Pertanika J. Sci. & Technol. 34 (2): 975 - 1004 (2026)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Genetic Algorithm-optimised Hybrid Framework Integrating Statistical Forecasting Models and Extreme Learning Machine for Dengue Case Forecasting in the United States of America

Dipankar Das[1]*, Arijit Chakraborty[2], Sajal Mitra[3], and Soumyabrata Nag[4]

[1]*Department of Computer Application and Science, Institute of Engineering & Management, University of Engineering and Management Kolkata, 700091 West Bengal, India*
[2]*Department of BCA, The Heritage Academy, 700107 West Bengal, India*
[3]*Department of Computer Science and Business Systems, Heritage Institute of Technology, 700107 West Bengal, India*
[4]*Department of Microbiology, All India Institute of Medical Sciences, Nagpur, 441108 Maharashtra, India*

ABSTRACT

Dengue, a vector-borne disease, has become a global calamity. A robust early warning mechanism could assist authorities in mitigating this hazard. The objective of the present study is to propose a dengue forecasting model employing a high-throughput dataset for eleven months of dengue progression in the United States of America (USA). The methodology integrates the efficacy of well-renowned statistical methods, i.e., Auto ARIMA, Auto ETS, and machine-intelligent methods, namely Extreme Learning Machine (ELM). The final architecture of the proposed approach is realised by using an evolutionary algorithm, i.e., Genetic Algorithm (GA), which demonstrates improvement of model performance and enhances out-of-sample predictive accuracy. Our proposed model obtained a Mean Error (ME) of 16.66, a Root Mean Square Error (RMSE) of 66.84, a Mean Absolute Error (MAE) of 55.22, a Mean Percentage Error (MPE) of 0.30, and a Mean Absolute Percentage Error (MAPE) of 0.79. Our approach convincingly identifies linearity, seasonality, and nonlinearity imprints of the dengue progression in the USA. The model outperforms nineteen other techniques, including seven traditional, three Artificial Intelligence (AI) -based, a generative AI, and eight hybrid methods. The significance of the findings lies in a rigorous validation technique, specifically non-parametric tests, underscoring the practicability of the proposed model in dealing with noisy or incomplete data environments inherent in coping with health-related time-series challenges.

ARTICLE INFO

Article history:
Received: 09 June 2025
Accepted: 11 March 2026
Published: 30 April 2026

DOI: https://doi.org/10.47836/pjst.34.2.16

E-mail addresses:
dipankr.das@gmail.com (Dipankar Das)
arijitchakraborty.besu@gmail.com (Arijit Chakraborty)
sajal.mitra@heritageit.edu (Sajal Mitra)
soumyabratanag@gmail.com (Soumyabrata Nag)
* Corresponding author

e-ISSN: 2231-8526

Potential applications of the research could be the development of a robust early warning mechanism that empowers public health efforts with enhanced epidemiological surveillance.

Keywords: Dengue, extreme learning machine, genetic algorithm, hybrid model, model optimisation, time-series forecasting

INTRODUCTION

Dengue is a vector-borne viral disease transmitted by Aedes mosquitoes. It is common in tropical and subtropical regions of Asia, Africa, and the Americas. Dengue often exhibits spatial heterogeneity and variations in distribution. The clinical symptoms of dengue range from ordinary nausea, vomiting, body aches, and fever to critical conditions such as dengue shock syndrome. It becomes a significant challenge in developing anti-dengue immunisation programmes, as highlighted by Mobin (2024), due to the presence of multiple serotypes in dengue and tropical and subtropical climatic conditions. Patil and Pandya (2021) established a positive link between weather fluctuations and dengue progression in their work. The epidemiological forecasting models developed by Mahanty et al. (2024) and Tomov et al. (2023) portrayed disease outbreaks over time.

The forecasting models are primarily classified as statistical, Artificial Intelligence (AI)-based, and hybrid. Their performance is often limited by non-linearity and interpretability issues. Often, relying on a single approach may restrict our insight into the disease outbreaks. It results in sketchy support for healthcare authorities in combating dengue. This work proposes a novel approach. It integrates Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing State Space Model (ETS), and a Genetic Algorithm (GA) optimised Extreme Learning Machine (ELM) method to portray the progression of dengue in the USA. This novel approach is expected to furnish more precise and timely dengue outbreak forecasts, allowing for more efficient public health measures and, eventually, lessening the disease's impact on afflicted communities.

LITERATURE REVIEW

In America, Sanchez-Gendrizet et al. (2022) applied the Long Short-Term Memory (LSTM) method to predict dengue outbreaks. They employed weekly data from 2016 to 2019. The proposed model was also used to predict the dengue prevalence rates and mosquito egg density in Natal, Brazil. The model achieved a high correlation. The values were between 0.87 and 0.92. The model was used to predict dengue infections and outbreaks 4 to 6 weeks in advance.

Saleh and Baiwei (2021) also used LSTM networks and Support Vector Regression (SVR) to forecast dengue infections, using satellite-based weather data for the years 1990-2008 in San Juan, Puerto Rico. In this regard, the LSTM performed better than the SVR.

It (the LSTM) achieved R^2 of 0.75, Mean Absolute Error (MAE) 8.76, Mean Squared Error (MSE) 245.61 and Root Mean Squared Error (RMSE) 15.67.

Hasan et al. (2024) proposed time series models pertaining to several contexts of healthcare planning and resource allocations. They also used external factors and interactions for dengue prevalence in Bangladesh. The ARIMA (1,1,1) had an MAE of 1266.08. Seasonal ARIMA (SARIMA) had MAE, MSE and RMSE values of 1246.442, 531.837 and 5175.86, respectively, for predicting the dengue prevalence in the forecasted year 2024.

Mohammed and Rahman (2023) employed time series analysis to forecast the monthly prevalence of dengue disease in Bangladesh. They specifically applied ARIMA and SARIMA models. The SARIMA (1,0,0) (1,1,1) model was the best-fit model with an RMSE of 9073.657, which was used to forecast the rate of dengue infection period from January 2022 to December 2022.

Sutriyawan et al. (2024) performed an ARIMA model study to predict the dengue incidence by analysing the data from 2014 to 2023, in Bandung, Indonesia. The best-fitted model is ARIMA (3,0,3), having Mean Absolute Percentage Error (MAPE) 33.34, RMSE 10.0493 and MAE 73.42905. Their proposed model predicted a peak for dengue cases in September 2024, estimated at 320 cases.

Rangarajan et al. (2019) forecasted dengue disease in Singapore, Taiwan, Thailand, Brazil, and Mexico using a sparse representation from electronic health records, Google trends, and time series data. Their proposed method, Autoregressive Likelihood Ratio (i.e., ARLR), achieved an average RMSE of 0.514, MAE of 0.531, and MAPE of 0.550. It forecasted dengue incidence in the USA four weeks ahead, which accomplished a 26% average reduction in RMSE, 21% in MAE, and 6% in MAPE for existing forecasting techniques.

Lu et al. (2024) proposed a Susceptible-Infected mosquito vectors and Susceptible-Infected-Recovered human hosts model (SI-SIR). According to expected climate conditions and mosquito bites in Selangor, Malaysia, they concluded that their model outperformed Multiple Linear Regression (MLR) and was better than LSTM at predicting climate. Forecasts from the model extended to 60 weeks. Before the Movement Control Order (MCO) in Malaysia, it had an MAPE of 13.97 and an MAE of 204.36.

Torres et al. (2025) employed an LSTM network-based model on Monte-Carlo dropout to classify and distinguish between healthy and infected Aedes Aegypti mosquitoes, the dengue virus's primary vectors. The proposed model, a 2-layer Bi-LSTM with 256 neurons, detected approximately 5% of uncertainty data from wing beat signals and achieved 94.87% accuracy, 94.87% precision, 94.87% recall, and 94.87% F1-score, which outperformed the K-Nearest Neighbours (KNN), Support Vector Machine (SVM), 2-layer LSTM with 64 neurons, and 2-layer LSTM with 256 neurons models.

Using weather-related datasets, Baker et al. (2021) employed thirteen machine-learning approaches to forecast weekly dengue incidences in San Juan, Puerto Rico, and in the city of Iquitos, Peru. The authors considered the dengue progression data of San Juan from 1999 to 2008 and the dengue progression data from 2000 to 2010 for Iquitos, which indicated that the Poisson Regression Model (PRM), Negative Binomial Regression Model (NBRM), and Random Forest (RF) models achieved the lowest MAE values of 25.6, 25.8, and 26.6, respectively.

Mendoza Chirinos (2024) developed a spatiotemporal neural network model that combines LSTM and Graph Convolution Network (GCN) layers to predict dengue outbreaks. The author used climatic and socioeconomic data assembled from meteorological records. The author's model achieved impressive results with an MSE of 11.0948 for the combination of cases, temperature, and precipitation features set.

The Artificial Neural Networks (ANN), Convolution Neural Networks (CNN), and LSTM models evaluated the effectiveness for diagnosing dengue cases in the study (Dhaked et al., 2025) using monthly surveillance and meteorological data from 2015 to 2019 in Jaipur, Rajasthan, India. Their proposed one-dimensional CNN (1-DCNN) model demonstrated significantly better accuracy than the ANN and LSTM models, achieving an MAE of 31.49, MSE of 3187.43, and RMSE of 56.45 in predicting dengue cases.

In summary, the investigation of dengue propagation and its effects primarily utilised time series and machine learning models, such as ARIMA, LSTM, SVM, ANN, CNN, and KNN. The performance of these models was evaluated using various statistical measures, including RMSE, MSE, MAPE, and MAE.

Forecasting dengue cases is not limited to possible earlier identification. It is also related to various aspects of potential solutions associated with dengue diseases, related to the protection of human health and the safety of social life. This leads to finding a suitable and relevant solution for identifying and managing a possible multi-objective outlook. Thus, the role of computational optimisation techniques is paramount. It presents substantial options for enhancing the forecasting approach for dengue cases and for building a decisive forecasting mechanism. Table 1 presents a statistical summary of these studies.

Motivation and Research Gap

Dengue fever is another major issue of global health that has been made worse by climate change. The use of AI models has been found to be effective in the prediction of dengue outbreaks. The creation of an AI-based prediction system would enable the health officials to receive crucial warnings and be better positioned to control an outbreak and ultimately save lives.

Wang et al. (2023) emphasised the fact that real-life time-series data can be mixed with both linear and non-linear components. They are also non- stationary because of the trends and seasonality. Pereira da Veiga et al. (2024) emphasised that the ARIMA is an

Table 1
Observations of the past investigations for dengue analysis

Author	Dengue disease investigation methods		Statistical measures used				
	ML-based	Time series	RMSE	MSE	MAE	MAPE	R^2
Sánchez-Gendriz et al. (2022)	✔	✔	-	-	-	-	✔
Saleh and Baiwei (2021)	✔	✔	✔	✔	✔	-	✔
Hasan et al. (2024)	✔	✔	✔	✔	✔	-	-
Mohammed and Rahman (2023)	-	✔	✔	-	-	-	-
Mendoza Chirinos (2024)	✔	✔	-	✔	-	-	-
Sutriyawan et al. (2024)	-	✔	✔	-	✔	✔	-
Dhaked et al. (2025)	✔	✔	✔	✔	✔	-	-
Rangarajan et al. (2019)	✔	✔	✔	-	✔	✔	-
Lu et al. (2024)	✔	✔	-	-	✔	✔	-
Baker et al. (2021)	✔	✔	-	-	✔	-	-
Flores et al. (2026)	✔	-	✔	-	-	✔	-
Bunprom et al. (2026)	-	✔	-	-	-	✔	-

effective model that can capture linear elements. Equally, the ETS can overcome trends and seasonality. In their work, Beck et al. (2025) also indicated that the ETS can face difficulties in addressing complex non-linearity.

Artificial Intelligence models can spot complex nonlinear trends. Wu et al. (2025) demonstrated that an ELM, which is an artificial intelligence model, can be applied successfully to non-linear trends of time-series data. However, overall, the models that are based on AI tend to be uninterpretable, and they need a significant amount of hyperparameter tuning to maximise their performance. Those problems indicate shortcomings of the existing models of dengue forecasting using the data of time-series dengue cases.

We found out that hybrid models are improved over individual models in most instances. They integrate numerous strengths and boost forecasting accuracy and strength. Considering the benefits of the ARIMA, ETS and ELM forecasting methods, we suggest a hybrid forecasting model that combines Auto ARIMA, Auto ETS, and ELM to predict dengue.

This combined strategy is expected to improve the out-of-sample predictive performance, thus improving the accuracy and reliability of dengue predictions. Moreover, a GA is used to optimise the ELM model. Though the application of GA is common in multi-objective optimisation, the ability of the technique to narrow down the number of hidden neurons in ELM to predict dengue has not been explored. ELM architecture is composed of one hidden layer with a very large number of neurons, the number of which is computed to give optimal model complexity and performance. GA systematically recognises the most appropriate number of concealed neurons in ELM by exploring a large solution space in a more systematic manner to reduce forecasting errors, which promotes a decrease in the costs of computation and the general complexity of the model. However, the evolution of the proposed model based on the hybridisation of different forecasting methods with the help of GA should deal with the following research issues:

- How effectively can the GA-tuned hybrid model, integrating Auto ARIMA, Auto ETS, and ELM, improve dengue case predictions in the USA compared to standalone models regarding accuracy and reliability?
- Can GAs optimise the hidden neurons in the GA-tuned ARIMA-ETS-ELM hybrid model for better dengue forecasting?
- What are the contributions of Auto ARIMA, Auto ETS, and ELM in enhancing the accuracy of the GA-tuned ARIMA-ETS-ELM model for dengue forecasting?

Accordingly, the objectives of the present study are outlined as follows:

- Analysing dengue case trends in the USA by exploring historical data.
- Creating a hybrid forecasting model that combines two automated statistical methods, Auto ARIMA and Auto ETS, with an AI approach, specifically ELM, to enhance out-of-sample predictive performance.
- Utilising GA to optimise the number of hidden neurons in ELM ensures an optimal balance between model complexity and forecasting accuracy.
- Comparing the forecasting accuracy of the proposed hybrid model with advanced automatic time-series forecasting models using benchmark metrics.
- Evaluating the reliability of the GA-tuned ARIMA-ETS-ELM model through statistical tests on its forecasts for effective dengue surveillance and early warning systems in public health.

METHODOLOGY

Our product is a new hybrid system that includes the ARIMA, ETS, and ELM models, which are predicted by the GA optimisation to predict cumulative dengue cases in the United States. To make use of a general-purpose GA to maximise the number of hidden nodes, which is a real-valued parameter. To systematically optimise the fitness of the candidate solutions, the GA utilised the conventional evolutionary operators, namely selection, crossover and mutation. The ARIMA-ETS-ELM model used to tune GA is in two layers. The former layer is the automated statistical modelling and involves two methods: ARIMA and ETS. The second layer involves the use of an AI-based forecasting method with an ELM. Here, a GA is used to optimise the ELM regarding the number of hidden nodes, thus making it easier to generate correct forecasts. Figure 1 displays a detailed description of the methodology applied in this research.

The out-of-sample forecasting accuracy of the suggested GA-optimised hybrid approach, which blends ARIMA, ETS, and ELM, is assessed. The following metrics are used in this study: Mean Error (ME), MAE, Mean Percentage Error (MPE), MAPE, and RMSE. ME is the average signed difference between forecasted and observed values, while RMSE emphasises more significant errors, making it sensitive to outliers. It helps assess the stability of a model's predictions. MAPE treats all errors uniformly. It is less sensitive to outliers compared to RMSE. The interpretation of MAPE is proper, where a model with MAPE ≤ 10 per cent is considered highly accurate (Yadav & Nath, 2019).

We compared its performance against nineteen automated techniques. Three component models ARIMA (Sucipto et al., 2025), ETS (Li, 2024), and ELM (Rahman et al., 2024); four traditional statistical methods Single Exponential Smoothing (SES) (Saragih et al., 2024), Holt's Linear Trend (HLT) (Ariyanto & Nugraha, 2024), Holt-Winters' (HW) (Santosa et al., 2024), Seasonal-Trend Decomposition using LOESS (STL), which separates a time series into seasonal, trend, and residual components (Yunisa & Siregar, 2023); three AI-based approaches – Neural Network Autoregressive (NNAR) (Ahmar & Boj, 2021), Multilayer Perceptron (MLP) (Chakraborty et al., 2024), and SVR (Parreño & Anter, 2024); one generative AI technique namely, Moirai (Woo et al., 2024); and eight hybrid models namely, GA-tuned ARIMA-ELM, GA-tuned ETS-ELM, ARIMA-ANN using equal weight, ETS-ANN using equal weight, and ARIMA-ETS-ANN using equal weight (Perone, 2021), ARIMA-ANN using CV-error (Talkhi et al., 2021), STL-ETS and STL-ARIMA (Hightower et al., 2024) are used. In these eight ANN-based hybrids, both GA-tuned techniques are applied for the ablation study to identify their contributions to the hybrid model.

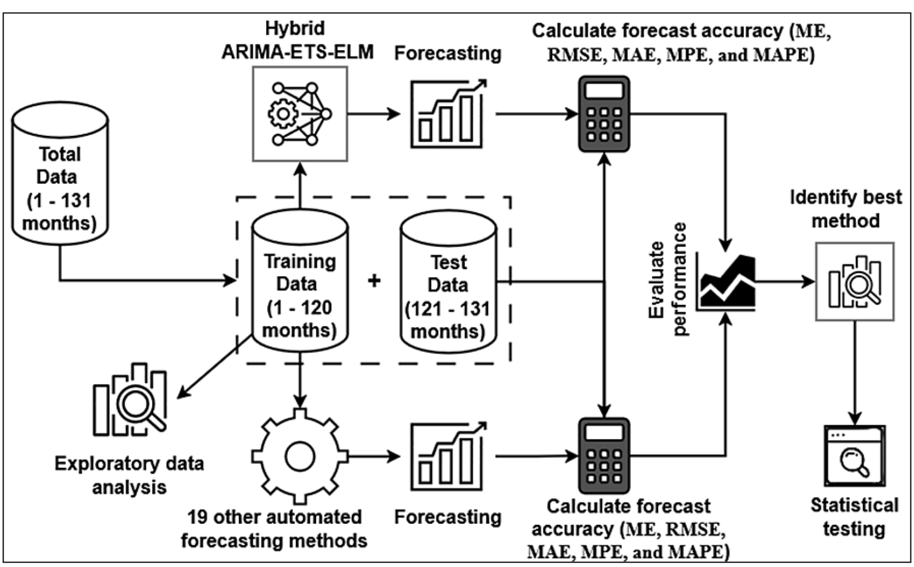

Figure 1. Overview of the research methodology

To assess the robustness of the forecasting performance of the identified best method between these twenty approaches, i.e., the GA-tuned ARIMA-ETS-ELM approach and the other nineteen techniques, we conducted statistical significance testing of the best method using the Mann-Whitney U test (Xiong et al., 2022) and the Wilcoxon Signed Rank test (Ahmadoni, 2025).

Dataset Construction

We obtained historical and high-throughput data regarding dengue cases from a publicly accessible open data source in the United States. The dataset comprises monthly dengue case statistics over 131 months, covering January 2014 to November 2024 (World Health Organisation, 2025). We utilised the monthly data to calculate the cumulative dengue cases in the USA. This cumulative case data is then split into a training set for model development and a test set for evaluating model performance. Comprehensive exploratory data analysis is conducted on the training set data to identify the trends, i.e., fitting trend lines (Das & Chakrabarti, 2021), seasonality using the Wilcoxon Signed-Rank test on the Buys-Ballot table's row variances (Nwogu et al., 2016), non-linearity employing Brock-Dechert-Scheinkman, i.e., BDS test, which is used detect nonlinear patterns in time series data (Jirasakuldech et al., 2023), and stationarity utilizing Augmented Dickey-Fuller, i.e., ADF test, which checks whether a time series is stationary (Ivanovski & Ivanovska, 2024) of the cumulative dengue cases.

The original dataset contained monthly dengue cases for each month. This monthly data was converted into cumulative dengue cases to capture the progressive trend of disease spread. Here, each observation represents the total number of cases up to that month (i.e., cumulative sum up to that month). This transformation smooths short-term fluctuations and highlights the long-term growth pattern of dengue incidence.

Cumulative dengue cases were modelled to highlight long-term outbreak growth, smoothing reporting irregularities and enabling hybrid model evaluation under mixed linear–nonlinear non-stationary conditions. It supports medium-term planning despite the monotonic nature of the cumulative series. This transformation can inflate the denominator, potentially reducing percentage-based errors like MAPE; therefore, model evaluation also relied on scale-dependent metrics (RMSE, MAE, ME) to ensure comparative consistency across all nineteen models.

Data Description

Figure 2 illustrates the cumulative data on dengue cases in the USA over time. The 131-month dataset is partitioned into a 120-month training set (1-120 months) and an

11-month test set (121-131 months). The 11-month test set, representing genuinely unseen future observations, serves as a real-time forward holdout that preserves temporal ordering, captures post-trend stabilisation, and reflects realistic one-year public-health forecasting horizons. Future work may incorporate rolling-origin evaluation for additional robustness. Table 2 provides the summary statistics and test results for the training set data of the cumulative dengue cases in the USA.

Table 2 demonstrates that the training dataset of cumulative dengue cases in the USA consists of 120 months of data without any missing values. The mean of 3435 cumulative cases has a median of 3393 and a standard deviation of 1823, indicating considerable variability. The minimum and maximum recorded cumulative cases are 54 and 6278, respectively.

The very high R^2 value of 0.976 and adjusted R^2 of 0.976 in the linear trend analysis and a statistically significant F-test with p-value < .001 confirm a strong linear trend in the cumulative cases. In contrast, the seasonality in the cumulative cases is not statistically significant, as indicated by the p-value of 0.7695 of the Wilcoxon Signed-Rank test. The BDS-test with p-value < 0.001 for both m=2 and m=3 affirms the presence of significant non-linearity in the cumulative dengue cases data. The cumulative dengue cases data is non-stationary, as suggested by the p-value of 0.2965 of the ADF test.

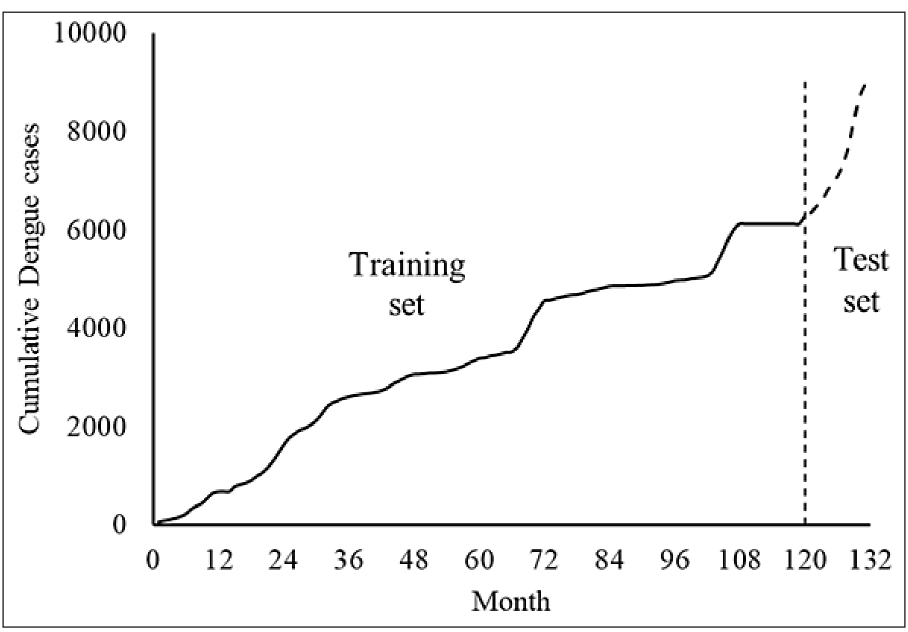

Figure 2. Cumulative dengue cases in the USA

Table 2
Summary statistics and test results for cumulative dengue cases in the USA

Statistic	Value
N (length of training set)	120
Missing	0
Mean	3435
Median	3393
Standard deviation	1823
Minimum	54
Maximum	6278
R^2 of linear trend line	.976
Adjusted R^2 of the linear trend line	.976
Significance of the F-test of the linear trend line	< .001
Significance of the Wilcoxon Signed-Rank test	.7695
Significance of the BDS test	<.001 (or embedding dimension of m=2), <.001 (or embedding dimension of m=3)
Significance of the ADF test	.2965

Proposed Model Development

Table 2 shows that the cumulative dengue case training data displays a strong linear trend. However, it also shows non-linearity and non-stationarity. These characteristics indicate that standalone models such as ARIMA and ETS may not fully capture the complex patterns within the data. Conversely, an ELM can capture non-linearity. Nevertheless, given the strong linear trend present in the data, ELM may struggle to model this effectively without incorporating trend modelling. A hybrid approach that combines ARIMA, ETS, and ELM may be more effective. ARIMA and ETS handle the trend component effectively, and ELM captures the non-linear dependencies. Figure 3 illustrates the integration of statistical (ARIMA and ETS) and AI (ELM) techniques in the proposed GA-tuned ARIMA-ETS-ELM approach forecasting framework.

The proposed hybrid GA-tuned ARIMA-ETS-ELM used the lagged values of the dengue time series and the fitted values obtained from ARIMA and ETS models as input to the ELM. During training, the fitted values from ARIMA and ETS on the training dataset served as external regressors. It enables the ELM to learn linear through ARIMA and trend, level, and seasonality through ETS components, along with the nonlinear components through ELM. During forecasting, the ARIMA and ETS models first generate forecasts for the test period. These forecasted values are then supplied as external inputs to the ELM to produce the final hybrid predictions. This ensures consistency between the training and forecasting phases. Leakage is prevented as follows: during training,

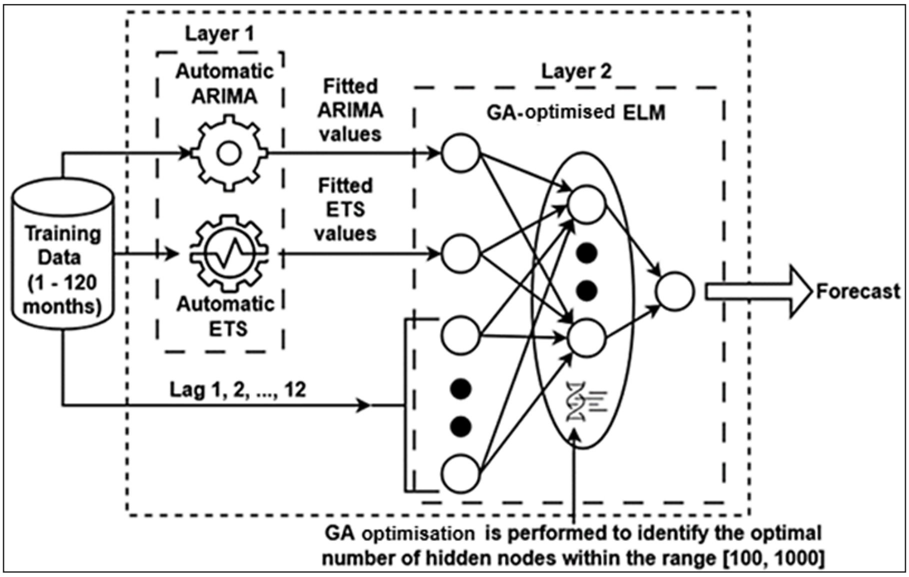

Figure 3. Proposed GA-tuned ARIMA-ETS-ELM hybrid approach

ELM uses fitted values generated strictly from the training set; during forecasting, ARIMA and ETS first generate forecasts for the test period, and these forecasted, not observed values, are then fed into ELM. Thus, no actual future observed test values enter the model during prediction, ensuring temporal integrity.

Mathematical Foundations of the Proposed Methodology

Considering the training data of the time series as $\{s_t\}_{t=1}^{N}$ (where *N* denotes the length of the training set). The methodology consists of two layers, as outlined below.

Working of Layer 1: The time series $\{s_t\}_{t=1}^{N}$ is initially modelled using automatic ARIMA and ETS techniques. The fitted values obtained from ARIMA and ETS are:

$$\hat{s}_t^{ARIMA} = \alpha_0 + \sum_{I=1}^{P} \alpha_i s_{t-i} + \sum_{j=1}^{q} \gamma_j e_{t-j} \quad [1]$$

$$\hat{s}_t^{ETS} = f(L_t, B_t, S_t, \epsilon_t) \quad [2]$$

In Equation 1, $\hat{s}_t^{ARIMA}$ represents the ARIMA fitted value at time *t*, α_0 is the intercept, α_i are autoregressive coefficients, γ_j are moving average coefficients, s_{t-i} are either lagged values or differenced values (if differencing is applied), and e_{t-j} denotes past error terms.

The order of the AR and MA components is p and q, respectively. In Equation 2, the ETS fitted values (at time t), $\hat{s}_t^{ETS}$, are derived from a combination of levels (L_t), trend (B_t), seasonality (S_t) and error (ϵ_t) components. ϵ_t affects the components, and not directly the fitted values. The function $f(\cdot)$ captures the relationship among these components. $f(\cdot)$ depends on the ETS model structure.

Working of Layer 2: In this layer, the ELM is employed to refine forecasting. The input to the ELM consists of the fitted values from ARIMA and ETS, along with the last 12 lagged values of the time series. The 12-lag structure reflects annual seasonality in monthly data, common epidemiological lags in vector-borne diseases, and standard practice in monthly time-series modelling. Future sensitivity analysis with alternative lags (6, 9, 18 months) could further evaluate robustness. Defining the lagged vector (v_t) as given in Equation 3:

$$v_t = [s_{t-1}, s_{t-2}, \dots, s_{t-12}] \tag{3}$$

The combined input vector (u_t) to ELM at time t as given in Equation 4:-

$$u_t = [v_t, \hat{s}_t^{ARIMA}, \hat{s}_t^{ETS}] \tag{4}$$

The ELM processes this input using a single hidden layer with L hidden nodes, where L is determined using a GA. The GA optimises L within the range [100, 1000], ensuring optimal performance. Each hidden node j (where $j = 1, 2, \cdots, L$) applies a transformation as follows:

$$h_j(u_t) = g\left(w_j^T u_t + b_j\right) \tag{5}$$

In Equation 5, w_j is the weight vector, w_j^T denotes its transpose, b_j is the bias, and $g(\cdot)$ is the activation function.

The final output of the ELM model is given by:

$$\hat{s}_t^{ELM} = \sum_{j=1}^{L} \beta_j h_j(u_t) \tag{6}$$

In Equation 6, β_j are the output weights connecting the hidden layer to the forecasted value and $h_j(u_t)$ is the activation (output). These weights are computed using the Moore-Penrose pseudo-inverse method:

$$\beta = H^{\dagger}T \tag{7}$$

In Equation 7, H is the hidden layer output matrix, as given in Equation 8, $H^{\dagger}$ is its Moore-Penrose pseudo-inverse, and β the output weight vector is defined in Equation 9, and T is the training target matrix, as given in Equation 10:

$$H = \begin{bmatrix} g(w_1^T u_1 + b_1) & \cdots & g(w_L^T u_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1^T u_N + b_1) & \cdots & g(w_L^T u_N + b_N) \end{bmatrix}_{N \times L} \tag{8}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}_{L \times 1} \tag{9}$$

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_N \end{bmatrix}_{N \times 1} \tag{10}$$

To enhance robustness, the proposed approach trains 20 different networks and aggregates their forecasts using the median operator:

$$\hat{s}_t^{Hybrid} = median\left(\hat{s}_t^{ELM(1)}, \hat{s}_t^{ELM(2)}, \dots, \hat{s}_t^{ELM(20)}\right) \tag{11}$$

In Equation 11, $\hat{s}_t^{ELM(k)}$ represents the forecast from the *k*-th ELM network.

The complete GA-tuned ARIMA-ETS-ELM approach can be mathematically summarised as:

$$\hat{s}_t^{Hybrid} = median\left(\sum_{j=1}^{L_k} \beta_j^{ELM(k)} g\left(w_j^{ELM(k)T}[v_t, \hat{s}_t^{ARIMA}, \hat{s}_t^{ETS}] + b_j^{ELM(k)}\right)\right)_{k=1}^{20} \tag{12}$$

In Equation 12, L_k is the number of hidden nodes for the k-th ELM network, selected via GA within the range [100, 1000]. Each weight, bias, and activation function parameter is unique to the corresponding ELM network.

Algorithm for Identifying the Optimal Number of Hidden Nodes of the GA-tuned ARIMA-ETS-ELM

The algorithm for determining the optimal no. of hidden nodes within the proposed GA-tuned ARIMA-ETS-ELM model is outlined in Figure 4. The range of hidden neurons [100, 1000] was chosen based on a heuristic approach. During the preliminary experiment, we observed that a sufficient search space for the GA was required to identify an optimal network size. A smaller range resulted in underfitting, and a broader range was able to minimise the RMSE and effectively regularise model complexity. The GA search range of 100-1000 hidden nodes was chosen because preliminary experiments showed underfitting below 100 neurons and RMSE stabilised above 100. A wide search space helps GA avoid local minima. The final optimal value (725) was determined via fitness minimisation rather than arbitrarily. The GA operated with a population size of 20 over 50 generations. A linear arithmetic crossover strategy was employed. The crossover probability and mutation probability were set to 0.8 and 0.1, respectively. In ELM, input weights and biases are randomly assigned, output weights are analytically computed, and the activation function is typically fixed, making hidden-node size the principal structural hyperparameter affecting the bias-variance trade-off.

Input: Training data, ARIMA and ETS models (to generate fitted values), Population size (for the GA), Maximum generations (upper limit for GA iterations), Search space (range of hidden neurons for ELM), Stopping criterion (number of generations without improvement to terminate GA)
Output: Optimised number of hidden neurons (hd*) minimising RMSE

1. Preprocessing
1.1. **Train** ARIMA model on training data and obtain fitted values.
1.2. **Train** ETS model on training data and obtain fitted values.

2. Define Fitness Function
2.1 **Train** ELM For a given number of hidden neurons (hd), with:
- **Inputs:** Lag 1, 2, ..., 12, Fitted ARIMA values, Fitted ETS values

2.2 **Compute** the RMSE of the model.
2.3 **Return** the negative RMSE as the fitness value (minimisation objective).

3. Initialize Genetic Algorithm
3.1 **Set** population size = 20 and generation limit = 50.
3.2 **Generate** an initial population of hidden neuron values (hd) randomly within [100, 1000].
3.3 **Apply** elitism: retain the top 5% best solutions in each generation.
3.4 **Set** crossover rate = 0.8 and mutation rate = 0.1.
3.5 **Set** stopping criterion to terminate if no improvement in 10 generations.

4. Evolution Process
Repeat *until termination criteria (max generations or early stopping) are met:*
4.1 Evaluate the fitness function for each individual in the population.
4.2 Select parent solutions based on fitness scores.
4.3 Apply crossover (0.8 probability) and mutation (0.1 probability) to generate offspring.
4.4 Update population with new offspring while retaining the elite solutions.
4.5 Track the best solution found so far.

5. Termination
5.1 If no improvement in fitness for 10 generations, stop.
5.2 Return the best number of hidden neurons (hd*) minimising RMSE.

Figure 4. Hidden node optimisation algorithm

Therefore, the present work only optimised the hidden-node count, because it directly controls model capacity and complexity in ELM.

Experimental Setup

Table 3 summarises the programming languages and packages/ libraries used in this study to ensure methodological transparency and support replication efforts.

The R scripts used in this study are available at the following GitHub repository: https://github.com/drddas-kolkata/GA-Optimized-ARIMA-ETS-ELM. The codes include data loading, GA-based hyperparameter tuning of the proposed ARIMA-ETS-ELM, building the final optimised hybrid model, generating the 11-month-ahead forecasts using the GA-Optimised-ARIMA-ETS-ELM, and assessing forecast accuracy of the model.

RESULTS AND DISCUSSION

Identification of Optimised GA-tuned ARIMA-ETS-ELM Model

We used a GA to find the optimal number of hidden neurons in the ELM for the GA-tuned ARIMA-ETS-ELM model, aiming to minimise the RMSE. Figures 5 and 6 present the relationship between the number of hidden neurons and the corresponding RMSE and the best RMSE values (i.e., fitness values) in different generations during the GA-based ELM tuning, respectively.

Table 3
Summary of packages and programming languages used in the experimental setup

Language/ Package (Version)	Purpose/ Description	Author
R (4.4.2)	Traditional, AI-based, and hybrid model development and statistical tests	(R Core Team, 2024)
Python (3.11.12)	Applying the Moirai model	(Python Software Foundation, 2025)
stats (R, v4.4.2)	To perform the Wilcoxon Signed-Rank and Mann-Whitney U Test	(R Core Team, 2024)
tseries (R, v0.10.58)	To perform the BDS and ADF tests	(Trapletti & Hornik, 2024)
forecast (R, v8.24.0)	Forecasting using ARIMA, ETS, NNAR, SES, HLT, and HW	(Hyndman et al., 2024; Hyndman & Khandakar, 2008)
GA (R, v3.2.4)	Implementing GA	(Scrucca, 2013; Scrucca, 2017)
nnfor (R, v0.9.9)	Forecasting using MLP and ELM model development	(Kourentzes, 2023)
TSSVM (R, v0.1.0)	Forecasting using SVR	(Ray et al., 2022;)
forecastHybrid (R, v5.0.19)	Forecasting using ARIMA-ANN, ETS-ANN, ARIMA-ETS-ANN, STL-ETS, and STL-ARIMA	*(Shaub & Ellis, 2020)*
uni2ts (Python)	Forecasting using Moirai	(Aksu et al., 2024; Liu et al., 2024; Woo et al., 2024)

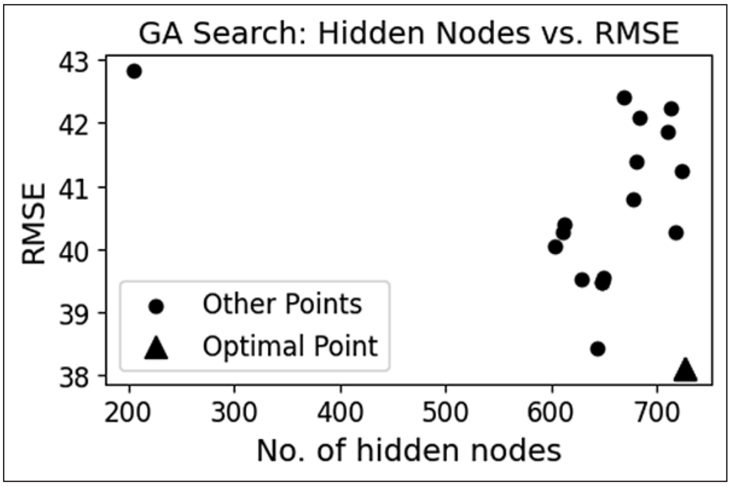

Figure 5. Identification of optimal hidden neurons in ELM via RMSE minimisation

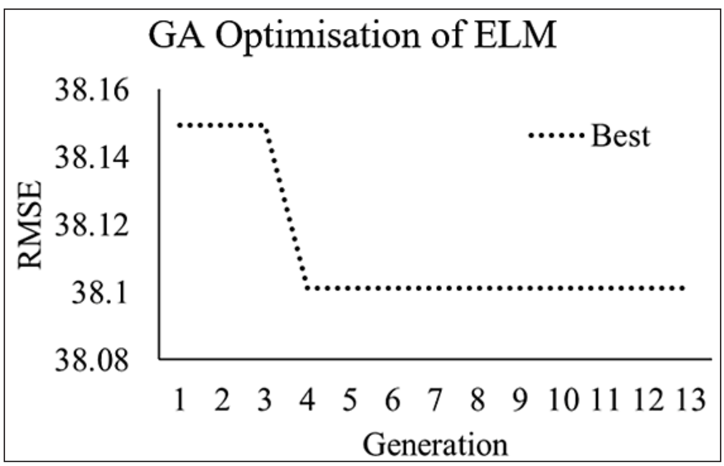

Figure 6. Evolution of the best RMSE across generations during GA optimisation of ELM.

In Figure 5, the triangle-shaped marker designates the optimal quantity of hidden neurons, which is 725, corresponding to the lowest RMSE value of 38.1. The final ELM architecture in the proposed GA-tuned ARIMA-ETS-ELM hybrid model is delineated as follows: 14 (input nodes: fitted ARIMA values, fitted ETS values, and lags 1 to 12 of the observed data) / 725 (hidden nodes: optimal number identified by the GA) / 1 (output node).

The GA-tuned ARIMA-ETS-ELM is a non-likelihood-based model. For this model, a pseudo-BIC was computed using the residual MSE and the number of hidden neurons to approximate model complexity. The resulting value (3770.09) is not directly comparable to the likelihood-based BIC values of ARIMA (1197.74) and ETS (1468.86) models.

The pseudo-BIC of the hybrid model provides only a rough indication of its relative complexity. This pseudo-BIC is used as an approximate complexity measure for the non-likelihood ELM framework. It is not directly comparable to likelihood-based BIC and serves only for internal structural interpretation and is not employed for cross-model ranking or performance comparison.

Overfitting with 725 hidden nodes was controlled using the Moore–Penrose pseudo-inverse solution, GA fitness based on RMSE minimisation, and median aggregation of 20 independently initialised ELM networks. The final evaluation was conducted strictly on unseen test data. We observed consistent improvements across 19 models, further supporting generalisation.

Forecasting Performance of the Proposed Approach

The line plot presented in Figure 7a shows that the observed values and forecasts are very similar. Meanwhile, the density plot in Figure 7b indicates that the distributions of the observed and forecasted values are closely aligned. The relationship indicates that the modes are close to one another; hence, no considerable amount of bias is present. The boxplot in Figure 7c shows that the values observed and predicted have close central tendencies, which indicates that the model can track the overall trend. The model further fails to produce excessive errors. The medians are very close, meaning that the forecast does not significantly overestimate or underestimate the mid-values. Figure 7d reveals the violin plot demonstrates that the overall shape of the observed distributions and forecasted distributions are almost equal and this is the way of knowing that the model is capturing the data. It is worth noting that the average values of the two violins are positioned at a similar point. This observation means that the average observed values are close to the average of the foregoing values. This therefore implies that little to no systematic bias in the forecasting of this model.

Table 4 presents the performance metrics of the GA-tuned ARIMA-ETS-ELM model, including ME, RMSE, MAE, MPE, and MAPE. It shows these metrics for a forecast horizon of 11 months, with individual values labelled as ME (11-month), RMSE (11-month), MAE (11-month), MPE (11-month), and MAPE (11-month). Additionally, Table 4 includes the average values for each of these metrics over the 11 months, denoted as average ME, average RMSE, average MAE, average MPE, and average MAPE.

The ME of 11 months was -52.73, and the average ME is 16.66, indicating that the model does not show a consistent over- or under-forecasting bias. The RMSE penalises significant errors. The RMSE -11 month and the average RMSE of the GA-tuned ARIMA-ETS-ELM were 141.79 and 66.84, respectively. It suggests a reasonable overall forecasting performance. The MAE of 11 months was 100.20, and the average MAE was 55.22. It indicates relatively low absolute error magnitudes.

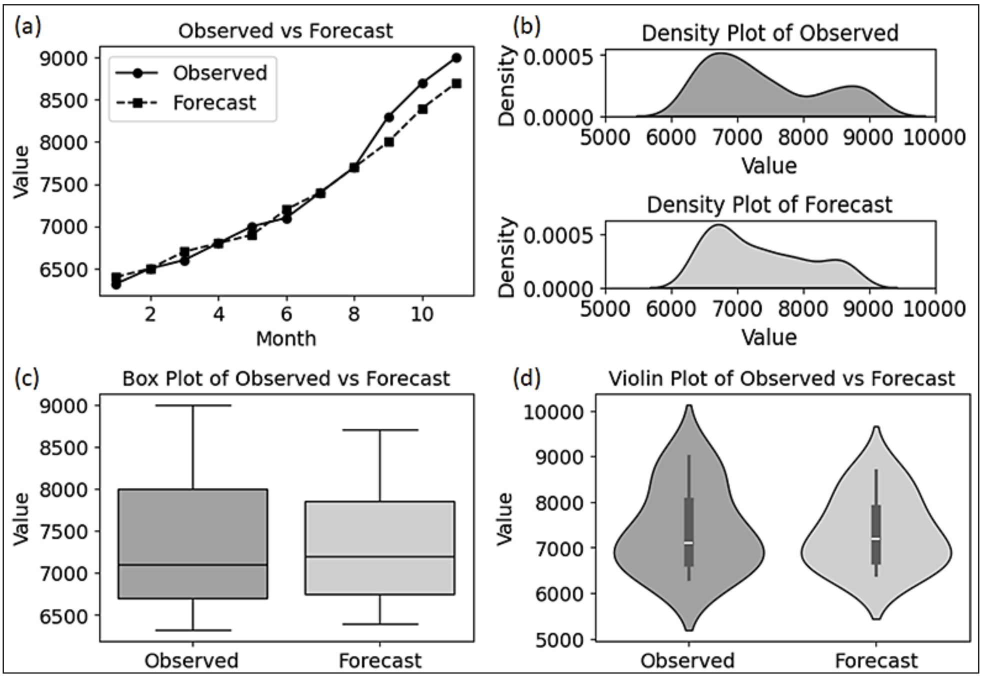

Figure 7. Forecast-performance of GA-tuned ARIMA-ETS-ELM: (a) Observed vs Forecast plot; (b) Density plots of observed and forecast values; (c) Box plot comparison; and (d) Violin plot comparison.

The MPE and MAPE are two percentage-based error measures. Low 11 months MPE of -0.58 and average MPE of 0.30 suggested minimal overall bias in forecast direction. All the MAPE values are well below the threshold of 10 for excellent performance. The MAPE of 11 months was 1.27, and the average MAPE was 0.79. It displays the GA-tuned ARIMA-ETS-ELM's strong forecasting ability.

Assessment of GA-tuned ARIMA-ETS-ELM Performance over Component Models

We performed a comparative analysis between the GA-tuned ARIMA-ETS-ELM and its components – Auto-ARIMA, Auto-ETS, and Auto-ELM, using out-of-sample forecasts over an 11-month horizon to evaluate the performance of the GA-tuned ARIMA-ETS-ELM relative to its individual components. The average ME of the GA-tuned ARIMA-ETS-ELM was 16.66. It was closest to zero among all. It exhibited that the GA-tuned ARIMA-ETS-ELM had the least forecast bias. The average RMSE (66.84) of the hybrid model was the lowest. It indicated that the hybrid model was more stable and accurate. It had the minimum average MAE (55.22) among all. The average MPE of the GA-tuned ARIMA-ETS-ELM was 0.30. It indicated that the model exhibited the least bias (in directional accuracy). It achieved the minimum average MAPE (0.79).

Table 4
Forecasting performance of the GA-tuned ARIMA-ETS-ELM model based on error metrics over an 11-month horizon

Metric	Value
ME 11-month	-52.73
Average ME	16.66
RMSE 11-month	141.79
Average RMSE	66.84
MAE 11-month	100.20
Average MAE	55.22
MPE 11-month	-0.58
Average MPE	0.30
MAPE 11-month	1.27
Average MAPE	0.79

The value was well below the commonly accepted threshold of 10% for excellent forecasting. Therefore, the GA-tuned ARIMA-ETS-ELM outperforms all its components across all evaluations. The GA-tuned ARIMA-ETS-ELM approach improved average MAPE over automatic ARIMA (74.7%), automatic ETS (61.2%), and automatic ELM (79.6%). Similar improvements were observed for average RMSE, also, i.e., 76.4% over automatic ARIMA, 62.2% over automatic ETS, and 80.9% over automatic ELM. It suggested improved forecasting. These findings highlight the advantage of combining statistical techniques with machine learning in developing a more accurate and robust forecasting framework. Table 5 presents these findings.

An ablation study was conducted to assess the individual contributions of Auto ARIMA and Auto ETS in the GA-tuned ARIMA-ETS-ELM framework. The GA-tuned ARIMA-ETS-ELM model was evaluated compared to two simplified variants: (1) the GA-tuned ARIMA-ELM model and (2) the GA-tuned ETS-ELM model. Table 6 displays the results. The GA-tuned ARIMA-ETS-ELM model outperformed both of its simplified variants, namely the GA-tuned ARIMA-ELM and GA-tuned ETS-ELM, across all evaluated metrics, achieving the lowest average ME, RMSE, MAE, MPE, and MAPE. The results indicate that incorporating ETS in the ARIMA-ELM model significantly augments average MAPE (41.2%) and average RMSE (41.1%). Adding ARIMA to the ETS-ELM combination yielded a substantial gain in average MAPE (40.2%) and average RMSE (39.8%). These findings confirm that both ARIMA and ETS individually enhance the performance of the GA-tuned ARIMA-ETS-ELM model. ARIMA captures linear and autocorrelated trends, ETS captures exponential smoothing behaviour and trend-seasonality interactions, ELM models the non-linear relationships, and GA ensures the optimal configuration of the ELM layer.

Table 5
Performance comparison of the GA-tuned ARIMA-ETS-ELM model with its component models (Auto ARIMA, Auto ETS, and Auto ELM) based on out-of-sample forecasting over an 11-month horizon

Sl.	Type	Model	Metric				
			Average ME	**Average RMSE**	**Average MAE**	**Average MPE**	**Average MAPE**
1.	Component model	Auto ARIMA	-37.28	283.7	215.33	-0.43	3.13
2.		Auto ETS	41.80	176.96	142.43	0.73	2.03
3.		Auto ELM	-230.85	349.21	260.76	-3.40	3.87
4.	Full model	GA-tuned ARIMA-ETS-ELM	16.66	66.84	55.22	0.30	0.79

Table 6
Performance comparison of the GA-tuned ARIMA-ETS-ELM model with its two simplified variants based on out-of-sample forecasting over an 11-month horizon

Sl.	Type	Model	Metric				
			Average ME	**Average RMSE**	**Average MAE**	**Average MPE**	**Average MAPE**
1.	Simplified variants	GA-tuned ARIMA-ELM	92.01	113.56	96.08	1.28	1.34
2.		GA-tuned ETS-ELM	90.48	111.10	94.08	1.27	1.32
3.	Proposed model	GA-tuned ARIMA-ETS-ELM	16.66	66.84	55.22	0.30	0.79

Performance Comparison of GA-tuned ARIMA-ETS-ELM with Others

Table 7 presents the performance of the GA-tuned ARIMA-ETS-ELM against eight other standalone models (excluding its component models) in terms of evaluated metrics. Similarly, Table 8 exhibits the performance of the GA-tuned ARIMA-ETS-ELM against six other hybrid techniques, excluding the GA-tuned hybrids used for the ablation study. For cumulative data, a naïve drift model would produce near-linear extrapolation ($\hat{Y}_{t+h} = Y_t + h.\frac{Y_t - Y_1}{t-1}$). Given the strong linear trend ($R^2 = 0.976$), such a model would perform reasonably well, but it cannot capture nonlinear growth acceleration or irregular dynamics. The proposed hybrid model demonstrates substantial RMSE and MAPE improvements over simpler linear alternatives.

The GA-tuned ARIMA-ETS-ELM model outperformed all seven standalones, a generative-AI, and six hybrid models, achieving the lowest average scores for ME, RMSE, MAE, MPE, and MAPE (Tables 7 and 8).

Table 9 provides the percentage gain in performance achieved by the GA-tuned ARIMA-ETS-ELM hybrid model compared to the nineteen (ten standalone, one generative-AI, and eight hybrid) models, based on average MAPE and average RMSE over an 11-month horizon.

Table 7
Performance comparison of the GA-tuned ARIMA-ETS-ELM model with other standalone forecasting techniques based on out-of-sample forecasting over an 11-month horizon

Sl.	Type	Model	Metric				
			Average ME	Average RMSE	Average MAE	Average MPE	Average MAPE
1.	Statistical model	STL	-322.30	408.47	322.30	-4.87	4.87
2.	Statistical model	SES	-489.15	601.48	489.15	-7.79	7.79
3.	Statistical model	HLT	41.80	176.96	142.43	0.73	2.03
4.	Statistical model	HW	-476.23	578.10	476.23	-7.52	7.52
5.	AI-based model	NNAR	-470.40	581.47	470.40	-7.46	7.46
6.	AI-based model	MLP	-25.54	206.7	161.99	-0.24	2.32
7.	AI-based model	SVR	-846.70	923.88	846.70	-14.30	14.30
8.	Generative AI model	Moirai	-373.18	462.96	373.18	-5.72	5.72
9.	Proposed Model	GA-tuned ARIMA-ETS-ELM	16.66	66.84	55.22	0.30	0.79

Table 8
Performance comparison of the GA-tuned ARIMA-ETS-ELM model with other hybrid techniques based on out-of-sample forecasting over an 11-month horizon

Sl.	Author	Technique	Metric				
			Average ME	Average RMSE	Average MAE	Average MPE	Average MAPE
1.	(Hightower et al., 2024)	STL-ARIMA	-170.00	307.58	218.07	-2.45	3.20
2.	(Hightower et al., 2024)	STL-ETS	-132.81	237.71	165.24	-1.85	2.36
3.	(Talkhi et al., 2021)	ARIMA-ANN using CV-error	-158.31	320.78	229.63	-2.29	3.39
4.	(Perone, 2021)	ARIMA-ANN using equal weight	-253.62	379.87	273.89	-3.80	4.12
5.	(Perone, 2021)	ETS-ANN using equal weight	-219.54	317.41	235.03	-3.21	3.45
6.	(Perone, 2021)	ARIMA-ETS-ANN using equal weight	-154.06	283.77	198.89	-2.19	2.89
7.	Proposed model	GA-tuned ARIMA-ETS-ELM	16.66	66.84	55.22	0.30	0.79

It also showed significant improvements in average MAPE and RMSE compared to all nineteen models (Table 9). The performance gains are notable, ranging from approximately 40% to 95% for average MAPE and from about 39% to 93% for average RMSE, indicating a considerable enhancement in accuracy.

Table 9
Percentage performance gain of the GA-tuned ARIMA-ETS-ELM model over nineteen forecasting techniques

Sl.	Technique	Performance gain (%) achieved by GA-tuned ARIMA-ETS-ELM	
		Average MAPE	Average RMSE
1.	Auto ARIMA (Component model)	74.74	76.44
2.	Auto ETS (Component model)	61.17	62.23
3.	Auto ELM (Component model)	79.57	80.86
4.	STL	83.78	83.64
5.	SES	89.86	88.89
6.	HLT	61.17	62.23
7.	HW	89.49	88.44
8.	NNAR	89.41	88.51
9.	MLP	65.94	67.66
10.	SVR	94.48	92.77
11.	Moirai	86.41	85.72
12.	GA-tuned ARIMA-ELM	41.19	41.14
13.	GA-tuned ETS-ELM	40.22	39.84
14.	STL-ARIMA (Hightower et al., 2024)	75.30	78.27
15.	STL-ETS (Hightower et al., 2024)	66.52	71.88
16.	ARIMA-ANN using CV-error (Talkhi et al., 2021)	76.69	79.16
17.	ARIMA-ANN using equal weight (Perone, 2021)	80.82	82.40
18.	ETS-ANN using equal weight (Perone, 2021)	77.13	78.94
19.	ARIMA-ETS-ANN using equal weight (Perone, 2021)	72.68	76.45

The results underscore the robust out-of-sample predictive capability of the proposed GA-tuned ARIMA-ETS-ELM approach when evaluated against traditional statistical models, contemporary generative AI methodologies, and other hybrid techniques.

Statistical Significance Testing of the GA-tuned ARIMA-ETS-ELM

The GA-tuned ARIMA-ETS-ELM approach outperformed all nineteen other techniques and was identified as the best candidate model for out-of-sample prediction of cumulative dengue cases in the USA. Statistical significance testing was conducted to assess the robustness of the forecasting performance achieved by the GA-tuned ARIMA-ETS-ELM approach. Table 10 presents the results. Both tests support the robustness of the GA-tuned ARIMA-ETS-ELM model. The Mann-Whitney U test indicates that both the distributions (forecasted values and the observed data) are close with no significant difference. The Wilcoxon Signed Rank test confirms that the model exhibits a low median forecast MAPE (≤ 1) for this 11-month forecast horizon. Therefore, these statistical pieces of evidence strengthen the GA-tuned ARIMA-ETS-ELM model's forecasting accuracy and reliability.

Table 10
Statistical significance testing of forecasting performance of the GA-tuned ARIMA-ETS-ELM model

Statistical test	Null hypothesis	Statistic	Significance	Decision
Mann-Whitney U test	The distribution of observed data and GA-tuned ARIMA-ETS-ELM forecasts is the same	61	1.00	Fail to reject the null hypothesis
Wilcoxon Signed-Rank test	Median MAPE of GA-tuned ARIMA-ETS-ELM for the forecasting horizon of 11 months ≤ 1	6.50	.9919	Fail to reject the null hypothesis

Statistical tests were applied to the 11-step forecast horizon to evaluate distributional alignment and median MAPE constraints. Nonparametric tests (Mann-Whitney/Wilcoxon) were used because forecast errors may not follow a normal distribution, the sample size is small ($n = 11$), and the focus was on distributional similarity and median constraints; the Diebold–Mariano test could be applied in future work for formal pairwise predictive accuracy comparison.

Theoretical Implications of the Research Findings

The theoretical implications of the findings from the present study are outlined as follows:

- Model integration: Auto ARIMA handles linear and autocorrelated patterns, while Auto ETS models trend and seasonality through exponential smoothing. ELM addresses non-linear relationships, which contribute strengths to the GA-tuned ARIMA-ETS-ELM approach.
- Evolutionary optimisation: The current work also explores GA optimisation, revealing that it enhances the model by optimally configuring the ELM component and improving forecast accuracy, emphasising evolutionary optimisation's importance in hybrid systems.
- Theoretical merit: The GA-tuned ARIMA-ETS-ELM performed better than several other models, including classical models, neural networks, and generative AI (e.g., Moirai). Statistical significance testing (Mann-Whitney U and Wilcoxon Signed Rank tests) highlights the reliability of the hybrid model, supporting its theoretical soundness and merit.
- Contribution to forecasting theory: The current study enhances the theoretical basis of hybrid forecasting by demonstrating that structured model integration and effective optimisation lead to greater accuracy in practical applications such as epidemiological surveillance.

The present study employed cumulative values of the monthly dengue cases to capture the long-term progression and growth pattern of dengue. This cumulative (monthly) data smoothed the irregularities caused by underreporting or short-term volatility in the data. This helps to improve the model's stability. Further aid in its ability to identify structural trends. Due to larger denominators, the cumulative data may tend to produce a lower MAPE. These results primarily reflect trend-level forecast accuracy. The proposed GA-tuned ARIMA-ETS-ELM framework can be equivalently applied to non-cumulative (monthly) dengue cases for assessing short-term predictive accuracy in future studies.

Practical Implications of the Research Findings

The multifaceted empirical essences of the present work are as follows:

- Enhanced forecasting accuracy for dengue surveillance: The model demonstrates low MAPE and RMSE across various forecast horizons. It facilitates accurate early prediction of dengue outbreaks, assisting public health authorities with timely resource allocation, intervention planning, and designing awareness campaigns.
- Real-World applications: The model is quite flexible and able to capture linear as well as seasonal and nonlinear patterns that make it suitable for a wide range of health-related time series problems. Performance of the proposed model has been rigorously validated using nonparametric statistical tests, which include the Mann-Whitney U test and the Wilcoxon Signed-Rank test. Therefore, the presented model can be well-applied to the settings where the data is noisy, incomplete, or uncertain.
- Model optimisation: There is an automated and efficient model optimisation. The ELM component is automatically tuned by the GA, which minimises the need to use manual parameter selection or expert assistance. This, in its turn, makes the model easier to use in government or institutional environments.
- Generalise: The GA-tuned ARIMA-ETS-ELM model can be used to forecast dengue and can be modified to suit other infectious diseases or epidemiology-related prediction tasks. It presents a flexible model which integrates both statistical and AI-enhanced predictive approaches to make efficient real-life decisions.

The present study used monthly dengue data of the United States from the World Health Organisation (2025) dataset. The dataset includes country-wise dengue data for multiple countries. The proposed GA-tuned ARIMA-ETS-ELM framework is practically adaptable and can be extended to other countries exhibiting distinct transmission dynamics.

The proposed GA-tuned ARIMA-ETS-ELM model can act as an early-warning tool for public health authorities. In case of any potential/future outbreak, the monthly forecasts can assist in planning vector control measures, allocating medical supplies, and preparing hospital resources in advance. The same framework can be extended to other infectious diseases that exhibit similar temporal dynamics.

Limitations and Future Direction

The research studies have established many factors that play a significant role in the dynamics of disease outbreaks that can significantly deteriorate the level of performance of the forecasting model, which include selective reporting, hetero-spatial conditions, and sensitivity in the estimation of the parameters. The study acknowledges possible reporting variability. The robustness is addressed indirectly through hybrid modelling that captures both linear and nonlinear shifts, nonparametric statistical validation, and the use of cumulative transformation. These reduce abrupt reporting spikes. Explicit structural break detection (e.g., Chow test, Bai–Perron) was not incorporated in the current work and is suggested for future work. The current paper concentrated on the development of the model against the USA dataset, in which the proposed model is prepared by using lagged values and fitted values of statistical models as inputs. These inputs might not be a complete capture of the external factors like weather, mobility or demographic data that affect the dengue spread. The research centred on the dengue cases, which were cumulative monthly. The current research did not investigate the fine-grained prediction (e.g., of week or city). This study focuses on point forecasts. The current work presented distributional visualisations (density and violin plots), but formal prediction intervals are not constructed. Incorporating bootstrapped or quantile-based interval estimation within the ELM framework is an important direction for future policy-grade decision support. Irrespective of such constraints, the study can lead to a hint in making decisions on public health. There are several directions of the research. The current framework is structurally extensible: ARIMA and ETS components can incorporate exogenous regressors (ARIMAX), and ELM can include additional input features such as climate, mobility, or vector indices, which are expected to enhance short-term predictive responsiveness and mechanistic interpretability. We can add exogenous variables (e.g., climate indicators, mosquito populations, rainfall) to the ARIMA-ETS-ELM model to investigate its prediction improvement. To identify the model's merit and generality, researchers may test this hybrid model on other regions' data. This will help to assess its usefulness and generalisation. The ARIMA-ETS-ELM can be refined with other sophisticated evolutionary algorithms (e.g., Particle Swarm Optimisation and Differential Evolution). This model can be adapted to create early warning systems for a real-time decision-support system that affects the health of the population.

CONCLUSION

Dengue adversely impacts millions worldwide. Several significant contributions to the present research frontier have been accomplished. Both traditional and AI-driven forecasting models illustrate the progression of Dengue, but they possess inherent limitations. These include challenges in identifying complex non-linear associations and issues related to the interpretability of the models.

To balance the trade-offs, the proposed model incorporates the GA-tuned ARIMA-ETS-ELM method for automatically forecasting cumulative dengue cases in the USA, wherein the ELM architecture within the hybrid model is optimised using a GA. It identifies the optimal no. of hidden neurons in the ELM. While GAs is primarily used for multi-objective optimisation, this study demonstrates their effectiveness in tuning the ELM to enhance forecasting accuracy. The present study thoroughly evaluates the proposed GA-tuned ARIMA-ETS-ELM model against nineteen leading forecasting methodologies. This comparison encompasses ten standalone, one generative-AI, and eight hybrid approaches. An ablation study was conducted on the GA-tuned ARIMA-ETS-ELM model to isolate the contributions of the ARIMA and ETS components. The GA-tuned ARIMA-ETS-ELM model is identified as the most effective method among the assessed techniques for forecasting cumulative dengue cases in the United States. The performance of this model was further validated through statistical significance tests, confirming its efficacy. The GA-tuned ARIMA-ETS-ELM model demonstrates significant forecasting accuracy, achieving an average ME of 16.66, a RMSE of 66.84, a MAE of 55.22, a MPE of 0.30, and a MAPE of 0.79. The metrics clearly illustrate that the performance is well within acceptable limits, providing strong evidence of the proposed model's innovative nature and retaining its prospect of making a substantial contribution to the field of study.

ACKNOWLEDGEMENT

The authors sincerely thank the editor and anonymous referees for their constructive comments and suggestions, which greatly improved the quality of this paper.

LIST OF ABBREVIATIONS

ADF	:	Augmented Dickey-Fuller
ANN	:	Artificial neural network
ARIMA	:	AutoRegressive Integrated Moving Average
BDS	:	Brock-Dechert-Scheinkman
CV	:	Cross-validation
ELM	:	Extreme learning machine
ETS	:	Error, trend, seasonal
GA	:	Genetic algorithm
HLT	:	Holt’s linear trend
HW	:	Holt-Winters
MAE	:	Mean absolute error
MAPE	:	Mean absolute percentage error
ME	:	Mean error
MLP	:	Multilayer perceptron
MPE	:	Mean percentage error
NNAR	:	Neural network AutoRegressive

RMSE	:	Root mean square error
SES	:	Simple exponential smoothing
STL	:	Seasonal-trend decomposition using LOESS
SVR	:	Support vector regression
USA	:	United States of America

REFERENCES

Ahmadoni, N. I. (2025). The role of cybersecurity governance in activating the disclosure of cybersecurity risk management and its impact on improving financial performance: An experimental study on banks listed on the Egyptian Stock Exchange. *Alexandria Journal of Accounting Research, 9*(1), 1-71. https://doi.org/10.21608/aljalexu.2025.397533

Ahmar, A. S., & Boj, E. (2021). Application of neural network time series (NNAR) and ARIMA to forecast infection fatality rate (IFR) of COVID-19 in Brazil. *JOIV: International Journal on Informatics Visualisation, 5*(1), 8-10. https://doi.org/10.30630/joiv.5.1.372

Aksu, T., Woo, G., Liu, J., Liu, X., Liu, C., Savarese, S., Xiong, C., & Sahoo, D. (2024). *GIFT-Eval: A benchmark for general time series forecasting model evaluation* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.10393

Ariyanto, E., & Nugraha, M. K. R. (2024). Population vs. poverty level in the future in Indonesia: Holt's linear trend method. *International Journal of Sustainable Development and Planning, 19*(7), 2763-2770. https://doi.org/10.18280/ijsdp.190733

Baker, Q. B., Faraj, D., & Alguzo, A. (2021). Forecasting dengue fever using machine learning regression techniques. In *Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS)* (pp. 157-163). IEEE. https://doi.org/10.1109/ICICS52457.2021.9464619

Bunprom, P., Siramaneerat, I., & Bhumkittipich, P. (2026). Forecasting dengue fever incidence in Thailand using ARIMA: Implications for public health planning. *Environment and Natural Resources Journal, 24*(1), 88-97. https://doi.org/10.32526/ennrj/24/20250069

Chakraborty, A., Das, D., Mitra, S., De, D., & Pal, A. J. (2024). Forecasting adversities of COVID-19 waves in India using intelligent computing. *Innovations in Systems and Software Engineering, 20*, 821-837. https://doi.org/10.1007/s11334-022-00486-y

Das, D., & Chakrabarti, S. (2021). Forecasting non-linear WPI of manufacture of chemicals and chemical products in India: An MLP approach. *International Journal of Advanced Technology and Engineering Exploration, 8*(82), 1193-1203. https://doi.org/10.19101/IJATEE.2021.874407

Dhaked, D., Sharma, O., Gopal, Y., & Gopal, D. (2025). Prediction of dengue patients using deep learning methods amid complex weather conditions in Jaipur, India. *Discover Public Health, 22*, Article 48. https://doi.org/10.1186/s12982-025-00448-2

Flores, A., Tito-Chura, H., Guzman-Valdivia, J., Morales-Gonzales, R., Rosado-Chavez, C., & Silva-Delgado, C. (2026). Multi-step forecasting of weekly dengue cases using data augmentation and convolutional recurrent neural networks. *Computers, 15*(2), Article 114. https://doi.org/10.3390/computers15020114

Hasan, P., Khan, T. D., Alam, I., & Haque, M. E. (2024). Dengue in tomorrow: Predictive insights from ARIMA and SARIMA models in Bangladesh: A time series analysis. *Health Science Reports, 7*(12), Article e70276. https://doi.org/10.1002/hsr2.70276

Hightower, A., Ziedan, A., Guo, J., Zhu, X., & Brakewood, C. (2024). A comparison of time series methods for post-COVID transit ridership forecasting. *Journal of Public Transportation, 26*, Article 100097. https://doi.org/10.1016/j.jpubtr.2024.100097

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 27*(3), 1-22. https://doi.org/10.18637/jss.v027.i03

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., & Yasmeen, F. (2024). *forecast: Forecasting functions for time series and linear models* (Version 8.23.0) [R package]. https://pkg.robjhyndman.com/forecast/

Ivanovski, Z., & Ivanovska, N. (2024). The augmented Dickey–Fuller test for the stationarity of the final public consumption and GDP time series of the Republic of North Macedonia. *UTMS Journal of Economics, 15*(2), 109-124.

Jirasakuldech, B., Emekter, R., & Bui, T. (2023). Non-linear structures, chaos, and bubbles in US regional housing markets. *Journal of Economics and Finance, 47*(1), 63-93. https://doi.org/10.1007/s12197-022-09598-4

Kourentzes, N. (2023). *nnfor: Time series forecasting with neural networks* (Version 0.9.9) [R package]. https://cran.r-project.org/package=nnfor

Li, K. (2024). Prediction of atmospheric temperature in Sacramento area based on ARIMA and ETS models. *Theoretical and Natural Science, 51*, 165-171. https://doi.org/10.54254/2753-8818/51/2024CH0197

Liu, X., Liu, J., Woo, G., Aksu, T., Liang, Y., Zimmermann, R., Liu, C., Savarese, S., Xiong, C., & Sahoo, D. (2024). *Moirai-MoE: Empowering time series foundation models with sparse mixture of experts* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.10469

Lu, X., Teh, S. Y., Tay, C. J., Abu Kassim, N. F., Fam, P. S., & Soewono, E. (2024). Application of multiple linear regression model and long short-term memory with compartmental model to forecast dengue cases in Selangor, Malaysia based on climate variables. *Infectious Disease Modelling, 10*(1), 240-256. https://doi.org/10.1016/j.idm.2024.10.007

Mahanty, C., Patro, S. G. K., Rathor, S., Rachapudi, V., Muzammil, K., Islam, S., Razak, A., & Khan, W. A. (2024). Forecasting of coronavirus active cases by utilising logistic growth model and fuzzy time series techniques. *Scientific Reports, 14*, Article 18039. https://doi.org/10.1038/s41598-024-67161-z

Mendoza Chirinos, J. (2024). *Spatio-temporal prediction of dengue outbreaks using neural networks with climatic and socioeconomic factors in Colombian municipalities* [Master's thesis, Universidad de los Andes]. https://hdl.handle.net/1992/75226

Mobin, M. A. (2024). Forecasting dengue in Bangladesh using meteorological variables with a novel feature selection approach. *Scientific Reports, 14*, Article 32073. https://doi.org/10.1038/s41598-024-83770-0

Mohammed, N., & Rahman, M. (2023). Forecasting dengue incidence in Bangladesh using seasonal ARIMA model: A time series analysis. In M. S. Satu, M. A. Moni, M. S. Kaiser, & M. S. Arefin (Eds.), *Machine intelligence and emerging technologies* (Vol. 491). Springer. https://doi.org/10.1007/978-3-031-34622-4_47

Nwogu, E. C., Iwueze, I. S., & Nlebedim, V. U. (2016). Some tests for seasonality in time series data. *Journal of Modern Applied Statistical Methods, 15*(2), Article 24. https://doi.org/10.22237/jmasm/1478002920

Parreño, S. J. E., & Anter, M. C. J. (2024). New approach for forecasting rice and corn production in the Philippines through machine learning models. *Multidisciplinary Science Journal, 6*(9), Article 2024168. https://doi.org/10.31893/multiscience.2024168

Patil, S., & Pandya, S. (2021). Forecasting dengue hotspots associated with variation in meteorological parameters using regression and time series models. *Frontiers in Public Health, 9*, Article 798034. https://doi.org/10.3389/fpubh.2021.798034

Pereira da Veiga, C., Pereira da Veiga, C. R., Girotto, F. M., Marconatto, D. A. B., & Su, Z. (2024). Implementation of the ARIMA model for the prediction of economic variables: Evidence from the health sector in Brazil. *Humanities and Social Sciences Communications, 11*, Article 1068. https://doi.org/10.1057/s41599-024-03023-3

Perone, G. (2021). Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalisations in Italy. *The European Journal of Health Economics, 23*, 917-940. https://doi.org/10.1007/s10198-021-01347-4

Python Software Foundation. (2025). *The Python language reference* (Version 3.11.12). https://docs.python.org/3/reference/

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rahman, M., Rashid, F., Roy, S. K., & Habib, M. A. (2024). Application of extreme learning machine (ELM) forecasting model on CO_2 emission dataset of a natural gas-fired power plant in Dhaka, Bangladesh. *Data in Brief, 54*, Article 110491. https://doi.org/10.1016/j.dib.2024.110491

Rangarajan, P., Mody, S. K., & Marathe, M. (2019). Forecasting dengue and influenza incidences using a sparse representation of Google Trends, electronic health records, and time series data. *PLOS Computational Biology, 15*(11), Article e1007518. https://doi.org/10.1371/journal.pcbi.1007518

Ray, M., Barman, S., Sinha, K., & Singh, K. N. (2022). *TSSVM: Time series forecasting using SVM model* (Version 0.1.0) [R package]. https://doi.org/10.32614/CRAN.package.TSSVM

Saleh, A. Y., & Baiwei, L. (2021). Dengue prediction using deep learning with long short-term memory. In *Proceedings of the 2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA)* (pp. 1-5). IEEE. https://doi.org/10.1109/eSmarTA52612.2021.9515734

Sanchez-Gendriz, I., de Souza, G. F., de Andrade, I. G. M., Neto, A. D. D., de Medeiros Tavares, A., Barros, D. M. S., de Morais, A. H. F., Galvão-Lima, L. J., & de Medeiros Valentim, R. A. (2022). Data-driven computational intelligence applied to dengue outbreak forecasting: A case study at the scale of the city of Natal, RN, Brazil. *Scientific Reports, 12*(1), Article 6550. https://doi.org/10.1038/s41598-022-10512-5

Santosa, R. G., Chrismanto, A. R., Raharjo, W. S., & Lukito, Y. (2024). LQ45 stock price forecasting: A comparison study of ARIMA (p, d, q) and Holt–Winter method. *International Journal of Information Technology and Computer Science Applications, 2*(2), 115-129. https://doi.org/10.58776/ijitcsa.v2i2.150

Saragih, S. A., Munar, A., & Hasibuan, W. R. (2024). Forecasting the amount of corn production in North Sumatra based on 2017–2021 data using the single and double exponential smoothing method (case study of Central Bureau of Statistics of North Sumatra). *Journal of Artificial Intelligence and Engineering Applications, 3*(2), 614-617. https://doi.org/10.59934/jaiea.v3i2.449

Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software, 53*(4), 1-37. https://doi.org/10.18637/jss.v053.i04

Scrucca, L. (2017). On some extensions to GA package: Hybrid optimisation, parallelisation and islands evolution. *The R Journal, 9*(1), 187-206. https://doi.org/10.32614/RJ-2017-008

Shaub, D., & Ellis, P. (2020). *forecastHybrid: Convenient functions for ensemble time series forecasts* (Version 5.0.19) [R package]. https://doi.org/10.1057/jors.1969.103

Sucipto, C. S., Sulandari, W., & Susanti, Y. (2025). Implementation of cross-validation on Hang Seng Index forecasting using Holt's exponential smoothing and Auto-ARIMA method. *BAREKENG: Jurnal Ilmu Matematika dan Terapan, 19*(1), 13-24. https://doi.org/10.30598/barekengvol19iss1pp13-24

Sutriyawan, A., Martini, M., Sutiningsih, D., Agushybana, F., Wahyuningsih, N. E., Adamu, V. E. V., Akbar, H., & Aba, M. (2024). Time series analysis of dengue incidence in Bandung City, Indonesia using an ARIMA model. *Journal of Microbiology, Epidemiology and Immunobiology, 101*(6), 803-811. https://doi.org/10.36233/0372-9311-570

Talkhi, N., Fatemi, N. A., Ataei, Z., & Nooghabi, M. J. (2021). Modelling and forecasting number of confirmed and deaths caused by COVID-19 in Iran: A comparison of time series forecasting methods. *Biomedical Signal Processing and Control, 66*, Article 102494. https://doi.org/10.1016/j.bspc.2021.102494

Tomov, L., Chervenkov, L., Miteva, D. G., Batselova, H., & Velikova, T. (2023). Applications of time series analysis in epidemiology: Literature review and our experience during the COVID-19 pandemic. *World Journal of Clinical Cases, 11*(29), 6974-6983. https://doi.org/10.12998/wjcc.v11.i29.6974

Torres, I., Nakano, M., Cime-Castillo, J. A., Escamilla-Hernandez, E., Lopez-Garcia, O., & Lanz Mendoza, H. (2025). Dengue-infected mosquito detection with uncertainty evaluation based on Monte Carlo dropout. *IEEE Latin America Transactions, 23*(2), 135-143. https://doi.org/10.1109/TLA.2025.10851361

Trapletti, A., & Hornik, K. (2024). *tseries: Time series analysis and computational finance* (Version 0.10-58) [R package]. https://doi.org/10.32614/CRAN.package.tseries

Wang, G., Zhuang, L., Mo, L., Yi, X., Wu, P., & Wu, X. (2023). BAG: A linear-nonlinear hybrid time series prediction model for soil moisture. *Agriculture, 13*(2), Article 379. https://doi.org/10.3390/agriculture13020379

World Health Organisation. (2025). *Global dengue surveillance*. Retrieved January 2025, from https://worldhealthorg.shinyapps.io/dengue_global/

Woo, G., Liu, C., Kumar, A., Xiong, C., Savarese, S., & Sahoo, D. (2024). Unified training of universal time series forecasting transformers. In *Proceedings of the 41st International Conference on Machine Learning (ICML 2024)* (Vol. 235, pp. 53140-53164). PMLR. https://ink.library.smu.edu.sg/sis_research/9906

Wu, Z., Zhou, J., & Yu, X. (2025). Forecast natural gas price by an extreme learning machine framework based on a multi-strategy grey wolf optimiser and signal decomposition. *Sustainability, 17*(12), Article 5249. https://doi.org/10.3390/su17125249

Xiong, Y., Ma, Y., Ruan, L., Li, D., Lu, C., Huang, L., & National Traditional Chinese Medicine Medical Team. (2022). Comparing different machine learning techniques for predicting COVID-19 severity. *Infectious Diseases of Poverty, 11*(1), Article 19. https://doi.org/10.1186/s40249-022-00946-4

Yadav, V., & Nath, S. (2019). Novel hybrid model for daily prediction of PM10 using principal component analysis and artificial neural network. *International Journal of Environmental Science and Technology, 16*(6), 2839-2848. https://doi.org/10.1007/s13762-018-1999-x

Yunisa, F. A., & Siregar, M. A. P. (2023). The application of seasonal trend decomposition using LOESS for export forecasting by economic commodity group in North Sumatra. *ZERO: Jurnal Sains, Matematika, dan Terapan, 7*(1), 61-67. https://doi.org/10.30829/zero.v7i1.17341