological data. *PLOS ONE, 11*(3), e0152688. https://doi.org/10.1371/journal.pone.0152688

Rossi, G., Karki, S., Smith, R. L., Brown, W. M., & Ruiz, M. O. H. (2018). The spread of mosquito-borne viruses in modern times: A spatio-temporal analysis of dengue and chikungunya. *Spatial and Spatio-Temporal Epidemiology, 26*, 113–125. https://doi.org/10.1016/j.sste.2018.06.002

Salim, N. A. M., Wah, Y. B., Reeves, C., Smith, M., Yaacob, W. F. W., Mudin, R. N., Dapari, R., Sapri, N. N. F. F., & Haque, U. (2021). Prediction of dengue outbreak in Selangor, Malaysia using machine learning techniques. *Scientific Reports, 11*, Article 939. https://doi.org/10.1038/s41598-020-79193-2

Loukas, S. (2020). *Time-series forecasting: Predicting stock prices using an LSTM model*. Towards Data Science. https://towardsdatascience.com/lstm-time-series-forecasting-predicting-stock-prices-using-an-lstm-model-6223e9644a2f

Sirisena, P. D. N. N., & Noordeen, F. (2014). Evolution of dengue in Sri Lanka: Changes in the virus, vector, and climate. *International Journal of Infectious Diseases, 19*, 6–12. https://doi.org/10.1016/j.ijid.2013.10.012

Sun, D., Wang, M., & Li, A. (2019). A multimodal deep neural network for human breast cancer prognosis prediction by integrating multidimensional data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16*(3), 841–850. https://doi.org/10.1109/TCBB.2018.2806438

Sylvestre, E., Joachim, C., Cécilia-Joseph, E., Bouzillé, G., Campillo-Gimenez, B., Cuggia, M., & Cabié, A. (2022). Data-driven methods for dengue prediction and surveillance using real-world and big data: A systematic review. *PLOS Neglected Tropical Diseases, 16*(1), e0010056. https://doi.org/10.1371/journal.pntd.0010056

World Health Organization. (2019). *Dengue and severe dengue*. https://www.who.int/health-topics/dengue-and-severe-dengue

World Health Organization. (2023). *Disease outbreak news: Dengue*. https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON498

Xu, J., Xu, K., Li, Z., Meng, F., Tu, T., Xu, L., & Liu, Q. (2020). Forecast of dengue cases in 20 Chinese cities based on the deep learning method. *International Journal of Environmental Research and Public Health, 17*(2), 453. https://doi.org/10.3390/ijerph17020453

Yu, H.-L., Yang, S.-J., Yen, H.-J., & Christakos, G. (2011). A spatio-temporal climate-based model of early dengue fever warning in southern Taiwan. *Stochastic Environmental Research and Risk Assessment, 25*(4), 485–494. https://doi.org/10.1007/s00477-010-0417-9

Zemouri, R., Devalland, C., Valmary-Degano, S., & Zerhouni, N. (2019). Neural networks: A future in pathology? *Annales de Pathologie, 39*(2), 119–129. https://doi.org/10.1016/j.annpat.2019.01.004