DOI: 10.21917/ijsc.2026.0590

ASSESSING THE EFFICACY OF TRADITIONAL AND ADVANCED PREDICTIVE TECHNIQUES IN ELECTRICITY DEMAND FORECASTING IN THE ASEAN REGION

Samuel John E. Parreno
Mathematics Division, University of Mindanao Digos College, Philippines

Abstract

This study examines various predictive models for forecasting electricity demand in the ASEAN region, characterized by rapid industrialization and diverse economic landscapes. Traditional methods like ARIMA and the Grey Model, while effective in certain scenarios, often fall short in capturing the complexities of electricity consumption due to evolving factors like technological advancements and economic policies. Advanced models such as Neural Network Autoregressive, Echo State Networks, Autoregressive Support Vector Machine, and Extreme Gradient Boosting have demonstrated superior capabilities in various fields, including finance and public health. Our research, leveraging data from the World Development Indicator database, evaluates these models specifically for per capita electricity consumption in ASEAN nations. The results highlight the advanced models, particularly XGBoost, for their accuracy and adaptability in handling the dynamic electricity demand, suggesting their greater efficacy over traditional models. This study not only provides valuable insights for electricity demand forecasting in rapidly evolving economies, but also emphasizes the importance of appropriate model selection for informed energy policy and planning. The findings have significant policy implications, aiding in the development of more efficient, sustainable, and equitable energy management across the ASEAN region, and pave the way for future research integrating real-time data and socio-economic factors.

Keywords:

Electricity Demand Forecasting, Predictive Modeling, ASEAN Region, Machine Learning, Time Series Analysis

1. INTRODUCTION

Electricity, a fundamental cornerstone of modern civilization, plays an indispensable role in powering industries, illuminating cities, and propelling technological advancements. Electricity holds a particularly prominent role in the Association of Southeast Asian Nations (ASEAN) region, which is composed of diverse and vibrant economies. Here, electricity transcends its conventional role as a mere utility. It emerges as the lifeblood that catalyzes economic growth, underpins social development, and nurtures the aspirations of over 650 million inhabitants [3]. As these nations collectively embark on a path marked by rapid industrialization and digitalization, the demand for electricity witnesses an unprecedented surge [4] [5]. This burgeoning demand poses a formidable challenge, one that lies at the very heart of sustainable development and resource management: forecasting electricity demand with both precision and reliability [6].

The ASEAN region, with its unique mosaic of emerging economies and varied energy landscapes, presents a distinct and complex set of challenges in the realm of electricity demand forecasting. Several dynamic factors fueled this complexity [7]. Rapid urbanization, a phenomenon witnessed across the region, brings with it an escalating demand for energy in urban centers. Industrial sectors, which are in varying stages of development across different ASEAN countries, contribute to a fluctuating and often unpredictable pattern of electricity consumption [8]. Additionally, the economic conditions within these nations, characterized by both periods of rapid growth and instability, further contribute to the intricacies of forecasting electricity demand [9].

Traditional forecasting methods have played a pivotal role in anticipating energy requirements. These methods, grounded in historical consumption data and straightforward predictive algorithms, have provided a foundation for understanding and planning for future energy needs. However, as the energy landscape evolves, these traditional approaches often fall short [10]. They struggle to encapsulate the full breadth of variables that now influence electricity demand [11]. Variables such as economic policies, technological advancements, environmental considerations, and social changes introduce a level of complexity that traditional models are ill-equipped to handle. This gap in forecasting capability highlights a pressing need for more advanced predictive techniques [12]. These advanced methodologies, leveraging the latest developments in data analytics and machine learning, are capable of assimilating the multifaceted and interconnected variables that define the energy landscape of the ASEAN region. By doing so, they promise to deliver forecasts that are not only more accurate but also more reflective of the dynamic nature of the region.

The realm of predictive modeling encompasses a variety of techniques, each with its unique strengths and applications across diverse fields. Among these, the Autoregressive Integrated Moving Average (ARIMA) model stands out as a foundational tool in time series forecasting. The ARIMA model, originated by Box and Jenkins [13], has gained recognition for its robust application in economic trend analysis [14]. This is evident from its widespread use in predicting stock market trends and assessing economic indicators. In the field of meteorology, ARIMA models have been pivotal in forecasting weather patterns, a testament to their reliability in handling time-dependent data [15]. Additionally, in public health, ARIMA models have been instrumental in predicting disease outbreaks and patient admission rates, showcasing their utility in diverse temporal data contexts [16] [17] [18] [19]. Moreover, the ARIMA model is effectively applied in forecasting electricity consumption in the Philippines, aiding in accurate prediction of energy demand patterns and facilitating efficient energy management and grid planning [20] [21]. The Grey Model, characterized by its ability to make effective predictions with limited or incomplete data, is another significant tool in the predictive modeling arsenal [22]. It has been particularly advantageous in marine economics and management analysis, where it is able to handle multi-variable systems [23]. Moreover, the Grey Model was effectively applied in studying COVID-19 trends in China, utilizing various forecasting models, including the GMQP (1,1), to analyze confirmed cases, deaths, and recovery trends, proving particularly useful during the early outbreak stages with limited data [24]. Another application of the Grey Model is in the prediction of material properties under changing conditions. In the study of Wu et al. [1], the Grey Model was used to predict the tensile strength of materials at temperatures ranging from 400 degrees Celsius to 1100 degrees Celsius.

In contrast, the Neural Network Autoregressive (NNAR) model represents the synergy of traditional time series analysis and the advanced capabilities of neural networks [25]. This hybrid model has been effectively employed in forecasting direct economic losses from earthquakes in Indonesia. This method is significant for providing policymakers with critical information to anticipate future losses from such natural disasters [26]. Additionally, in the telecommunications sector, NNAR models have been adept at load forecasting, a critical function for network capacity planning and optimization [27]. Echo State Neural Networks, emerging from the domain of reservoir computing, demonstrate exceptional efficiency in learning and predicting dynamic temporal patterns [28]. These networks have found their niche in applications such as speech recognition, where they process and predict complex audio sequences [29]. In the finance sector, their ability to handle the non-linear dynamics of financial time series has made them valuable for predicting stock market fluctuations [30]. Furthermore, their application in robotics, particularly in dynamic systems modeling, underscores their versatility in handling time-varying data [31]. The Autoregressive Support Vector Machine (AR-SVM) combines the time series analysis capabilities of autoregressive models with the robustness and precision of Support Vector Machines [32]. This fusion has been beneficial in predicting financial time series, where it navigates the volatility and non-linearity inherent in financial markets [33]. In the biomedical field, SVM models are used for analyzing biomedical signals, contributing to advancements in health monitoring and diagnostics [34]. Additionally, their application in forecasting health-related events, such as disease outbreaks and patient admissions, highlights their relevance in public health planning and response [35]. Lastly, Extreme Gradient Boosting (XGBoost), renowned for its efficiency and performance in machine learning, has made significant inroads in various applications [36]. Its use in credit scoring has revolutionized the way financial institutions assess creditworthiness, enhancing accuracy and reducing risk [37]. In the arena of customer behavior analysis, XGBoost aids businesses in understanding and predicting consumer preferences, driving marketing strategies [38]. Moreover, its application in recommendation systems exemplifies its capability to handle large datasets and complex algorithms, making it a cornerstone in the field of machine learning and data science [39]. These predictive models, with their diverse applications, not only underscore the breadth of their capabilities but also highlight their potential utility in the specific context of electricity demand forecasting in the ASEAN region.

The diverse applications of these models, ranging from financial markets to public health, demonstrate their adaptability and robustness in analyzing and predicting complex patterns in time-series data. Such versatility is crucial in addressing the dynamic and often unpredictable nature of electricity demand, especially in the rapidly evolving ASEAN region. This observation seamlessly leads us to the core objectives of our study. The primary objective of our research is to rigorously evaluate and compare the performance of these diverse models—ARIMA, Grey Model, Neural Network Autoregressive, Echo State Neural Network, Autoregressive Support Vector Machine, and XGBoost—in the specific context of per capita electricity consumption. By doing so, we aim to identify the model that best encapsulates the intricacies of electricity demand patterns in these countries. Our study leverages the extensive applications of these predictive models in various fields to address a vital question: How can these proven models be effectively utilized in forecasting electricity demand in the ASEAN region? The answer to this question holds the potential to significantly influence energy policy and planning, ensuring sustainable and efficient fulfillment of the energy needs of the region.

2. MATERIALS AND METHODS

2.1 DATASET

The dataset utilized in this research is the electric power consumption dataset (measured in kWh per capita), sourced from the World Development Indicator database managed by the World Bank. This dataset encompasses a comprehensive aggregation of per capita electric power consumption data, extending from 1971 through 2014. It should be noted, however, that for Cambodia, the dataset is confined to the years 1995 to 2014. The scope of the study encompasses nine ASEAN nations: Indonesia, Thailand, Singapore, Vietnam, Malaysia, the Philippines, Myanmar, Cambodia, and Brunei Darussalam, selected based on data accessibility. The exclusion of Lao PDR from the study was due to the unavailability of relevant data. In the study conducted by Parreño [2], an examination of the stationarity of electric power consumption across each nation was undertaken. It was discovered that, with the notable exception of Brunei, the electric power consumption patterns in most ASEAN countries exhibited nonstationary characteristics. This finding of nonstationarity was a critical consideration in the development of the forecasting models employed in this research, ensuring that the models accurately reflect the dynamic nature of electric power consumption within these nations.

2.2 ARIMA MODELS FOR REGIONAL FORECASTING

The ARIMA model, an acronym for Autoregressive Integrated Moving Average, is a prominent forecasting tool used for time series data. The model is based on the concept that past values of a time series can be used to predict future values. Its popularity stems from its ability to model various standard temporal structures found in time series data. The methodology of the ARIMA model hinges on three key parameters: p (the order of the autoregressive term), d (the number of differencing required to make the time series stationary), and q (the order of the moving average term). The ARIMA model can be represented mathematically [19] as,

$$Y_t = \mu + \sum_{j=1}^{p} \phi_j Y_{t-j} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_j, \quad \forall t \in \mathbb{Z} \tag{1}$$

Before applying the ARIMA model, it is crucial to check whether the time series is stationary. This means its statistical properties like mean, variance, and autocorrelation are constant over time. The Augmented Dickey-Fuller (ADF) test is used for this purpose. If the series is not stationary, differencing is applied to transform it into a stationary series. The next step is to determine the order of the ARIMA model, denoted as ARIMA (p,d,q), The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were used for identifying p and q. After identifying the order of the model, the next step is to estimate its parameters. This was done using the Akaike Information Criterion (AIC), which helps in selecting a model that balances goodness of fit with simplicity. Once the model is fitted, the diagnostic checks were performed to ensure the residuals are white noise. This was done using plots the Residuals Autocorrelation and Partial Autocorrelation, as well as the Ljung-Box test [15].

2.3 GREY MODEL

In this paper, we used the GM(1,1) variant of the grey model. In forecasting scenarios where data is limited or incomplete, the grey model, especially the GM(1,1) variant, serves as a powerful tool. The methodology of the Grey GM(1,1) model involves several key steps. Initially, the model considers a nonnegative sequence as the basis for generating a 1-AGO (Accumulating Generation Operator) sequence, which is essential for the construction of the model. This is done to reveal any hidden patterns in the data. The new sequence, $X^{(1)}$, is generated from the original data sequence, $X^{(0)}$, as follows,

$$X^{(1)}(i)=\sum_{j=1}^{i} X^{(0)}(j), \quad i=1,2,\ldots,n \tag{2}$$

where n is the number of data points. Subsequently, the GM(1,1) model is constructed, governed by the differential equation

$$\frac{dX^{(1)}}{dt}+aX^{(1)}=b \tag{3}$$

where a and b are the model parameters estimated, and $\frac{dX^{(1)}}{dt}$ is the rate of change of the accumulated sequence. These parameters were estimated using the least square method, with the equation linearized for ease of computation. The differential equation is solved to provide a forecast. The solution of the GM(1,1) model is given by,

$$\hat{X}^{(1)}(i+1)=\left(X^{(0)}(1)-\frac{b}{a}\right)e^{-ai}+\frac{b}{a} \tag{4}$$

Post-solution, an inverse transformation is applied to revert the forecasted data to its original scale [40].

2.4 NEURAL NETWORK AUTOREGRESSIVE

The Neural Network Autoregressive model, represented as NNAR(p,k), is a forecasting method that integrates neural networks with time series data. It utilizes past values (lags) of the time series as inputs to a neural network. The model combines these inputs linearly with weights and biases. Mathematically, the input combination is given by,

$$z_j=b_j+\sum_{i=1}^{n} w_{i,j}x_i \tag{5}$$

where, z_j is the input to the j^{th} neuron, b_j is the bias term for that neuron, $w_{i,j}$ are the weights, and x_i are the input values. The inputs are then transformed using sigmoid function, a nonlinear function denoted by

$$s(z)=\frac{1}{1+e^{-z}} \tag{6}$$

This function maps the input to a value between 0 and 1, adding non-linearity to the model. This nonlinearity allows the model to capture complex relationships in the data. The NNAR(p,k) notation indicates p lagged inputs and k neurons in the hidden layer. The parameters of the network are learned from the data through backpropagation algorithms that aim to minimize error measures. Regularization techniques, specifically using a decay parameter, are then applied to prevent overfitting [41].

2.5 ECHO STATE NETWORK

The Echo State Network (ESN) is a type of recurrent neural network that is specifically tailored for time series prediction. Its unique feature is the reservoir, a large and randomly initialized network of neurons. The reservoir updates its state at each time step, combining information from the current input and its previous state using tanh function, a nonlinear activation function. This update process is represented by the equation,

$$x(t)=f\left(W_{\text{in}}u(t)+Wx(t-1)\right) \tag{7}$$

where $x(t)$ is the reservoir state at time t, $u(t)$ is the input, W_{in} are the input weights, W is the recurrent weight matrix, and f is the activation function. The output of the ESN is computed using a readout function that linearly combines the state of the reservoir. This can be described by the equation

$$y(t)=W_{\text{out}}x(t) \tag{8}$$

where $y(t)$ is the output and W_{out} are the output weights. These weights are trained using linear regression methods, minimizing the error between the output of the network and the actual data.

Additionally, our ESN implementation include a leakage rate, denoted as a, to control the update speed of the reservoir's state. The state update equation incorporating the leakage rate is

$$x(t)=(1-a)x(t-1)+af\left(W_{\text{in}}u(t)+Wx(t-1)\right) \tag{9}$$

This highlights how this rate moderates the influence of new inputs versus the existing state of the reservoir [42] [43].

2.6 AUTOREGRESSIVE SUPPORT VECTOR MACHINE

Support Vector Regression (SVR) is a sophisticated machine learning method used primarily for forecasting and regression tasks. It operates by determining a function, denoted as

$$f(x)=\langle w,x\rangle+b \tag{10}$$

This aims to capture the relationship between the input features x and the output. In this function, $\langle w,x\rangle$ represents the dot product of the weight vector w and the feature vector x, with b being the bias term. A key aspect of SVR is its focus on minimizing $\frac{1}{2}\|\omega\|^2$, balancing the complexity of the model against its ability to tolerate deviations greater than a certain threshold, ϵ. This balancing act incorporates an ϵ-insensitive loss function, where deviations within a margin of ϵ from the actual values are not penalized. The formulation uses slack variables ξ_i and ξ_i* for each data point to manage this aspect, and involves a regularization parameter C, which controls the trade-off between the flatness of the function and the amount up to which deviations larger than ϵ are tolerated. In cases where the relationship between the input and output is non-linear, SVR employs the kernel trick, using functions like linear, polynomial, RBF, or sigmoid kernels. These kernels transform the input data into a higher-dimensional space where linear separation is possible. The kernel function, denoted as $K(x_i,x_j)$, essentially calculates the dot product of the vectors in this higher-dimensional space.

The actual regression function in SVR is derived from the solution to an optimization problem involving Lagrange multipliers, represented by α_i and α_i^*. The final regression function is expressed as

$$f(x)=\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^*\right)K(x_i,x)+b \tag{11}$$

which is used for making predictions or forecasts [32].

2.7 XGBOOST

XGBoost, or eXtreme Gradient Boosting, is a sophisticated machine learning algorithm that is particularly effective for forecasting in time series analysis, thanks to its mathematical foundation. At the heart of XGBoost is an objective function that combines a loss function and a regularization term. This function is expressed as

$$\text{Obj}(\theta)=\sum_{i=1}^{n}l(y_i,\hat{y}_i)+\sum_{k=1}^{K}\Omega(f_k) \tag{12}$$

where $l(y_i,\hat{y}_i)$ quantifies the difference between the predicted and actual values, and $\Omega(f_k)$ is the regularization term that penalizes complexity to avoid overfitting. During training, XGBoost expands the loss function to its second-order Taylor series, utilizing both the first-order (gradient) and second-order (Hessian) derivative information. This approach significantly enhances the performance of the model over traditional gradient boosting methods. The regularization term in the objective function,

$$\Omega(f_k)=\gamma T+\frac{1}{2}\lambda\|\omega\|^2 \tag{13}$$

where T is the number of leaves, ω is the scores on the leaves, γ is the parameter for tree complexity, and λ is the L2 regularization on leaf weights, is key in preventing overfitting.

XGBoost builds its model additively. In each iteration, it introduces a new tree that optimizes the objective function, following

$$\hat{y}_i^{(t)}=\hat{y}_i^{(t-1)}+\eta f_t(x_i) \tag{14}$$

with η being the learning rate and $f_t(x_i)$ the output of the new tree. When applying XGBoost to the ASEAN electric consumption data, the time series data is first transformed into a format suitable for machine learning. This process involves creating lagged variables to encapsulate the temporal dependencies of the time series. The model is trained on this dataset, tuning hyperparameters like the learning rate, depth of trees, and number of trees to enhance performance. The effectiveness of the model is then evaluated using Mean Absolute Error and Root Mean Square Error [17] [44] [45].

3. RESULTS AND DISCUSSION

3.1 ARIMA MODEL RESULTS

Since the electric consumption data for most ASEAN nations are nonstationary [2], the time series data need to be differenced first before building the ARIMA model. The Table.1 shows the fitted ARIMA models for electric power consumption for each ASEAN nations. The variation in the ARIMA model parameters (p,d,q) across countries highlights the differing dynamics in the electric power consumption trends in each nation. The performance of the ARIMA model, indicated by the MAPE values, varies from about 2.78% for Singapore to over 6.9% for Cambodia, reflecting the differing levels of model accuracy in capturing power consumption trends.

Table.1. Performance of Fitted ARIMA Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	ARIMA	RMSE	MAPE
Brunei Darussalam	ARIMA(1,1,0)	345.6957	5.39317%
Cambodia	ARIMA(1,2,0)	8.439079	6.900228%
Indonesia	ARIMA(2,2,0)	8.963071	4.442399%
Malaysia	ARIMA(3,1,0)	105.1567	3.033633%
Myanmar	ARIMA(2,2,2)	6.135964	6.423116%
Philippines	ARIMA(1,1,0)	21.7223	4.585378%
Singapore	ARIMA(1,1,1)	172.5642	2.784062%
Thailand	ARIMA(2,1,0)	50.67693	3.680094%
Vietnam	ARIMA(1,2,0)	13.64435	4.59009%

3.2 GREY MODEL RESULTS

Table 2 presents the performances of the fitted Grey Models for forecasting electric power consumption in various ASEAN nations. The results reveal variety of dynamics in the electric power consumption patterns across the region. For instance, the model parameter for Indonesia are a=-0.06904143 and b=60.95087, resulting in an RMSE of 135.4473 and a high MAPE of 91.84311%, suggesting a notable discrepancy between the actual and fitted values, indicating significant variability. In contrast, Cambodia, with a=-0.1476733 and b=14.95296, shows a much lower RMSE of 5.598126 and a MAPE of 4.908046%, indicating a closer fit between the output of the model and the actual data. Notably, the grey model for Singapore, with parameters a=-0.03453686 and b=2457.047, exhibits an RMSE of 821.288 and a MAPE of 19.13825%. This reflects the complexity and high fluctuation in electric power consumption patterns in Singapore. The performance of the grey model, as indicated by the RMSE and MAPE values, varies significantly across the ASEAN countries, reflecting the distinct characteristics and challenges in modeling electric power consumption in each nation. For instance, the Philippines shows a relatively low MAPE of 6.602833%, indicating a relatively more consistent performance of the model in this country.

Table.2. Performance of Fitted Grey Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	a	b	RMSE	MAPE
Brunei Darussalam	-0.03893318	2166.069	949.8664	25.26895%
Cambodia	-0.1476733	14.95296	5.598126	4.908046%
Indonesia	-0.06904143	60.95087	135.4473	91.84311%
Malaysia	-0.05298033	539.2747	314.7542	21.45852%
Myanmar	-0.05694217	11.42241	18.96123	22.05486%
Philippines	-0.02299563	249.1481	30.68312	6.602833%
Singapore	-0.03453686	2457.047	821.288	19.13825%
Thailand	-0.05451121	315.3965	269.1985	38.62651%
Vietnam	-0.1101007	-2.60247	451.0755	88.81626%

3.3 NEURAL NETWORK AUTOREGRESSIVE RESULTS

The Table.3 outlines the performance of NNAR models for electric power consumption forecasting across various ASEAN nations. Electric power consumption data of each nation was modeled using the NNAR(1,1) configuration. The NNAR models, averaged over 20 networks, each being a 1-1-1 network with 4 weights, demonstrated varying degrees of forecasting accuracy among the ASEAN countries. For example, Indonesia demonstrated an estimated variance (σ^2) of 93.57, along with an RMSE of 9.67327 and a MAPE of 6.390644%, indicating a moderate level of accuracy in the fit of the model. In contrast, NNAR model for Brunei exhibited a much higher σ^2 of 106811, suggesting greater variability in the fitted model. NNAR model for Singapore, with a σ^2 of 21014, managed to achieve a relatively low MAPE of 2.085161%. This suggests that despite the high variance, the model was able to capture the general trend of electric power consumption with reasonable accuracy. Meanwhile, countries like the Philippines and Myanmar, with their lower estimated variances and RMSEs, indicate different levels of model fit quality and predictability in their electric power consumption patterns.

Table.3. Performance of Fitted NNAR Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	NNAR	σ^2	RMSE	MAPE
Brunei Darussalam	NNAR(1,1)	106811	326.8197	6.596919%
Cambodia	NNAR(1,1)	57.37	7.574161	6.467324%
Indonesia	NNAR(1,1)	93.57	9.67327	6.390644%
Malaysia	NNAR(1,1)	12057	109.8058	3.356124%
Myanmar	NNAR(1,1)	38.15	6.176324	5.724485%
Philippines	NNAR(1,1)	350.8	18.73017	3.890609%
Singapore	NNAR(1,1)	21014	144.9632	2.085161%
Thailand	NNAR(1,1)	2119	46.03151	3.150477%
Vietnam	NNAR(1,1)	155.1	12.45478	4.356473%

3.4 ECHO STATE NETWORK RESULTS

The Table.4 details the performance of fitted ESN models for electric power consumption in various ASEAN nations. Most nations, specifically Indonesia, Thailand, Singapore, Vietnam, Malaysia, Philippines, Myanmar, and Brunei Darussalam, utilized an ensemble of 34 ESN models with the configuration ESN (1,17,1), indicating 1 lag, 17 reservoir nodes, and a single output. Cambodia, however, employed 16 models with the configuration ESN (1,8,1). These differences in model configurations are reflective of the distinct power consumption patterns and data characteristics in each country. For instance, the ESN model for Indonesia, ESN (1,17,1), yielded an RMSE of 21.37499 and a MAPE of 13.81274%, indicating a moderate level of alignment between the outputs of the model and the actual values. ESN model for Thailand demonstrated an RMSE of 58.40638 and a MAPE of 8.428777%, suggesting a relatively higher level of model accuracy. ESN model for Singapore showed a higher RMSE of 184.3397, but a lower MAPE of 4.01458%, indicating that while the magnitude of errors was larger, the proportion of errors relative to actual values was smaller. In the case of Cambodia, which used a smaller ensemble of ESN models (ESN (1,8,1)), the RMSE was 14.99147 with a MAPE of 21.30261%. This higher MAPE indicates less accuracy in the fit of the model compared to other nations.

Table.4. Performance of Fitted ESN Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	ESN	RMSE	MAPE
Brunei Darussalam	ESN(1,17,1)	224.3887	5.545757%
Cambodia	ESN(1,8,1)	14.99147	21.30261%
Indonesia	ESN(1,17,1)	21.37499	13.81274%
Malaysia	ESN(1,17,1)	105.0565	7.253703%
Myanmar	ESN(1,17,1)	6.681808	7.053579%
Philippines	ESN(1,17,1)	9.937318	2.338732%
Singapore	ESN(1,17,1)	184.3397	4.01458%
Thailand	ESN(1,17,1)	58.40638	8.428777%
Vietnam	ESN(1,17,1)	49.75843	11.54566%

3.5 AUTOREGRESSIVE SUPPORT VECTOR MACHINE RESULTS

The Table.5 provides a detailed analysis of the performance of AR-SVM models for electric power consumption in ASEAN nations. AR-SVM model for Indonesia, utilizing 7 support vectors, shows an RMSE of 16.95347 and a MAPE of 22.23923%, indicating the level of deviation of the predictions of the model from the actual values. In contrast, AR-SVM model for Thailand, with 13 support vectors, has an RMSE of 59.97657 and a lower MAPE of 9.129444%, suggesting a different scale of error in the output of the model relative to actual consumption data. The AR-SVM model for Singapore, using 9 support vectors, demonstrates a significantly higher RMSE of 191.6974, yet maintains a low MAPE of 3.982123%. This suggests that while the absolute errors might be larger, they are relatively small compared to the actual values, indicating a good model fit in proportionate terms. In the case of the Philippines, which employed 22 support vectors, the model achieved an RMSE of 20.79329 and a MAPE of 4.164425%, suggesting a relatively more accurate fit. On the other hand, Brunei Darussalam, with 18 support vectors, exhibited the highest RMSE in the table at 371.8799, and a MAPE of 8.101263%, indicating considerable variability in the fit of the model.

Table.5. Performance of Fitted AR-SVM Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	Number of Support Vectors	RMSE	MAPE
Brunei Darussalam	18	371.8799	8.101263%
Cambodia	8	12.3071	12.38234%
Indonesia	7	16.95347	22.23923%
Malaysia	9	126.5497	6.850145%
Myanmar	17	10.5982	7.792609%
Philippines	22	20.79329	4.164425%
Singapore	9	191.6974	3.982123%
Thailand	13	59.97657	9.129444%
Vietnam	7	38.29419	24.01097%

3.6 XGBOOST RESULTS

The Table.6 showcases the performance of fitted XGBoost models for electric power consumption in ASEAN nations. The XGBoost model for Indonesia, which underwent 100 iterations, started with a train RMSE of 329.2199653 and progressively improved to 0.7431893. This dramatic improvement indicates the efficiency of the model in learning and adjusting to the characteristics of the dataset. The final RMSE of 0.7431934 and an exceptionally low MAPE of 0.003123305% reflect the high accuracy of the model in fitting the electric power consumption data. Similarly, the model for Thailand demonstrated significant progress during training, reducing the train RMSE from an initial 1206.978461 to 1.448969. This indicates the capability of the model to adapt and learn from the dataset, resulting in a final RMSE of 1.448977 and a very low MAPE of 0.001210328%. XGBoost model for Singapore, starting with a train RMSE of 5428.63625, showed considerable improvement, eventually reaching an RMSE of 10.47943. Despite the higher RMSE compared to other nations, the low MAPE of 0.001332788% suggests that the model could capture the general trend of the data effectively. The model for Vietnam also exhibited consistent improvement, with the train RMSE decreasing from 465.090110 to 1.215772, leading to a final RMSE of 1.215777 and a MAPE of 0.002943752%. This trend of significant reduction in RMSE during the training process was observed across all nations, indicating the robustness of XGBoost in modeling and fitting complex datasets. These results underscore the effectiveness of XGBoost models in capturing the nuances of electric power consumption patterns. The consistently low MAPE values across all countries further reinforce the accuracy and reliability of the model, making XGBoost a compelling tool for predictive modeling in diverse contexts.

Table.6. Performance of Fitted XGBoost Models for Electric Power Consumption in ASEAN Nations

ASEAN Nation	RMSE	MAPE
Brunei Darussalam	11.2106	0.001738513%
Cambodia	0.3674934	0.001363503%
Indonesia	0.7431934	0.003123305%
Malaysia	5.8119	0.002816197%
Myanmar	0.2413716	0.001959335%
Philippines	0.8513808	0.001186033%
Singapore	10.47941	0.001332788%
Thailand	1.448977	0.001210328%
Vietnam	1.215777	0.002943752%

3.7 COMPARISON OF MODELS

The Fig.1 visually contrasts the model-derived estimates against the actual electric power consumption data in Brunei Darussalam. It can be observed that the XGBoost model outshines others. This superiority is further underscored by its lowest RMSE and MAPE values, highlighting its exceptional accuracy in the dataset. Compared to the ARIMA and Grey models, which are traditionally strong in linear trends and small datasets respectively, the ability of XGBoost to handle non-linearities and complex interactions gives it an edge. NNAR and ESN models, despite being good at capturing non-linear trends, do not match the performance of XGBoost.

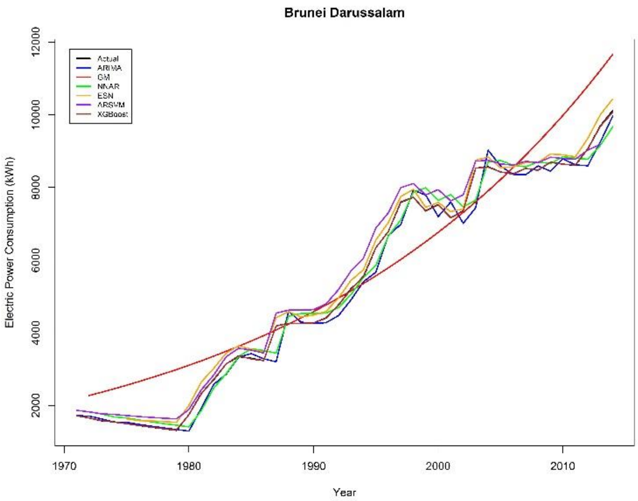

Fig.1. Graphical comparison of the fitted values from different models with Brunei Darussalam electric consumption data.

The line graph presented in Fig.2 offers a visual comparison between the actual electric power consumption data of Cambodia and the fitted values from various models, effectively highlighting the differences in accuracy. The superiority of XGBoost model is clearly demonstrated, as it aligns most closely with the actual consumption figures, a fact underscored by its lowest RMSE and MAPE. While ARIMA and Grey models have their merits in certain scenarios, they fall short in achieving the same accuracy level. Similarly, although the NNAR and ESN models are effective in a range of applications, they do not match the efficiency demonstrated by XGBoost.

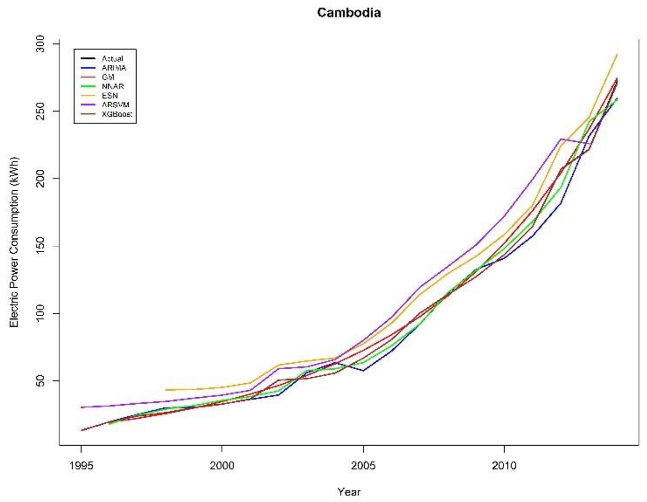

Fig.2. Graphical comparison of the fitted values from different models with Cambodia electric consumption data

In Fig.3, a detailed line graph comparison showcases how different fitted values of the models align with the actual electric consumption data of Indonesia, revealing the effectiveness of each model. Moreover, the graph demonstrates the notable proficiency of the XGBoost model in Indonesia, as indicated by its fitted line closely mirroring the actual values, along with the lowest RMSE and MAPE, signifying its superior predictive accuracy. ARIMA and Grey models, despite their usefulness in specific conditions, do not provide the same level of accuracy. NNAR and ESN models, although capable in various applications, do not reach the efficiency of XGBoost.

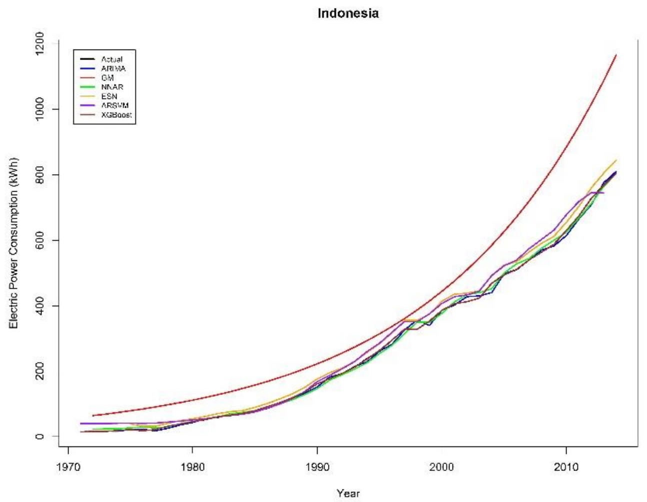

Fig.3. Graphical comparison of the fitted values from different models with Indonesia electric consumption data

For Malaysia, XGBoost again demonstrates the best performance. The complex consumption patterns of the country might be challenging for the linear approaches of ARIMA and the assumptions of the Grey model. NNAR and ESN, while effective in capturing non-linear relationships, do not match the accuracy of XGBoost (Fig.4).

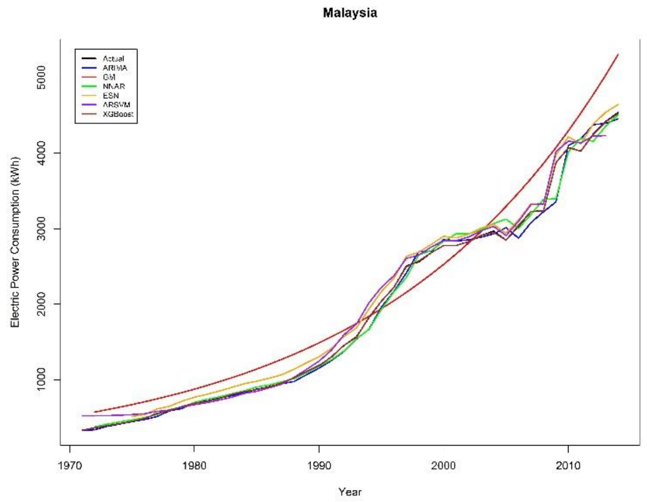

Fig.4. Graphical comparison of the fitted values from different models with Malaysia electric consumption data

In Myanmar, the performance of XGBoost is unmatched, indicating its strong adaptability and accuracy. The ARIMA model, though useful for certain types of time series data, and the Grey model, ideal for small or incomplete datasets, are not as effective. NNAR and ESN models also lag behind in comparison to the capabilities of XGBoost. This finding is illustrated in Fig.5.

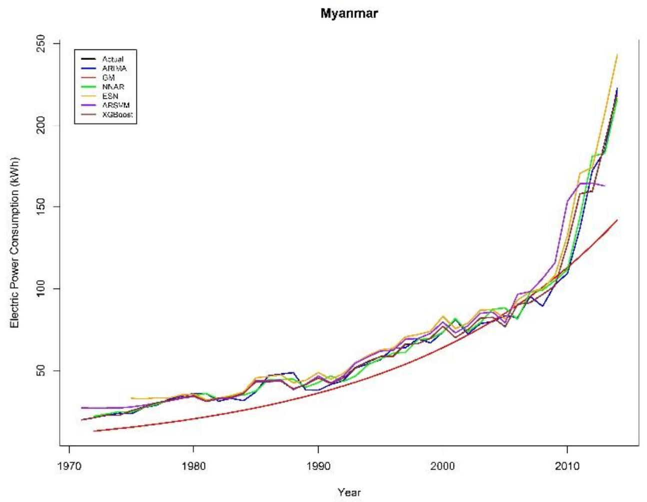

Fig.5. Graphical comparison of the fitted values from different models with Myanmar

The Philippines sees the best results with the XGBoost model. Its ability to manage large and complex datasets gives it a clear advantage over ARIMA and Grey models. NNAR and ESN models, while beneficial in specific contexts, do not reach the same level of predictive power as XGBoost in this case (Fig.6).

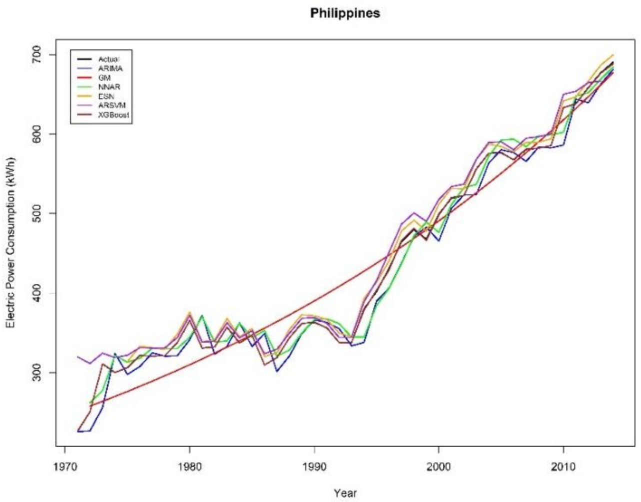

Fig.6. Graphical comparison of the fitted values from different models with Philippines electric consumption data

The diagram in Fig.7 provides a comparative graphical analysis of electricity consumption of Singapore against the fitted values from different predictive models. The XGBoost leads in predictive accuracy. Traditional models like ARIMA struggle with the complexity of the data, and the Grey model is limited by its inherent assumptions of linearity. NNAR and ESN models, which are generally more adaptable to non-linear data, still do not match the comprehensive learning algorithm of XGBoost.

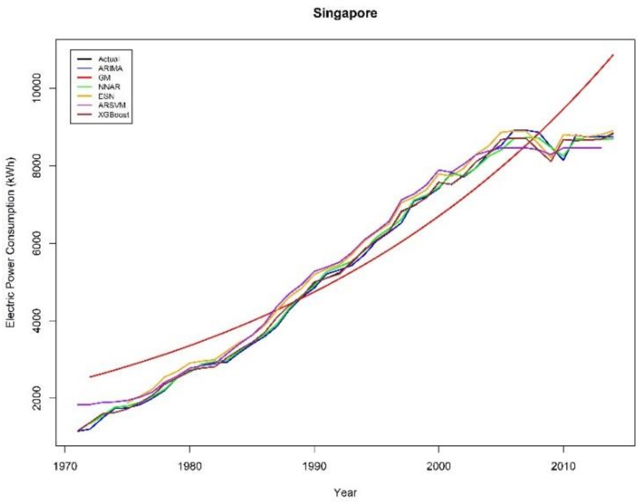

Fig.7. Graphical comparison of the fitted values from different models with Singapore electric consumption data

In Thailand, the superior performance of XGBoost is again evident. While the ARIMA model is decent for linear trends and the Grey model can be beneficial for datasets with limited information, their accuracy falls short compared to the advanced machine learning capabilities of XGBoost. NNAR and ESN models, though capable of handling complex patterns, are outperformed by the robustness and precision of XGBoost (Fig.8).

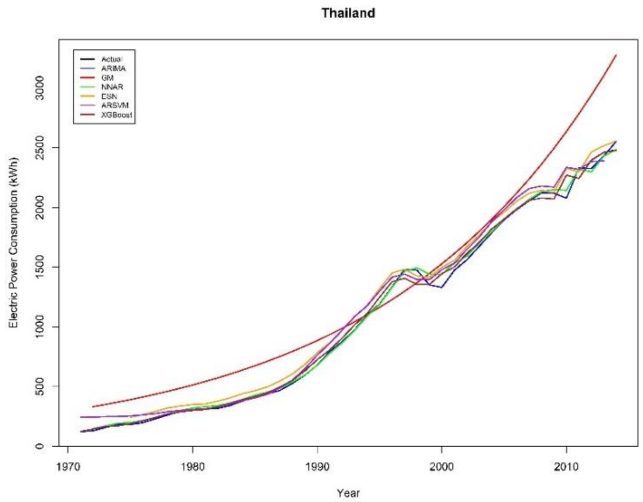

Fig.8. Graphical comparison of the fitted values from different models with Thailand electric consumption data

The XGBoost model again shows its prowess in Vietnam, surpassing other models in forecasting accuracy. While ARIMA and Grey models have their merits in simpler scenarios, they fall behind in capturing the intricate consumption patterns that XGBoost manages to model effectively (Fig.9).

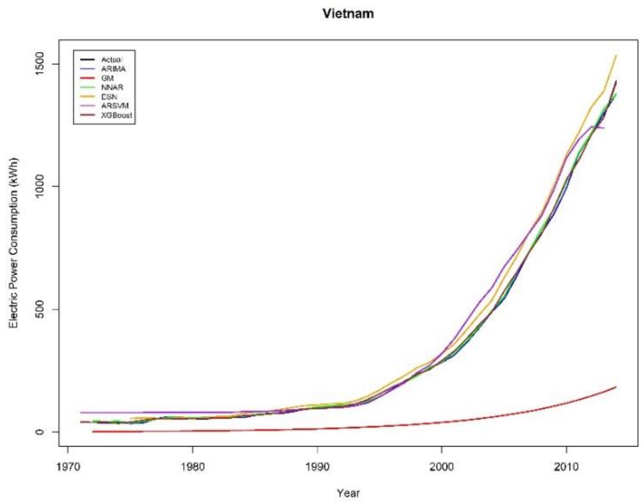

Fig.9. Graphical comparison of the fitted values from different models with Vietnam electric consumption data

In this study, we examined several predictive models for electricity demand forecasting in the ASEAN region, each with its own set of assumptions and limitations. The ARIMA model, known for its linear approach, assumes future values of a time series as a function of past observations and random errors. The effectiveness of this model might be compromised in rapidly changing economies due to its reliance on historical data and the requirement for stationarity in time series data [46]. In contrast, the Grey Model, designed for limited or incomplete data sets, may not fully exploit its strengths in contexts where comprehensive data is available [22]. It simplifies complex systems, potentially overlooking the multifaceted factors influencing electricity consumption in ASEAN countries. The NNAR model, capable of capturing non-linear relationships, faces challenges of complexity, potential overfitting, and being data and computationally intensive. Its "black box" nature could pose interpretational challenges for policy formulation [47] [26]. Similarly, the ESN model, while innovative, is sensitive to initial conditions and parameters. The performance of the model heavily depends on the size of reservoir, and there is a risk of instability, especially with volatile data like electricity consumption patterns [48]. The AR-SVM requires careful kernel selection and parameter tuning, presenting complexity in application. Its predictions, while robust, offer limited interpretability, a common drawback in machine learning models [49]. Finally, XGBoost, known for its efficiency in handling complex datasets, necessitates meticulous hyperparameter tuning. While it includes regularization to prevent overfitting, improper tuning or lack of data diversity can still lead to overfitting. It is also computationally intensive, posing challenges in resource-constrained scenarios [17] [45].

4. CONCLUSIONS

This study presents a comprehensive evaluation of various predictive models for forecasting electricity demand in the ASEAN region. The ARIMA, Grey Model, NNAR, ESN, AR-SVM, and XGBoost models each exhibit unique strengths and limitations in handling the intricate dynamics of electricity consumption patterns. While traditional models like ARIMA and Grey Model provide foundational insights, advanced models like XGBoost demonstrate superior accuracy and adaptability in capturing complex, non-linear patterns in the data. The findings of this research not only offer valuable insights into the effectiveness of these models in the context of electricity demand forecasting but also underscore the importance of selecting appropriate predictive tools to inform energy policy and planning in rapidly evolving economies. As the ASEAN region continues to grow and diversify, the ability to accurately forecast electricity demand will remain a critical component in ensuring sustainable and efficient energy management.

Future work should focus on integrating real-time data and IoT-based analytics to enhance the accuracy of electricity demand forecasts in the ASEAN region. Exploring the incorporation of socio-economic and environmental variables can provide a more holistic understanding of consumption patterns. Additionally, the development of user-friendly interfaces for these predictive models would facilitate wider adoption and application by policymakers and energy sector stakeholders.

5. AVAILABILITY OF DATA AND MATERIALS

The datasets analyzed and R scripts used to execute the methodology are available at https://github.com/samjohnparr/ASEAN_Predictive_Models.

REFERENCES

[1] W.Z. Wu, J. Jiang and Q. Li, “A Novel Discrete Grey Model and its Application”, *Mathematical Problems in Engineering*, Vol. 2019, pp. 1-6, 2019.

[2] S.J. Parreno, “Assessing the Stationarity of Per Capita Electricity Consumption: Time Series Analysis in ASEAN Countries”, *International Journal of Energy Economics and Policy*, Vol. 14, No. 1, pp. 46-52, 2024.

[3] Y. Hakimah, A.T. Nugraha, A. Surya, A.R.W. Ananda and P. Astuty, “Does the Electricity Consumption Determine the Economic Growth and Energy Prices in ASEAN Countries?”, *Journal of Security and Sustainability Issues*, Vol. 9, No. 2, pp. 489-503, 2019.

[4] Y. Liu, Z. Sheng and D. Azhgaliyeva, “Toward Energy Security in ASEAN: Impact of Regional Integration, Renewables and Energy Efficiency”, *Achieving Energy Security in Asia: Diversification, Integration and Policy Implications*, Vol. 87, pp. 291-318, 2019.

[5] D.H. Vo and A.T. Vo, “Renewable Energy and Population Growth for Sustainable Development in the Southeast Asian Countries”, *Energy, Sustainability and Society*, Vol. 11, No. 1, pp. 1-30, 2021.

[6] N.A. Utama, K.N. Ishihara, Q. Zhang and T. Tezuka, “2050 ASEAN Electricity Demand: Case Study in Indonesia and Cambodia”, *Zero-Carbon Energy Kyoto 2010, Green Energy and Technology*, Vol. 32, pp. 32-39, 2011.

[7] M.A. Al Matin, S. Takeda, Y. Tanaka, S. Sakurai and T. Tezuka, “LCOE Analysis for Grid-Connected PV Systems of Utility Scale Across Selected ASEAN Countries”, *Economic Research Institute for ASEAN and East Asia*, Vol. 78, pp. 1-50, 2019.

[8] IEA, “Electricity Market Report-December 2020”, Available at: https://www.iea.org/reports/electricity-market-report-december, Accessed in 2020.

[9] E.J.M. Sahid, T.C. Sin and G.C. Hock, “Energy Security in ASEAN Region: Its Challenges”, *IOP Conference Series: Earth and Environmental Science*, Vol. 268, No. 1, pp. 1-8, 2019.

[10] A. Deka, N. Hamta, B. Esmaeilian and S. Behdad, “Predictive Modeling Techniques to Forecast Energy Demand in the United States: A Focus on Economic and Demographic Factors”, *Journal of Energy Resources Technology*, Vol. 138, No. 2, pp. 1-9, 2016.

[11] E. Sharma, “Energy Forecasting based on Predictive Data Mining Techniques in Smart Energy Grids”, *Energy Informatics*, Vol. 1, No. 1, pp. 367-373, 2018.

[12] K. Nti, M. Teimeh, O. Nyarko-Boateng and A.F. Adekoya, “Electricity Load Forecasting: A Systematic Review”, *Journal of Electrical Systems and Information Technology*, Vol. 7, No. 1, pp. 1-19, 2020.

[13] G.E. Box, G.M. Jenkins, G.C. Reinsel and G.M. Ljung, “*Time Series Analysis: Forecasting and Control*”, 2015.

[14] S. Hua, “Back-Propagation Neural Network and ARIMA Algorithm for GDP Trend Analysis”, *Wireless Communications and Mobile Computing*, Vol. 2022, pp. 1-9, 2022.

[15] L. Mao, Y. Huang, X. Zhang, S. Li and X. Huang, “ARIMA Model Forecasting Analysis of the Prices of Multiple Vegetables Under the Impact of the COVID-19”, *PLoS One*, Vol. 17, No. 7, pp. 1-26, 2022.

[16] A. Wagner and K. Cleland, “Using Autoregressive Integrated Moving Average Models for Time Series Analysis of Observational Data”, *BMJ*, Vol. 383, pp. 1-5, 2023.

[17] M.S. Rahman, A.H. Chowdhury and M. Amrin, "Accuracy Comparison of ARIMA and XGBoost Forecasting Models in Predicting the Incidence of COVID-19 in Bangladesh", *PLOS Global Public Health*, Vol. 2, No. 5, pp. 1-13, 2022.

[18] A. Zhao, R. Zhang, H. Zhang and S. He, "Prediction of Global Omicron Pandemic using ARIMA, MLR and Prophet Models", *Scientific Reports*, Vol. 12, No. 1, pp. 1-9, 2022.

[19] S. Siamba, A. Otieno and J. Koech, "Application of ARIMA and Hybrid ARIMA Models in Predicting and Forecasting Tuberculosis Incidences among Children in Homa Bay and Turkana Counties, Kenya", *PLOS Digital Health*, Vol. 2, No. 2, pp. 1-19, 2023.

[20] S.J. Parreño, "Forecasting Electricity Consumption in the Philippines using ARIMA Models", *International Journal of Machine Learning and Computing*, Vol. 13, pp. 1-14, 2022.

[21] S.J. Parreno, "Forecasting the Total Non-Coincidental Monthly System Peak Demand in the Philippines: A Comparison of Seasonal Autoregressive Integrated Moving Average Models and Artificial Neural Networks", *International Journal of Energy Economics and Policy*, Vol. 13, No. 5, pp. 544-552, 2023.

[22] N. Xie, "A Summary of Grey Forecasting Models", *Grey Systems: Theory and Application*, Vol. 12, No. 4, pp. 703-722, 2022.

[23] L. Xuemei, Y. Cao, J. Wang, Y. Dang and Y. Kedong, "A Summary of Grey Forecasting and Relational Models and its Applications in Marine Economics and Management", *Marine Economics and Management*, Vol. 2, No. 2, pp. 87-113, 2019.

[24] Zhang and Z. Jiang, "A New Grey Quadratic Polynomial Model and its Application in the COVID-19 in China", *Scientific Reports*, Vol. 11, No. 1, pp. 1-9, 2021.

[25] O. Triebe, N. Laptev and R. Rajagopal, "AR-Net: A Simple Auto-Regressive Neural Network for Time-Series", *Proceedings of International Conference on Machine Learning*, Vol. 7, pp. 1-12, 2019.

[26] U. Azmi, S. Soehardjoepri and I.M.H. Putri, "Application of Neural Network Autoregression (NNAR) and ARIMA-GARCH based on Interpolation for Forecasting Direct Economic Losses of Earthquake in Indonesia", *AIP Conference Proceedings*, Vol. 2641, No. 1, pp. 1-10, 2022.

[27] A.P. Velasco, D.L.L. Polestico, G.P.O. Macasieb, M.B.V. Reyes and F.B. Vasquez, "Load Forecasting using Autoregressive Integrated Moving Average and Artificial Neural Network", *International Journal of Advanced Computer Science and Applications*, Vol. 9, No. 7, pp. 1-9, 2018.

[28] H. Jaeger, "The Echo State Approach to Analysing and Training Recurrent Neural Networks with an Erratum Note", *German National Research Center for Information Technology GMD Technical Report*, Vol. 148, No. 34, pp. 1-13, 2001.

[29] H. Shrivastava, A. Garg, Y. Cao, Y. Zhang and T. Sainath, "Echo State Speech Recognition", *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, Vol. 98, pp. 5669-5673, 2021.

[30] J. Dan, W. Guo, W. Shi, B. Fang and T. Zhang, "Deterministic Echo State Networks BASED Stock Price Forecasting", *Abstract and Applied Analysis*, Vol. 2014, pp. 1-7, 2014.

[31] A. Racca and L. Magri, "Robust Optimization and Validation of Echo State Networks for Learning Chaotic Dynamics", *Neural Networks*, Vol. 142, pp. 252-268, 2021.

[32] J. Kim, "Financial Time Series Forecasting using Support Vector Machines", *Neurocomputing*, Vol. 55, No. 1, pp. 307-319, 2003.

[33] A. Gupta, M. Pratama, Z. Ma, J. Li and M. Prasad, "Financial Time Series Forecasting using Twin Support Vector Regression", *PLoS One*, Vol. 14, No. 3, pp. 1-7, 2019.

[34] A. Subasi, "Classification of EMG Signals using PSO Optimized SVM for Diagnosis of Neuromuscular Disorders", *Computers in Biology and Medicine*, Vol. 43, No. 5, pp. 576-586, 2013.

[35] H.H. Al-Ahmadi, "The Significance of Software Engineering to Forecast the Public Health Issues: A Case of Saudi Arabia", *Frontiers in Public Health*, Vol. 10, pp. 1-9, 2022.

[36] J. Ma, Z. Yu, Y. Qu, J. Xu and Y. Cao, "Application of the XGBoost Machine Learning Method in PM2.5 Prediction: A Case Study of Shanghai", *Aerosol and Air Quality Research*, Vol. 20, No. 1, pp. 128-138, 2020.

[37] C. Qin, Y. Zhang, F. Bao, C. Zhang, P. Liu and P. Liu, "XGBoost Optimized by Adaptive Particle Swarm Optimization for Credit Scoring", *Mathematical Problems in Engineering*, Vol. 2021, pp. 1-18, 2021.

[38] K. Cai and M.R. Rodavia, "XGBoost Analysis based on Consumer Behavior", *Frontiers in Computing and Intelligent Systems*, Vol. 5, No. 2, pp. 85-89, 2023.

[39] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, Vol. 9, pp. 785-794, 2016.

[40] C. Hsu and C.H. Wang, "Forecasting the Output of Integrated Circuit Industry using a Grey Model Improved by the Bayesian Analysis", *Technological Forecasting and Social Change*, Vol. 74, No. 6, pp. 843-853, 2007.

[41] R.J. Hyndman and G. Athanasopoulos, "*Forecasting: Principles and Practice*", 2021.

[42] A. Sun, J. Lin, C. Yang, X. Yin, Z. Li and P. Guo, "Stock Price Forecasting: An Echo State Network Approach", *Computer Systems Science and Engineering*, Vol. 36, No. 3, pp. 509-520, 2021.

[43] X. Chen and H. Zhang, "Grey Wolf Optimization-based Deep Echo State Network for Time Series Prediction", *Frontiers in Energy Research*, Vol. 10, pp. 1-18, 2022.

[44] L. Zhang, W. Bian, W. Qu, L. Tuo and Y. Wang, "Time Series Forecast of Sales Volume based on XGBoost", *Journal of Physics: Conference Series*, Vol. 1873, No. 1, pp. 1-7, 2021.

[45] Z. Qinghe, X. Wen, H. Boyan, W. Jong and F. Junlong, "Optimised Extreme Gradient Boosting Model for Short Term Electric Load Demand Forecasting of Regional Grid System", *Scientific Reports*, Vol. 12, No. 1, pp. 1-9, 2022.

[46] S. Stevenson, "A Comparison of the Forecasting Ability of ARIMA Models", *Journal of Property Investment and Finance*, Vol. 25, No. 3, pp. 223-240, 2007.

[47] K.C. Lam and O.S. Oshodi, “Forecasting Construction Output: A Comparison of Artificial Neural Network and Box-Jenkins Model”, *Engineering, Construction and Architectural Management*, Vol. 23, No. 3, pp. 302-322, 2016.

[48] D. Niu, L. Ji, Y. Wang and D. Liu, “Echo State Network with Wavelet in Load Forecasting”, *Kybernetes*, Vol. 41, No. 10, pp. 1557-1570, 2012.

[49] R. Samsudin, A. Shabri and P. Saad, “A Comparison of Time Series Forecasting using Support Vector Machine and Artificial Neural Network Model”, *Journal of Applied Sciences*, Vol. 10, No. 11, pp. 950-958, 2010.