Comparative Analysis of ARIMA, LSTM, and TCN for Time Series Forecasting

Fangxudong Liu

{fliu0404@uni.sydney.edu.au}

The university of Sydney, Faculty of science, Sydney, NSW, 2006, Australia

Abstract. The review summarizes ARIMA under the Box–Jenkins framework, LSTM as a gated recurrent architecture for long-term dependencies, and TCN as a causal, dilated convolutional model with residual learning. Overall, LSTM/TCN generally show advantages under strong nonlinearity and long-dependency settings, while ARIMA can remain competitive for short-horizon and relatively stable series. The findings support hybrid architectures and feature selection as promising directions for improving robustness and practical deployment. In power load prediction, studies have also shown that by optimizing the input lag term and network structure through feature selection and genetic algorithms, the prediction performance of LSTM can be further improved. However, it also reflects that the effect of deep models is highly dependent on structure selection and hyperparameter setting, requiring systematic parameter tuning or search strategies.

Keywords: Time series forecasting; ARIMA model; Temporal Convolutional Network (TCN)

1 Introduction

Time series forecasting is a key technology in intelligent transportation, smart grids, and renewable energy grid integration. Its results directly support decisions such as congestion warning and traffic control, grid dispatch and supply-demand balance, and wind power grid integration and reserve capacity allocation [1-3]. However, these real-world series often exhibit nonlinearity, sudden fluctuations, and long-range dependencies, which challenge the stability and accuracy of traditional linear models in complex scenarios [1, 2].

In traditional time series methods, ARIMA (Autoregressive Integrated Moving Average) is widely used because of its clear structure, strong interpretability, and low training and inference costs. Its classic modeling process usually follows the Box-Jenkins approach: model identification is performed through ACF/PACF, and parameter estimation and diagnostic testing are completed before it is used for prediction [4]. In scenarios with short prediction steps, relatively stable sequences, and more pronounced linear dependencies, ARIMA may still be competitive. For example, short-term wind speed prediction studies show that within short prediction intervals of 10 minutes to 4 hours, ARIMA can achieve prediction results that are no

ICIAAI 2026, May 22-24, Singapore, Singapore

DOI 10.4108/eai.22-5-2026.2365141

worse than or even better than backpropagation neural networks, while having lower computational costs, thus possessing engineering application value [2]. However, ARIMA's main limitations are also prominent: its linear assumptions and dependence on stationarity make it prone to performance degradation when facing strong nonlinearity, sudden fluctuations, or complex dynamic changes, usually requiring differencing, seasonal term expansion, or additional preprocessing to enhance its adaptability [1,4].

In addition to recurrent networks, Temporal Convolutional Network (TCN) has also attracted attention as a convolutional sequence model. TCN usually uses causal convolution to avoid the leakage of future information and expands the receptive field without significantly increasing the number of parameters by dilating convolution, while improving the training stability of deep networks with residual connections [1]. In short-term traffic flow prediction, the TCN framework has shown effective modeling ability for long-short-term dependencies through structural optimization and comparative experiments, and outperforms many deep learning baseline models in error performance, demonstrating the advantages of the convolutional structure in parallel computing and training efficiency [1]. In wind power prediction, TCN has been reported to have a stable training process and strong generalization ability, and outperforms several traditional and deep baseline models in prediction accuracy [2]. Furthermore, hybrid architectures have begun to emerge. For example, the combination of TCN and Informer improves multi-step prediction performance by "extracting hidden temporal features first and then making predictions", indicating that combining feature extraction with long sequence modeling may be an important direction for improving prediction performance [6].

Although the aforementioned studies have made progress in traffic, power, and wind power scenarios, different models exhibit significant trade-offs in interpretability, computational cost, long-dependency modeling capabilities, training stability, and robustness. Furthermore, these trade-offs vary across different tasks and data characteristics. Therefore, it is necessary to systematically compare ARIMA, LSTM, and TCN in multiple representative application scenarios within a unified analytical framework.

Based on this, this paper compares and reviews the theoretical mechanisms and empirical conclusions of ARIMA, LSTM, and TCN in three typical tasks: traffic flow prediction, power load prediction, and wind power prediction. It summarizes the advantages and limitations of different models in different scenarios and concludes the discussion section with key challenges (such as noise robustness, feature selection, model optimization, and engineering deployment requirements), proposing feasible future research directions

2 Theoretical background

2.1 ARIMA

ARIMA is a classic linear time series model, usually written as ARIMA(p, d, q), where p represents the order of autoregression (AR), d represents the difference order (used to transform non-stationary series into approximately stationary ones), and q represents the order of moving average (MA). Its modeling process generally follows the Box-Jenkins approach: first, model identification is performed using tools such as autocorrelation (ACF) and partial autocorrelation (PACF), then parameter estimation and diagnostic testing are performed, and finally, it is used for prediction. [4] The advantages of ARIMA are: interpretable model structure, low training and inference costs, and often stable performance in scenarios with obvious short-term linear dependence and controllable noise. For example, in a study on short-term wind speed forecasting, the authors used measured wind speed sequences to build an ARIMA model and compared it with backpropagation neural networks (BPN). The report pointed out that ARIMA outperformed NNT in short forecasting steps of 10 minutes to 4 hours, and the results of the two were similar on data with "flat terrain and relatively stable sequences". From an engineering perspective, a simpler forecasting model can be adopted to facilitate energy management [2]. Its limitations are also clear: ARIMA assumes that the relationship is mainly characterized by linear terms and is sensitive to non-stationarity, strong seasonality, or structural abrupt changes. It usually requires differencing, seasonal term expansion (such as SARIMA), or additional preprocessing [4].

2 LSTM (Recurrent Neural Network)

LSTM is a gated recurrent network proposed to alleviate the gradient vanishing/exploding problem in long sequence training of traditional RNNs. Its core is "memory unit (cell state) + three gate structures (input gate, forget gate, output gate)", which determines the retention, update and output of information through the gating mechanism, thus making it better at capturing long-distance dependencies [7]. In energy series such as load or wind power, LSTM is often used to fit nonlinear and multi-scale fluctuations. Taking short-term power load forecasting as an example, some studies have searched for the key hyperparameters and structure configuration of LSTM in an automated/optimized way, to obtain prediction results with lower error on a given dataset; this work also reflects a commonality of deep models in this type of task: performance is highly related to "structure selection/hyperparameter setting", and requires systematic parameter tuning or search strategies [8]. It should be noted that although LSTM has strong expressive power, its training cost is relatively high, and it is sensitive to data scale, feature engineering and regularization strategies; if the training data is insufficient or the distribution drift is obvious, overfitting or generalization decline is likely to occur [8].

2.3 Temporal Convolutional Network (TCN)

TCN is a type of one-dimensional convolutional structure for sequence modeling. Typical TCNs ensure that future information is not leaked in time through "causal convolution" and expand the receptive field without significantly increasing the number of parameters through "dilated convolution", thereby covering longer historical dependencies; at the same time, they are often combined with structures such as residual connections and layer normalization to improve training stability and deep expression ability [1]. Compared with recurrent networks, the convolutional computation of TCN is easier to parallelize, and the training process is usually more stable. In the experiment of wind power prediction, the authors reported that TCN converges quickly, the training process is stable, and the loss of the training set and the test set are relatively close, showing strong generalization ability [2]. In general, TCN is often regarded as a sequence model option that "preserves the ability to retain long dependencies" while taking into account "training efficiency and stability", but its effect still depends on factors such as the number of convolutional layers, dilation rate design, input features and prediction stride [1].

3. Case studies

3.1 Traffic flow time series prediction

Traditional traffic flow prediction often employs statistical time series methods such as ARIMA. However, real traffic flow is affected by external disturbances such as peak congestion, weather, accidents, and holidays, exhibiting obvious nonlinearity, sudden fluctuations, and long-term dependence characteristics. In such complex scenarios, statistical models relying on linear structures and stationarity assumptions often struggle to stably depict the dynamic changes in traffic conditions, and prediction errors are more easily amplified during periods of high volatility. Therefore, recent research has gradually shifted towards deep learning models to improve short-term prediction accuracy and robustness [1, 3].

In the direction of convolutional time series modeling, some studies have proposed a short-term traffic flow prediction framework based on TCN. This framework uses causal convolution to ensure that the time sequence does not leak future information and utilizes dilated convolution to expand the receptive field, thereby more effectively capturing long-term and short-term dependencies. This study also uses the Taguchi method to systematically optimize the network structure and compares it with various deep learning baseline models. Experimental results show that the optimized TCN is significantly better than the control model in terms of error index, demonstrating the advantages of convolutional structures in long-dependency modeling and parallel computing [1]. In the direction of recurrent neural network modeling, another study constructed an LSTM prediction model based on the actual traffic flow time series of OpenITS and compared it with ARIMA and backpropagation neural networks (BPNN). The results show that LSTM has a lower overall prediction error and can maintain a relatively stable fitting effect during periods of stronger fluctuation, such as morning and evening peak hours; in contrast, ARIMA's error is more likely to increase during peak hours, reflecting its insufficient adaptability to nonlinear fluctuations and sudden changes [3]. Combining these two types of studies, it can be seen that deep learning methods are better able to cope with nonlinear and strong disturbance characteristics in traffic flow prediction than traditional statistical models: TCN has a significant advantage in overall error performance through convolutional long-dependency modeling, while LSTM shows stronger stability and robustness in peak fluctuations and complex disturbance scenarios. These conclusions provide more reliable technical support for real-time traffic prediction and traffic management decisions in intelligent transportation systems [1,3].

3.2 Power load time series forecasting

Traditional power load forecasting mainly relies on statistical time series models, such as ARIMA and ARMA. However, due to the obvious nonlinear and random characteristics of power load data, the prediction accuracy of traditional models is limited under complex load scenarios.

In terms of statistical model research, the ARIMA noise robustness study constructed a seasonal prediction model based on hourly load data of the Polish power system and evaluated the model stability by superimposing different noise signal ratios (NSR). The experimental design adopted multiple noise scenarios, gradually increasing the noise intensity from 10% to 200%, and evaluated the model performance changes through RMS and AIC indices. The results showed that when the NSR was below 30%, the model prediction remained stable, while the prediction error increased significantly under high noise conditions, and even the model irreversibility problem occurred when the NSR reached 200% [9].

In terms of deep learning hybrid model research, a short-term power load forecasting framework integrating feature selection, deep learning, and heuristic optimization was proposed. The study first uses the Modified Mutual Information method to screen high-dimensional load features to reduce the impact of the dimensionality curse; then, it constructs a Factored Conditional Restricted Boltzmann Machine deep learning model for load prediction and optimizes the model parameters using the Genetic Wind Driven Optimization algorithm. Experimental data are derived from hourly load data of the PJM power market's FE, Dayton, and EKPC grids. The training set covers 2014–2016, the test set is 2017, and MAPE is used as the main evaluation index [2].

Experimental results show that the proposed hybrid model achieves optimal prediction performance on all three grid datasets. Specifically, the MAPE for the FE grid is approximately 2.12%, for the Dayton grid approximately 1.4%, and for the EKPC grid approximately 1.6%. The overall error is significantly lower than that of benchmark models such as LSTM, MI-ANN, and AFC-ANN, while also showing better convergence speed and generalization ability [2]. The combined results of these two studies show that traditional statistical models still have stable predictive capabilities in low-to-medium noise load sequences, while deep learning hybrid models exhibit higher accuracy and stronger adaptability in high-dimensional nonlinear load scenarios, providing a more promising technical path for short-term load forecasting of smart grids.

3.3 Wind power time series prediction

Wind power/wind speed prediction is crucial for grid dispatch and reserve capacity configuration. However, wind power series are driven by meteorological conditions and exhibit strong randomness and nonlinearity, making traditional methods prone to accuracy degradation in multi-step prediction and drastic fluctuation scenarios. Therefore, related research usually evolves from "statistical baseline → deep learning → hybrid architecture". Early comparative studies used ARIMA as the statistical baseline and compared it with a backpropagation neural network (BPNN): based on 10-minute sampled wind speed data (18,090 observations) from the Andalusia region of Spain, prediction step sizes of 10 minutes, 1h, 2h, and 4h were set. The results showed that ARIMA was more stable in terms of correlation and error indicators in the short time interval and had a lower computational cost [2]. With the introduction of deep learning, research has begun to emphasize stronger temporal feature extraction: for example, TCN (causal/dilated convolution and residual connection) is used for short-term wind power prediction, and compared with SVM, MLP, GRU, LSTM, etc., the reported MAE of TCN on the test set is 2.16-2.85 MW and RMSE is 0.023 MW, which is significantly better than the control model [2]. Another approach is to strengthen the structure inside the recurrent network and combine it with feature selection: LSTM-EFG reduces the error in multiple wind farms by strengthening the forget gate and using "sequence-related features + clustering" to select wind turbine features that are highly related to the target unit. Among them, the lowest MSE of Palm Springs wind farm reaches 5.5311 (spectral clustering scheme), and the reported accuracy improvement relative to SVR, KNN, and traditional LSTM can reach 18.30%, 16.84% and 13.10%, respectively [10]. The latest research further integrates convolutional feature extraction and long sequence modeling into an end-to-end hybrid architecture: TCN-Informer first extracts hidden temporal features using TCN, then uses Informer to process long dependency prediction, and performs 1-step to 4-step multi-step prediction on Kaggle wind power data (18 features, 10-minute intervals, 2019/05–2020/03), comparing it with baselines such as Informer, CNN, GRU, LSTM, XGBoost, and LightGBM; in its 1-step results, TCN-Informer achieves MSE=0.0518, MAE=0.1467, RMSE=0.2277, and SMAPE=0.3535, and still maintains relative best performance in 4-step [6]. The study also compared optimizers, reporting that AdaBelief had the lowest average error (e.g., MSE mean=0.0573, MAE mean=0.1524, RMSE mean=0.2394, SMAPE mean=0.3560), indicating that training strategies also further affect prediction performance [6]. In summary, the statistical model (ARIMA) can still serve as a strong baseline in ultra-short-term and resource-constrained deployments; deep learning significantly reduces error through stronger nonlinear modeling and feature extraction (TCN) or gating structure enhancement and relevance feature selection (LSTM-EFG); while the hybrid architecture (TCN-Informer) is more advantageous in multi-step prediction, but it is usually accompanied by higher model complexity and computational overhead, requiring a trade-off between accuracy and engineering cost [5, 6, 10, 11].

3.4 Comparative analysis of different models

Based on the above examples, the differences between ARIMA, LSTM, and TCN can be summarized into three main lines: interpretability and computational cost, modeling ability for nonlinearity and long-term dependencies, and training stability and deployment friendliness.

(1) Interpretability and Cost: ARIMA has a clear structure and fast training, making it suitable for resource-constrained online forecasting. In short-term wind speed forecasting, the authors also emphasized that ARIMA and neural networks have similar results but lower computation time. However, its linear assumption limits its ability to characterize complex nonlinear fluctuations [1,10].

(2) Nonlinearity and Long-Term Dependencies: LSTM is better at handling long-term dependencies and nonlinear sequences through its gated memory mechanism. In tasks such as load forecasting, related studies have found better LSTM configurations through systematic search, indicating that deep models have a higher performance ceiling, but they also rely more on parameter tuning strategies [7, 8].

(3) Training Stability and Parallelism: TCN can be trained in parallel using convolutional structures and expands its receptive field by expanding causal convolutions. In the wind power prediction comparison experiment, TCN showed a lower error index and more stable training, demonstrating its advantage in short-term prediction scenarios [2, 4]. When the sequence is closer to linear and rapid deployment is required, ARIMA can be discussed first; when nonlinearity, long dependence is obvious, and there is sufficient data, LSTM can be discussed; when it is desirable to improve training efficiency and stability while covering a long historical window, TCN can be discussed. All the above conclusions should be bound to the experimental settings and conclusions given in each paper to avoid concluding directly across datasets [5, 11].

4 Discussion

In time series forecasting tasks, deep learning models demonstrate stronger nonlinear modeling and feature extraction capabilities compared to traditional statistical methods. Studies of three cases—traffic flow, power load, and wind power—reveal that deep learning models typically exhibit higher prediction accuracy and stronger generalization performance in complex time series scenarios.

In traffic flow forecasting, time series models based on deep structures can effectively capture the dynamic evolution characteristics of traffic flow. For example, TCN, through dilated convolution and causal convolution structures, can better model long-term and short-term dependencies, thereby improving short-term traffic flow forecasting performance [1]. Meanwhile, LSTM-based traffic flow forecasting models also demonstrate better prediction accuracy than ARIMA in experimental comparisons, indicating that deep models have a greater advantage in handling long sequences and volatility [3]. In contrast, traditional ARIMA relies on linearity and stationarity assumptions, which often limit its application in scenarios with strong random fluctuations like traffic flow [1].

In power load forecasting, power load is affected by multiple factors such as weather and social activities, exhibiting obvious nonlinear and random characteristics. Traditional time series models struggle to stably characterize this complex mapping relationship [2]. LSTM captures long-term dependencies through a gating mechanism, achieving lower errors in load forecasting. Furthermore, by optimizing the input lag terms and network structure through feature selection and genetic algorithms, the forecasting performance can be further improved [8]. In wind power forecasting, wind power exhibits strong fluctuations and intermittency, making deep learning models more advantageous. Improved LSTM, by strengthening the forget gate and combining it with correlation feature extraction, can effectively improve the accuracy and convergence speed of wind power forecasting [10]. TCN alleviates the long-term dependency problem through residual connections and dilated causal convolutions, demonstrating a more stable training process and stronger generalization ability in short-term wind power forecasting [2]. In addition, the hybrid framework of TCN and Informer further improves wind power forecasting performance by "extracting time-series features first and then performing long-series forecasting" [6]. Overall, from the three application cases, it can be seen that ARIMA may still be competitive when the data is relatively stable, and the forecasting step is short, but in complex nonlinear scenarios, deep learning models such as LSTM and TCN have a greater advantage [2, 8, 11].

4.1 Challenges and limitations

Although deep learning performs well in time series forecasting, it still faces challenges in data, models, and actual deployment. These factors may affect the stability and generalizability of model predictions.

Data issues. Time series forecasting is highly dependent on the quality of historical data, while real-world data often contains noise, missing data, and distribution drift. In power load forecasting, noise significantly affects the model's identification and prediction capabilities. When the noise level increases, ARIMA modeling and parameter estimation may become significantly unstable, leading to a decrease in prediction performance [9]. Therefore, the quality of data preprocessing and noise reduction is crucial for prediction stability [9].

Meanwhile, power load data is usually driven by multiple factors. Increased data dimensionality increases the difficulty of feature selection and modeling, requiring a more systematic data processing and feature extraction process [2]. Wind power forecasting is also strongly influenced by meteorological factors, and the correlation and selection method of input features directly affect model performance [10].

Model issues. Deep learning models have complex structures, high computational overhead, and high model parameter tuning costs. While LSTM is suitable for modeling long dependencies, it is time-consuming to train, has many parameters, and the key gating components in its structure have a significant impact on performance, making the model sensitive to structure selection and hyperparameter settings [7]. To improve performance, research often requires combining feature selection and genetic algorithms to find suitable input lag and network configuration, which also increases complexity [8]. Although TCN is relatively stable to train, network design (convolution kernel, dilation rate, number of residual layers) also affects the effect, and increasing the model depth will also bring resource consumption and deployment pressure [2,4]. In contrast, ARIMA has the advantages of a simple structure and strong interpretability, but its linear assumption and dependence on stationarity make it difficult to handle strongly nonlinear time series such as traffic and wind power [1]. Theoretically, the modeling and recognition of ARIMA usually rely on ACF/PACF and differencing steps, which makes it more difficult to recognize in complex fluctuating data [4].

4.2 Future directions

Hybrid/ensemble model design. In the future, more "two-stage or hybrid models" can be adopted, combining the advantages of different models: for example, one model is responsible for feature extraction, and another model is responsible for long-sequence prediction.

In wind power prediction, the TCN+Informer framework first uses TCN to extract hidden temporal features, then uses Informer to complete long-sequence prediction, and uses an optimizer to improve training performance. This hybrid structure of "feature extraction + predictor" can significantly improve prediction accuracy [6].

The same idea can be applied to traffic and power prediction: for example, in traffic flow prediction, TCN is first used to capture long-term and short-term dependencies, and then a prediction module is added to perform multi-step prediction, which is consistent with the motivation of the traffic TCN framework [1].

Multi-factor feature fusion and feature selection. In the future, "external influencing factors" should be incorporated into the model input more systematically, and feature selection should be used to reduce redundancy, allowing the model to learn more stable patterns. In power load forecasting, studies have pointed out that load is affected by factors such as temperature, humidity, calendar, and social behavior. It is proposed to perform data preprocessing and feature selection before training a deep model to address load nonlinearity and randomness [2].

Furthermore, LSTM studies on load forecasting have explicitly used feature selection methods to screen effective features and used genetic algorithms to search for optimal time lag and network structure, demonstrating that "feature fusion + feature selection + lag optimization" is a feasible route to improve performance [8].

In wind power scenarios, LSTM-EFG improves prediction accuracy by using correlation features and clustering to screen more useful wind power features, which is also a similar approach [10].

Based on this paper's theoretical review of ARIMA, LSTM, and TCN, and comparative evidence in three representative scenarios—traffic flow, power load, and wind power—a core conclusion can be drawn: model selection is not about "which is always the best," but rather depends on the trade-off between data features, prediction step size, and engineering constraints. Overall, ARIMA has advantages in terms of clear structure, low inference cost, and strong interpretability, making it suitable as a strong baseline for resource-constrained or relatively linear sequences; however, its linearity and stationarity assumptions also limit its ability to characterize complex nonlinear fluctuations and sudden changes [1, 2]. In contrast, LSTM, relying on gated memory mechanisms, is better at handling long-range dependencies and nonlinear sequences, but its performance is highly dependent on structure and hyperparameters, and its training and parameter tuning costs are higher [7, 8]. TCN, with its advantages of causal/dilated convolution and parallel computing, is more attractive in terms of training stability and efficiency, and can cover a longer historical window [2, 4]. Based on the above comparison, this paper forms more actionable selection recommendations: when the task emphasizes rapid deployment and the sequence is approximately linear, ARIMA can be given priority; when nonlinearity and long dependencies are obvious and data is sufficient, LSTM is more suitable; when it is desirable to improve training efficiency and stability while ensuring long dependency modeling, TCN is often more advantageous; at the same time, all conclusions should be tied to specific experimental settings to avoid concluding directly across datasets [3, 11]. Looking to the future, the focus of time series forecasting research will gradually shift from "single model performance improvement" to system optimization of "robustness and deployability". First, hybrid/two-stage architectures deserve to be explored: for example, the idea of using TCN as a feature extractor and then connecting it to a long sequence predictor can provide a more stable structural path for multi-step forecasting [6]. Second, external influencing factors should be introduced more systematically and feature selection should be performed to reduce redundancy and improve generalization ability; at the same time, automatic parameter tuning or structural optimization strategies should be combined to reduce the dependence of deep models on human experience [5, 8]. Third, noise, missing data, and outliers are inevitable in real systems. Subsequent work needs to pay more attention to data cleaning, robust training, and standardized evaluation processes to ensure that the model still works stably under distribution changes; power load noise research also shows that model performance changes significantly with noise intensity, thus emphasizing the importance of robustness evaluation [9]. Finally, at the deployment level, a balance needs to be struck between accuracy and computational overhead to ensure that the model can run reliably in real-time scenarios such as transportation and power grids [3, 5].

5 Conclusion

Overall, the comparison across traffic flow, power load, and wind power tasks shows that the applicability of ARIMA, LSTM, and TCN depends on the linearity of the sequence, the intensity of fluctuations, the length of dependencies, the prediction step size, and deployment constraints. ARIMA has a clear structure and low cost, making it suitable as an interpretable baseline for short-term and relatively stable scenarios. When the sequence has significant nonlinearity, sudden fluctuations, or long-range dependencies, LSTM and TCN are better able to capture complex dynamics. LSTM focuses on long-term dependency modeling, while TCN has advantages in parallel computing and training stability.

Looking to the future, research focus can shift from simply pursuing lower errors to optimizing robust and deployable systems: first, exploring two-stage or hybrid architectures to decouple feature extraction from long-sequence prediction to improve multi-step prediction stability; second, strengthening multi-factor feature fusion and feature selection to reduce redundancy and improve cross-dataset generalization ability; and third, establishing more comprehensive preprocessing and evaluation processes to address noise, missing data, and distribution drift in real-world data, achieving a feasible balance between accuracy and computational cost.

References

[1] Zhao, W.: Deep Temporal Convolutional Networks for Short-Term Traffic Flow Forecasting. Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC). pp. 1-6 (2019)

[2] Hafeez, G.: Electric load forecasting based on deep learning and optimized by heuristic algorithm. IEEE Access. 8, pp. 81615-81629 (2020)

[3] Zheng, J., & Huang, Y.: Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning. Journal of Physics: Conference Series. 1544, pp. 012012 (2020)

[4] Zhu, X., Liao, K., & Wang, Y.: Short-term prediction for wind power based on temporal convolutional network. Power System Technology. 44(10), pp. 3811-3818 (2020)

[5] Shumway, R. H., & Stoffer, D. S.: Time Series Analysis and Its Applications: With R Examples. Springer. pp. 1-560 (2017)

[6] Gong, M.: Short-term wind power forecasting model based on temporal convolutional network and Informer. Energy. 283, pp. 129171 (2023)

[7] Greff, K., Srivastava, R. K., Koutn k, J., et al.: LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems. 28(10), pp. 2222-2232 (2017)

[8] Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A.: Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. Energies. 11(7), pp. 1636 (2018)

[9] Chodakowska, E.: ARIMA models in electrical load forecasting and their robustness to noise. Energies. 14(23), pp. 7952 (2021)

[10] Yu, J.: LSTM-EFG for wind power forecasting based on sequential correlation. Proc. IEEE Sustainable Power and Energy Conference (iSPEC). pp. 1-5 (2019)

[11] Palomares-Salas, J. C., et al.: ARIMA vs. Neural Networks for Wind Speed Forecasting. Proc. International Conference on Renewable Energies and Power Quality (ICREPQ). pp. 1-6 (2009)