Improving Short-Term Electrical Load Forecasting with Dilated Convolutional Neural Networks: A Comparative Analysis

Tuan Anh Nguyen [1*], Thanh Ngoc Tran [2]
[1, 2] Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Viet Nam
Email: [1] nguyenanhtuan@iuh.edu.vn, [2] tranthanhngoc@iuh.edu.vn
*Corresponding Author

***Abstract*—Short-term load forecasting (STLF) is vital for grid stability and resource optimization for energy systems. Accurate forecasting helps maintain a stable power supply, reduce costs, and improve decision-making. Traditional convolutional neural networks (CNNs) capture local patterns well but struggle with long-term dependencies under fluctuating conditions. This study introduces an optimized Dilated Convolutional Neural Network (DCNN) to enhance accuracy in short- and long-term load forecasting. The key contribution is a new DCNN framework that expands the receptive field without adding computational complexity, effectively capturing multi-level temporal dependencies. This improves performance, stability, and accuracy in volatile conditions. The methodology applies dilated convolution techniques to a real-world electricity load dataset with 13,440 hourly data points. Preprocessing includes normalization and outlier removal. Hyperparameter tuning optimizes dilation rates, kernel sizes, and learning rates. Results show that the DCNN outperforms traditional models, achieving the lowest Mean Absolute Percentage Error (MAPE) of 0.0096. These results surpass CNN (MAPE: 0.0116), GRU (MAPE: 0.0102), and Long Short-Term Memory (LSTM) (MAPE: 0.0272) models. The DCNN also maintains efficiency and stability with volatile data. In conclusion, optimized dilated convolution techniques significantly enhance load forecasting, offering scalable, robust solutions for modern energy management systems requiring fast, accurate, and reliable predictions.**

***Keywords*—Dilated Convolution; Load Forecasting; Deep Learning; Energy Management; Time Series.**

I. INTRODUCTION

STLF plays a pivotal role in the operation and management of energy systems, ensuring grid stability and optimizing resource allocation. Accurate STLF enables energy providers to maintain a reliable power supply, reduce operational costs, and make informed decisions in energy distribution. The growing integration of renewable energy sources and the increasing complexity of modern power grids have heightened the demand for precise and efficient forecasting models. Researchers and engineers have developed various forecasting methods to meet this demand, from traditional statistical techniques to advanced artificial intelligence-based approaches. Conventional statistical models, such as the dynamic average method [1]-[10], the ARIMA Sliding Average Integrated Regression Model [11]-[16], and the linear regression model [17]-[26], are favored for their simplicity and interpretability. While these models perform well with continuous, consistent time-series data, they struggle with irregular fluctuations and nonlinear relationships, limiting their effectiveness in dynamic energy environments.

Advancements in deep learning have brought significant improvements to load forecasting. Models based on CNNs [27]-[41] and recurrent neural networks (RNNs) [42]-[50], including variants like LSTM [51]-[74] and Gated Recurrent Unit (GRU) [75]-[82], have gained considerable attention. RNNs and LSTMs are adept at processing long data sequences and capturing temporal dependencies, but they are computationally intensive and prone to overfitting without extensive datasets. Conversely, CNNs excel in feature extraction for short-term time-series data but face limitations in capturing multiscale patterns and long-term dependencies due to their restricted receptive fields, affecting forecasting accuracy in highly variable conditions. The Dilated Convolution technique has emerged as a promising solution to address the limitations of traditional CNNs. By expanding the receptive field without increasing computational complexity, dilated convolution enables models to capture broader temporal dependencies in time-series data. This capability is especially advantageous for load forecasting, where maintaining accuracy amid large and unstable fluctuations is crucial. The technique meets the stringent demands of modern energy management systems by offering both efficiency and robustness.

This study introduces an innovative application of Dilated Convolution Techniques in short-term load forecasting, conducting a comparative analysis with traditional CNN, GRU, and LSTM models. The analysis focuses on key performance metrics, including forecasting accuracy, computational efficiency, and model stability. Additionally, the study evaluates the robustness and scalability of the Dilated Convolution approach using a real-world electricity load dataset. The primary contribution of this research is developing an optimized DCNN model that significantly enhances forecasting accuracy, reduces computational costs, and improves model stability under fluctuating load conditions. By addressing gaps in existing forecasting methods, this study highlights the potential of Dilated Convolution Techniques to deliver fast, accurate, and reliable predictions, paving the way for broader applications in dynamic and intelligent energy management systems.

Despite advancements in CNN-based forecasting methods, existing models struggle with capturing long-term dependencies and adapting to dynamic load fluctuations due to limited receptive fields. Although dilated convolutions help address these issues, optimizing their performance for large-scale, real-world datasets remains challenging. This study aims to bridge this gap by presenting an optimized Dilated Convolution Technique that balances computational efficiency with forecasting accuracy, offering a robust solution for dynamic energy environments. However, potential limitations include performance trade-offs with massive datasets and sensitivity to hyperparameter tuning, such as dilation rates and filter sizes. Recognizing these challenges ensures a balanced assessment of the technique's capabilities, enhancing its credibility and applicability across diverse real-world forecasting scenarios.

II. THEORETICAL BASIS

A. Short-Term Load Forecasting

Short-term load forecasting predicts electricity load demand over a short period, usually from a few minutes to a week. This forecast plays an essential role in the power system, helping to adjust supply and demand, optimize operating costs, minimize risks, and ensure energy security. Many factors, such as weather, time of day, day of the week, and socio-economic events, often influence load demand. Therefore, STLF requires a forecasting model capable of capturing complex relationships and rapidly changing time series data. Traditional methods for STLF include linear regression, autoregressive integrated moving average model (ARIMA), and generalized nonlinear regression (GLM) model. However, these methods are limited in predicting time series that are nonlinear and irregular. With the development of deep learning techniques, neural networks and machine learning-based models have shown more powerful predictive capabilities thanks to their ability to extract and learn from complex data automatically.

B. Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a deep learning model designed explicitly for processing image data, with widespread applications in computer vision such as facial recognition [83], image classification [84], and object detection [85]. The structure of CNNs consists of four main components: the convolutional layer, the activation layer, the pooling layer, and the fully connected layer [86].

Convolutional Layer: A filter (or filters) is slid through the input image to create a featured map in a convolutional layer. Each filter is small (3x3 or 5x5) and is applied to the entire input image to create a new featured map. The mathematics of this process can be represented as follows: For the input image I and filter F, the characteristic map is calculated by the convolutional product:

$$S(i,j) = (F * I)(i,j) = \sum_m \sum_n F(m,n) I(i-m, j-n) \tag{1}$$

Where, $\mathrm{S(i,j)}$ is the value at position (i,j) in the feature map, $F(m,n)$ is the filter of size m x n, $(i-m, j-n)$ represents the corresponding region in the input image.

$*$ denotes the convolution operation.

This process enables CNNs to recognize spatial image features like edges, corners, and textures [87].

Activation Layer: After the filter is applied, the values on the characteristic map are passed through a nonlinear trigger function, usually ReLU (Rectified Linear Unit). The ReLU function is defined as:

$$ReLU(x) = max(0, x) \tag{2}$$

Where, x is the input of the activation function, which selects the more excellent value between 0 and x. If x is less than 0, the output will be 0; if x is greater than 0, the output will remain x.

The ReLU function enhances the neural network's ability to learn nonlinear features while reducing the vanishing gradient problem, making the model more efficient during training.

This is beneficial for mitigating the vanishing gradient problem, thereby speeding up the training process [88].

Pooling Layer: Pooling typically uses max pooling or average pooling to reduce the spatial size of featured maps, highlight essential features, and reduce the number of parameters. Maximum compounding is defined as:

$$P(i,j) = max_{k,l \in window}\, I(i+k, j+l) \tag{3}$$

Where, $I(i+k, j+l)$ is a small region in the feature map, the window refers to the pooling region (e.g., 2×2 or 3×3).

Max Pooling ensures that the most essential features in an image are retained while reducing noise. This allows CNNs to detect objects regardless of variations in scale or position [89].

Fully Connected Layer: Data from the fully connected layer is flattened and fed into one (or more) fully connected layers. Each neuron in this layer is connected to all the neurons in the previous layer, each with its weight. The output of this class is:

$$y = W_x + b \tag{4}$$

Where, x is the input from the previous layer. W is the weighted matrix, and b is the bias vector.

The fully connected layer often uses the Softmax function for multi-class classification tasks, particularly in image and object classification [84].

C. Dilated Convolutional Neural Network

The DCNN is a variant of CNN designed to extend the range of recognition without increasing the number of parameters. Dilated convolution is a technique that uses a filter with a distance (dilation) between points in a filter. Precisely, this distance is adjusted by the dilation_rate parameter, allowing the filter to "look further" in the input data without increasing the filter size. With dilated convolution, each filter can reach data points at greater distances, helping to capture long-term patterns in the time series. This feature is particularly useful in load forecasting, where demand may depend on long-cycle patterns. Another advantage of dilated convolution is that it helps maintain the resolution of the input data, as there is no need to resort to pooling classes that lose details. The structure of DCNN usually consists of multiple expansion product layers with increasing 'dilation_rate' values in successive layers, helping the model to capture features at various levels in the time series.

For a traditional convolutional product, the output at the location can be calculated as: $y[i]$

$$y[i] = \sum_{k=1}^{K} x[i+k].w[k] \tag{5}$$

For expansion condensation, we use a dilation factor d to widen the distance between the sampling points in the filter. The product of expansion in position will be calculated using the formula: $y[i]$

$$y[i] = \sum_{k=1}^{K} x[i+d.k].w[k] \tag{6}$$

With, x is the Input, W is the filter or kernel with size K, i is the Position of the output y, d the dilation rate, which is a positive integer. When $d = 1$, the expansion conjugation product becomes a regular convolutional product.

D. Evaluation Criteria For Power Load Forecasting

Accurate evaluation of power load forecasting models is critical in optimizing energy management and ensuring the reliability of modern power grids. Metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and MAPE are commonly used to assess forecasting accuracy. Additionally, a recent study by Schreck et al. [90] highlights that MAPE, MAE, and RMSE are key evaluation metrics widely used in power load forecasting research for Local Energy Markets.

MAE: calculates the average magnitude of forecast errors, expressed in absolute terms. It is simple to compute and unaffected by outliers. It also provides insights into the average deviation in measurable units such as megawatts (MW).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{7}$$

MSE: Emphasizes more significant errors by squaring them, making it practical for identifying and minimizing substantial deviations. However, its unit (MW^2) can reduce interpretability.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

RMSE: The square root of MSE, brings the error metric back to the original unit (MW). It balances sensitivity to significant errors with better interpretability, making it a popular choice for assessing overall forecasting performance.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

MAPE: Normalizes errors as percentages of the actual values, making it helpful in comparing datasets with varying scales. It is particularly effective for evaluating performance across datasets with fluctuating power demands.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

Multiple error metrics, such as MAE, MAPE, MSE, and RMSE, form a comprehensive framework for evaluating the performance of power load forecasting models. Each metric captures unique aspects of model accuracy, contributing to a holistic understanding of forecasting performance. MAE and MAPE are particularly effective for measuring average forecast errors

III. SUGGESTED METHODOLOGY

A. Overview of the Recommendation Method

The proposed method in this study is to use a DCNN to predict short-term loads, taking advantage of the ability to expand the receiving field without increasing the number of parameters. DCNN makes it possible for forecasting models to learn both short-term and long-term patterns in the load data series, improving accuracy compared to traditional methods. The DCNN model will consist of multiple expansion product layers with incremental dilation_rate values to optimize the ability to learn from the data. In addition, the model will be refined, and the forecast efficiency through the MAPE index

B. Structure of the DCNN Model with the Proposed Method

The constructed model used in the paper is structured as follows: Two-Layer Convolution 1D (Conv1D). The first layer uses 32 filters with a kernel size of 2 and an expansion (dilation_rate) ratio of 2, which allows for an extended sensing range, helping to detect samples with longer distances in the data. The second layer has 64 filters, using a size two kernel with an expansion ratio of 4 to learn more complex relationships. Both classes use the ReLU (Rectified Linear Unit) trigger function. After the convolution layers, the flattened layer flattens the output into a one-way vector, preparing for fully connected layers. This is followed by a Dense class with 64 neurons, which uses the ReLU trigger function, which helps the model learn nonlinear features, and finally, a Dense class with one output neuron, which is suitable for the problem of predicting continuous values.

The model is compiled with the Adam optimizer, an efficient and popular deep learning algorithm, and the MSE loss function suitable for regression problems. Metric MAE is used to evaluate performance during training. The model is trained with X_train and y_train data in 100 epochs and uses X_val and y_val data to assess the validation set. With this design, the model effectively leverages sequential features in the data and optimizes predictability through fully connected convolution layers.

C. Algorithmic Flowchart

The flowchart in Fig. 1 outlines implementing and evaluating a forecasting model, explicitly utilizing a Convolutional Neural Network (CNN) with dilated convolution techniques for electrical load forecasting. The process begins with collecting and preprocessing input data, including cleaning, normalization, or transformation into a suitable model training format. The CNN model is then constructed with dilated convolution layers, where the dilation technique allows the model to expand the receptive field of the filters without increasing the kernel size. This enables the model to capture long-range dependencies in time-series data effectively. The data is split into test (X_test, y_test), training (X_train, y_train) and validation (X_val, y_val) datasets.

The training process is carried out using the training dataset, while the model's performance is periodically evaluated on the validation dataset. Dilated convolution layers play a critical role in extracting complex temporal features and long-term patterns that conventional convolution layers may overlook. Once training is completed, the model generates predictions on the test dataset (X_test). The forecasting results are assessed using the MAPE. If the MAPE does not meet the threshold, the process loops back to improve and optimize the model. When the MAPE satisfies the predefined criteria, the final results are produced. The flowchart in Fig. 1 demonstrates a systematic approach to developing an accurate and optimized forecasting model, emphasizing the critical role of dilated convolution in learning intricate data relationships with computational efficiency.

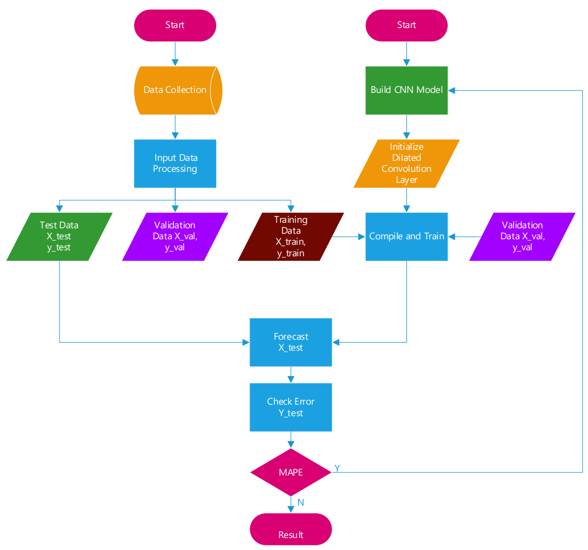

Fig. 1. Algorithm flowchart

IV. RESULT AND DISCUSSION

A. Data

Table I summarizes the Queensland electricity demand dataset, which records half-hourly demand from July 1, 2019, offering valuable insights for STLF. Structured with 4,745 samples, the dataset employs a sliding window approach to generate sequential input-output pairs, where each input X consists of 7 consecutive demand values, and the output Y represents the next demand value. The dataset is split into training, testing, and validation sets, with proportions of 64%, 20%, and 16%, respectively. It has been reshaped to be compatible with deep learning architectures such as LSTM and CNN, ensuring the temporal structure of the data is preserved for capturing dependencies in electricity demand. This processed dataset facilitates accurate STLF and enables practical applications in energy management and grid optimization, particularly for predicting demand fluctuations and ensuring grid stability, serving as a robust foundation for developing intelligent energy systems tailored to Queensland's needs.

TABLE I. HISTORICAL LOAD DATA IN QUEENSLAND

Date	**00:00**	**00:30**	**………….**	**23:00**	**23:30**
1/1/19	5507.31	5362.77	………….	6024.93	5841.3
2/7/19	5620.65	5467.6	………….	6124.78	5932.52
………….	……………..	………….	………….	………….	………….
6/10/19	5380.77	5248.99	………….	5649.78	5856.31
7/10/19	5443.88	5300.68	………….	6119.54	5883.86

B. Model Specifications

Table II summarizes the configurations and parameters of the models, including ARIMA, CNN, Dilated CNN variants, LSTM, and GRU. The table details filters, kernel size, dilation rates, activation functions, dense units, loss functions, optimizers, and epochs used for training, clearly comparing the settings across models.

TABLE II. MODEL PARAMETERS

Model	**ARIMA**	**LSTM**	**GRU**
Parameters	p=1, d=1, q=1		
Loss Function	N/A	MSE	MSE
Optimizer	N/A	Adam	Adam
Epochs	N/A	100	100
Activation		ReLU	ReLU
Dense Units		1	1
Units		50	50

Model	**CNN**	**DCNN1**	**DCNN2**	**DCNN3**
Loss Function	MSE	MSE	MSE	MSE
Optimizer	Adam	Adam	Adam	Adam
Epochs	100	100	100	100
Filters	32	32	32, 64	32, 64, 16
Kernel Size	2	2	2	2
Dilation Rate	1	2		
Activation	ReLU	ReLU	ReLU	ReLU
Dense Units	64	64	64	64
Dilation Rates			2, 4	1, 2, 2
Units	50			

C. Result

Fig. 2 presents the execution time for the seven models: **ARIMA, CNN, LSTM, GRU, DCNN1, DCNN2, and DCNN3**. The runtime comparison highlights significant differences in computational efficiency among these models.

GRU is the slowest model, taking **53.33 seconds** to execute due to its recurrent architecture and long-term memory dependencies. Similarly, **LSTM follows closely behind**, requiring **50.51 seconds**, as it also relies on sequential processing, which increases computational cost.

In contrast, **ARIMA is the fastest model**, completing execution in just **0.38 seconds**, making it the most computationally efficient choice. This result indicates that ARIMA is well-suited for real-time applications with necessary quick predictions.

Among the deep learning models, **DCNN1, CNN, and DCNN2 offer moderate execution times**, with DCNN1 taking **28.15 seconds**, CNN requiring **31.14 seconds**, and DCNN2 completing execution in **32.22 seconds**. These models balance efficiency with predictive power.

Meanwhile, **DCNN3 has a longer runtime of 42.85 seconds**, placing it between CNN-based architectures and recurrent models like LSTM and GRU.

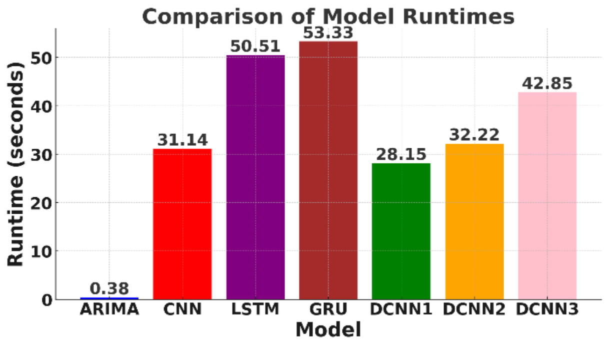

Fig. 2. Runtimes model

Fig. 3 compares the actual load and forecasted values from the ARIMA, CNN, GRU, LSTM, and Dilated CNN models. The graph clearly illustrates each model's performance in tracking load data fluctuations. **ARIMA** (orange line) shows significant limitations, with a flat forecast line failing to reflect the actual data's peaks and troughs. This indicates that ARIMA is unsuitable for complex and non-linear fluctuating data and is better suited for linear time series or simple fluctuating patterns. **CNN** (blue line) improves short-term fluctuation capture compared to ARIMA but still exhibits deviations at the peaks and troughs due to its limited ability to remember long-term relationships in the time series. **GRU** (green line) and **LSTM** (purple line) demonstrate superior performance in load forecasting. GRU closely follows the peaks and troughs, thanks to its ability to learn short-term and long-term patterns. LSTM also captures complex patterns but may experience slight delays at specific peaks and troughs. The **Dilated CNN versions** (cyan, pink, and red lines) exhibit the best performance, accurately reflecting both short-term and long-term details in the data. The expanded receptive field of the network enables these models to maintain high accuracy without increasing the number of parameters. However, the dilation rate must be carefully adjusted to avoid missing short-term details.

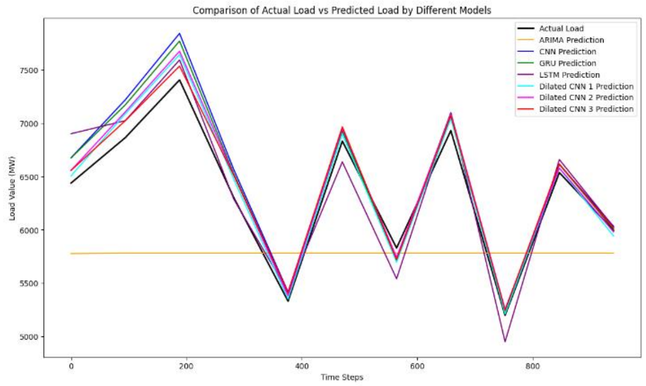

Fig. 3. The load forecasting chart of different models

Fig. 4 presents the MAPE values of various forecasting models, including ARIMA, CNN, GRU, LSTM, and Dilated CNN versions, providing an overview of the performance of each model. MAPE is an index that measures the accuracy of forecasts, with lower MAPE values indicating higher accuracy. The analysis results reveal significant differences in performance among the models. The **ARIMA** model has the highest MAPE value, approximately 0.11, reflecting poor forecasting performance when applied to load data with complex and unstable fluctuations. It is the least efficient model among those tested, as ARIMA is better suited for linear or stable time series, while it struggles with strongly fluctuating load data. **CNN (Convolutional Neural Network)** shows better forecasting performance with a MAPE value of around 0.0116, significantly lower than ARIMA. This model performs more effectively when forecasting short-term cycles and local patterns in the data, thanks to the capabilities of the pooling layers. However, CNN is still not optimal compared to deep regression models like GRU or LSTM, better at capturing long-term relationships. **GRU** achieves a very low MAPE value, approximately 0.0102, demonstrating better forecasting performance than CNN and almost equivalent to Dilated CNN models. With its ability to retain long-term relationships, GRU performs well on highly volatile and complex load data, making it a suitable and practical choice for load forecasting applications. Meanwhile, **LSTM** has a higher MAPE value than GRU, approximately 0.0272. Although LSTM is renowned for memorizing long-term patterns, this model tends to be more complex and requires significant computational resources. The performance of LSTM in this case is inferior to GRU, potentially due to overfitting or unnecessary complexity for this dataset. Finally, **Dilated CNN versions** (Dilated 1, 2, and 3 CNN) show the best forecasting performance with very low MAPE values. Specifically, Dilated 2 CNN achieves the lowest MAPE value of 0.0096, followed by Dilated 3 CNN at 0.0101 and Dilated 1 CNN at 0.0108. The Dilated CNN versions effectively forecast short-term load data by expanding the receptive field without increasing the number of parameters. This allows the models to capture local and long-term patterns in the time series, providing the highest accuracy among the compared models.

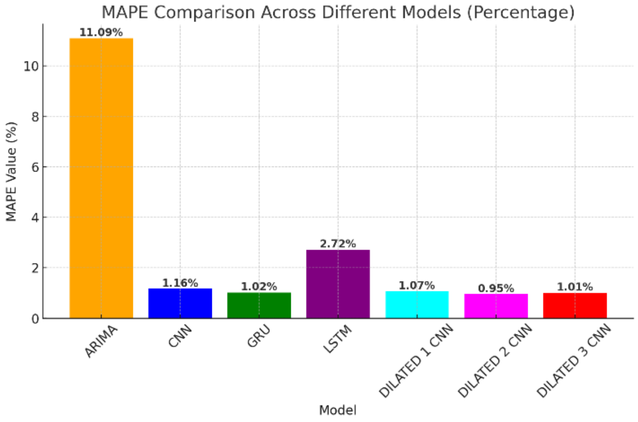

Fig. 4. The MAPE values of the models

D. Discussion

- **Computational Efficiency and Sensitivity to Hyperparameters**

Execution time is crucial in selecting the appropriate model for real-world applications. ARIMA has the fastest execution time (0.38 seconds) due to its simple linear structure with three parameters (p, d, q). However, this simplicity limits its handling of nonlinear and highly dynamic data.

On the other hand, LSTM and GRU have the most extended execution times (50.51 seconds and 53.33 seconds, respectively) due to their sequential nature and memory retention mechanisms. These models consume significant computational resources and are highly sensitive to hyperparameters such as the number of hidden units, number of epochs, and optimizer settings.

Meanwhile, CNN and its variants (DILATED CNN) demonstrate moderate execution times (28.15 - 42.85 seconds), depending on the number of convolutional layers and dilation rate. These models are generally less sensitive to hyperparameters than LSTM and GRU, making them more stable during training.

- **Forecasting Accuracy Based on MAPE**

The MAPE is used to evaluate model accuracy, where lower MAPE values indicate higher accuracy.

The results show that ARIMA has the highest MAPE, suggesting that while it operates efficiently, its forecasting accuracy is poor when dealing with nonlinear or highly volatile data. Therefore, ARIMA is not an optimal choice when high-precision forecasting is required.

Among deep learning models, LSTM exhibits a higher MAPE than CNN and DILATED CNN, indicating that it may not be fully optimized for this forecasting task. Although LSTM performs well for long-term sequential data, it may struggle with short-term patterns and fluctuating data.

In contrast, CNN and its DILATED CNN variants achieve the lowest MAPE values, demonstrating superior feature extraction capabilities. Specifically, DILATED CNN achieves the highest forecasting accuracy, leveraging dilated convolutions to capture long-range dependencies in the data without losing local information.

- **Main Findings of the Present Study**

The study highlights that DILATED CNN outperforms traditional models like ARIMA, CNN, LSTM, and GRU regarding forecasting accuracy. DILATED CNN's ability to capture both short-term and long-term dependencies through expanded receptive fields has proven effective in managing volatile energy load data.

- **Comparison with Other Studies**

Our findings align with studies indicating that deep learning models, especially DILATED CNN, provide superior forecasting accuracy compared to previous research. Prior works [91] demonstrated similar trends where models incorporating dilated convolutions outperformed standard CNNs and RNN-based architectures in handling complex time-series data.

In recent years, several studies have further supported these findings. For instance, a 2022 study on time-series analysis with smoothed CNNs found that CNNs could increase accuracy by up to 30% and train models twice as fast as other algorithms such as Recurrent Neural Networks (RNNs), GRUs, and LSTMs [92]. A 2023 review on deep learning models for time series forecasting highlighted that models incorporating dilated convolutions, such as Temporal Convolutional Networks (TCNs), have improved performance over standard RNN-based architectures in handling complex time-series data [93].

Furthermore, a 2024 study on leveraging hybrid deep learning models for enhanced multivariate time series forecasting proposed hybrid CNN-RNN and TCN-RNN models, significantly outperforming both baseline and state-of-the-art models [94]. Moreover, a 2022 study introduced a hybrid Temporal Convolutional Network and Prophet model for time series forecasting, demonstrating that the proposed method was faster with similar forecasting accuracy compared to LSTM and RNN models [95].

These findings align with our results, reinforcing the conclusion that dilated CNNs offer superior forecasting accuracy and efficiency in handling complex time-series data.

- **Implications and Explanation of Findings**

The improved performance of DILATED CNN is attributed to its ability to maintain computational efficiency while expanding the receptive field, allowing for better capture of temporal dependencies. This makes it highly suitable for applications requiring fast, accurate load forecasting, particularly in dynamic energy environments with fluctuating demands.

- **Strengths**

This study presents several notable strengths. First, it comprehensively compares multiple forecasting models, including ARIMA, CNN, LSTM, GRU, and DILATED CNN, offering valuable insights into their performance across various metrics. The analysis highlights the differences in computational efficiency and forecasting accuracy and examines how each model handles dynamic and non-linear data. Additionally, the study integrates robust statistical validation techniques, such as ANOVA and Tukey's HSD tests, to ensure the reliability and significance of the results. This methodological rigor enhances the credibility of the findings and provides a solid foundation for future research.

- **Practical Applications**

Load forecasting models are valuable in theoretical research and are crucial in various practical applications, particularly in the energy sector. In power system management, short-term load forecasting helps optimize resource allocation and ensure grid stability, with DILATED CNN effectively handling volatile data and ARIMA suited for stable systems with linear trends. LSTM and GRU process long-term sequential data for renewable energy integration to capture generation cycles influenced by weather conditions. In industrial load management, CNN and DILATED CNN identify recurring consumption patterns and optimize energy use in complex production environments. ARIMA analyzes short-term price trends in energy market forecasting and pricing, while DILATED CNN handles complex, volatile market data for more accurate predictions. For smart cities, CNN and DILATED CNN analyze large IoT datasets to optimize public services like lighting and air conditioning in real-time. Lastly, LSTM and GRU accurately forecast load demand in microgrid management, enhancing energy storage efficiency and renewable energy distribution. These models' flexibility and accuracy make them essential for modern energy management systems.

- **Future Work and Recommendations**

Future work should incorporate these statistical tests to strengthen the reliability of model performance comparisons. Based on these findings, the following recommendations can be made: If speed is prioritized over accuracy, ARIMA is a suitable choice, especially for simple forecasting tasks or real-time applications where computational efficiency is more critical than precision. LSTM and GRU can offer better performance for modeling long-term sequential data but require more computational resources and careful hyperparameter tuning to minimize forecasting errors. CNN and DILATED CNN provide the best accuracy, with DILATED CNN being the most effective model due to its enhanced feature extraction through dilated convolutions. This makes DILATED CNN an ideal choice for complex forecasting tasks, mainly when dealing with highly volatile data with hidden patterns.

Moreover, future research should focus on:

- Applying these models to more extensive, real-world datasets with diverse features to evaluate their scalability and robustness.
- Integrating additional external factors like weather data and economic indicators to improve forecasting accuracy.
- Exploring hybrid models that combine the strengths of GRUs and DILATED CNNs to enhance accuracy and computational efficiency.
- Investigating explainability techniques like SHAP (Shapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) to improve model interpretability.

By addressing these areas, future studies can contribute to more robust, interpretable, and efficient forecasting models suitable for various applications.

V. Conclusion

Based on the model performance analysis, **GRU and Dilated CNN** are the most effective approaches for short-term load forecasting, capturing nonlinear fluctuations with high accuracy and efficiency. With its simplified architecture, GRU optimizes computational resources while maintaining accuracy, making it ideal for real-world applications. **Dilated CNN** enhances feature extraction and long-term dependency modeling without increasing computational costs.

However, each model has limitations—**ARIMA** struggles with nonlinear and volatile data, **LSTM and GRU** require extensive computational resources, and **CNN** is limited in capturing long-term dependencies. Selecting the appropriate model necessitates balancing accuracy, efficiency, and data complexity.

This study highlights **Dilated CNN's superior forecasting accuracy** and **GRU's efficiency**, contributing to the theoretical foundation of load forecasting. However, sensitivity to data quality and hyperparameters remains a challenge. Future research should explore **hybrid GRU-Dilated CNN models**, apply them to **diverse datasets**, and improve **model interpretability** using techniques like **SHAP**, ensuring robust and scalable forecasting solutions for real-world energy systems.

References

[1] U. I. Lezama Lope, A. Benavides-Vázquez, G. Santamaría-Bonfil, and R. Z. Ríos-Mercado, "Fast and Efficient Very Short-Term Load Forecasting Using Analogue and Moving Average Tools," *IEEE Lat. Am. Trans.*, vol. 21 no. 9 pp. 1015–1021, 2023, doi: 10.1109/TLA.2023.10251808.

[2] Y. Miao and Z. Chen, "Short-term Load Forecasting Based on Echo State Network and LightGBM," *2023 IEEE Int. Conf. Predict. Control Electr. Drives Power Electron.*, no. 52177087, pp. 1–6, 2023, doi: 10.1109/PRECEDE57319.2023.10174609.

[3] U. Samal and A. Kumar, "Enhancing Software Reliability Forecasting Through a Hybrid ARIMA-ANN Model," *Arab. J. Sci. Eng.*, 2023, doi: 10.1007/s13369-023-08486-1.

[4] C. Wei, D. Pi, M. Ping, and H. Zhang, "Short-term load forecasting using spatial-temporal embedding graph neural network," *Electron. Power Syst. Res.*, vol. 225, p. 109873, 2023, doi: 10.1016/J.EPSR.2023.109873.

[5] M. Abdurohman and A. G. Putrada, "Forecasting Model for Lighting Electricity Load with a Limited Dataset using XGBoost," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 8, no. 2, pp. 571–580, 2023, doi: 10.22219/kinetic.v8I2.1687.

[6] P. D. Lê *et al.*, "Applying statistical analysis for assessing the reliability of input data to improve the quality of short-term load forecasting for a Ho Chi Minh City distribution network," *Science & Technology Development Journal - Engineering and Technology*, vol. 2, no. 4. pp. 223–239, 2020, doi: 10.32508/stdjet.v2i4.614.

[7] S. Chen, R. Lin, and W. Zeng, "Short-Term Load Forecasting Method Based on ARIMA and LSTM," *Int. Conf. Commun. Technol. Proceedings, ICCT*, vol. 2022, pp. 1913–1917, 2022, doi: 10.1109/ICCT56141.2022.10073051.

[8] H. Xu, G. Fan, G. Kuang, and Y. Song, "Construction and Application of Short-Term and Mid-Term Power System Load Forecasting Model Based on Hybrid Deep Learning," *IEEE Access*, vol. 11, pp. 37494–37507, 2023, doi: 10.1109/ACCESS.2023.3266783.

[9] X. Guo, Q. Zhao, D. Zheng, Y. Ning, and Y. Gao, "A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network

considering the real-time electricity price," *Energy Reports*, vol. 6, pp. 1046–1053, 2020, doi: 10.1016/j.egyr.2020.11.078.

[10] I. Arora, J. Gambhir, and T. Kaur, "Data Normalisation-Based Solar Irradiance Forecasting Using Artificial Neural Networks Auto Regressive Moving Average model," *Arab. J. Sci. Eng.*, 2020, doi: 10.1007/s13369-020-05140-y.

[11] A. Ajitha, M. Goel, M. Assudani, S. Radhika, and S. Goel, "Design and development of Residential Sector Load Prediction model during COVID-19 Pandemic using LSTM based RNN," *Electr. Power Syst. Res.*, vol. 212, p. 108635, 2022, doi: 10.1016/j.epsr.2022.108635.

[12] A. Mansouri, A. H. Abolmasoumi, and A. A. Ghadimi, "Weather sensitive short term load forecasting using dynamic mode decomposition with control," *Electr. Power Syst. Res.*, vol. 221, p. 109387, 2023, doi: 10.1016/j.epsr.2023.109387.

[13] S. Aoufi, A. Derhab, M. Guerroumi, H. Guemmouma, and H. Lazali, "LITE-FORT: Lightweight three-stage energy theft detection based on time series forecasting of consumption patterns," *Electr. Power Syst. Res.*, vol. 225, p. 109840, 2023, doi: 10.1016/J.EPSR.2023.109840.

[14] Q. Liu, J. Cao, J. Zhang, Y. Zhong, T. Ba, and Y. Zhang, "Short-Term Power Load Forecasting in FGSM-Bi-LSTM Networks Based on Empirical Wavelet Transform," *IEEE Access*, vol. 11, pp. 105057–105068, 2023, doi: 10.1109/ACCESS.2023.3316516.

[15] E. A. Siqueira-Filho, M. F. A. Lira, A. Converti, H. V. Siqueira, and C. J. A. Bastos-Filho, "Predicting Thermoelectric Power Plants Diesel/Heavy Fuel Oil Engine Fuel Consumption Using Univariate Forecasting and XGBoost Machine Learning Models," *Energies*, vol. 16, no. 7, 2023, doi:10.3390/EN16072942.

[16] H. Kuang, Q. Guo, S. Li, and H. Zhong, "Short-term Power Load Forecasting Method in Rural Areas Based on CNN-LSTM," *Proc. 2021 IEEE 4th Int. Electr. Energy Conf. CIEEC 2021*, pp. 7–11, 2021, doi: 10.1109/CIEEC50170.2021.9510777.

[17] A. K. Kür, "Forecasting Uganda 's Net Electricity Consumption Using a Hybrid PSO-ABC Algorithm," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3021-3031, 2023, doi: 10.1007/s13369-018-3383-z.

[18] Y. Deng, X. Wang, and Y. Liao, "ASA-Net: Adaptive Sparse Attention Network for Robust Electric Load Forecasting," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 4668–4678, 2024, doi: 10.1109/geot.2023.3300695.

[19] X. Chen, M. Yang, Y. Zhang, J. Liu, and S. Yin, "Load Prediction Model of Integrated Energy System Based on CNN-LSTM," *2023 3rd Int. Conf. Energy Eng. Power Syst. EEPS 2023*, pp. 248–251, 2023, doi: 10.1109/EEPS58791.2023.10257124.

[20] S. H. Rafi, N. Al-Masood, S. R. Deeba, and E. Hossain, "A short-term load forecasting method using integrated CNN and LSTM network," *IEEE Access*, vol. 9, pp. 32436–32448, 2021, doi: 10.1109/ACCESS.2021.3060654.

[21] S. Atef and A. B. Eltawil, "Assessment of stacked unidirectional and bidirectional long short-term memory networks for electricity load forecasting," *Electr. Power Syst. Res.*, vol. 187, p. 106489, 2020, doi: 10.1016/j.epsr.2020.106489.

[22] L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: CatBoost and XGBoost approaches," *Expert Syst. Appl.*, vol. 241, 2024, doi: 10.1016/j.Eshwa.2023.122686.

[23] M. Hasan, L. C. Hau, Y. S. Lim, J. Wong, S. Nadiah, and M. Sapihie, "Short-term Load Forecasting for Peak Demand Reduction with Limited Historical Data," *2023 11th Int. Conf. Smart Grid Clean Energy Technol.*, pp. 79–85, 2023, doi: 10.1109/ICSGCE59477.2023.10420080.

[24] H. Aldosari, "A Hybrid Support Vector Machine with Grasshopper Optimization Algorithm based Feature Selection for Load Forecasting," *2024 Int. Conf. Distrib. Comput. Best. Tech.*, pp. 1–4, 2024, doi: 10.1109/ICDCOT61034.2024.10515470.

[25] Z. Yang, T. Hu, J. Zhu, W. Shang, Y. Guo, and A. Foley, "Hierarchical High-Resolution Load Forecasting for Electric Vehicle Charging: A Deep Learning Approach," *IEEE J. Emerg. Sel. Top. Ind. Electron.*, vol. 4, no. 1, pp. 118–127, 2022, doi: 10.1109/jestie.2022.3218257.

[26] J. W. Xiao, P. Liu, H. Fang, X. K. Liu, and Y. W. Wang, "Short-Term Residential Load Forecasting with Baseline-Refinement Profiles and Bi-Attention Mechanism," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 1052–1062, 2023, doi: 10.1109/TSG.2023.3290598.

[27] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020, doi: 10.1109/ACCESS.2020.3028281.

[28] L. Wensheng, W. Kuihua, F. Liang, L. Hao, W. Yanshuo, and C. Can, "A Region-Level Integrated Energy Load Forecasting Method Based on CNN-LSTM Model with User Energy Label Differentiation," *2020 5th Int. Conf. Power Renew. Energy, ICPRE 2020*, pp. 154–159, 2020, doi: 10.1109/ICPRE51194.2020.9233226.

[29] S. Wu *et al*., "Power Load Forecasting Method Based on Random Matrix Theory and CNN-LSTM Model," *2022 IEEE 2nd Int. Conf. Digit. Twins Parallel Intell. DTPI 2022*, pp. 1–6, 2022, doi: 10.1109/DTPI55838.2022.9998910.

[30] S. Pelekis *et al*., "A comparative assessment of deep learning models for day-ahead load forecasting: Investigating key accuracy drivers," *Sustainable Energy, Grids and Networks*, vol. 36, p. 101171, 2023.

[31] M. F. Alsharekh, S. Habib, D. A. Dewi, W. Albattah, M. Islam, and S. Albahli, "Improving the efficiency of multistep short-term electricity load forecasting via R-CNN with ML-LSTM," *Sensors*, vol. 22, no. 18, p. 6913, 2022.

[32] Y. Yang, E. U. Haq, and Y. Jia, "A Novel Deep Learning Approach for Short and Medium-Term Electrical Load Forecasting Based on Pooling LSTM-CNN Model," *2020 IEEE/IAS Ind. Commer. Power Syst. Asia, I CPS Asia 2020*, pp. 26–34, 2020, doi: 10.1109/ICPSAsia48933.2020.9208557.

[33] W. Xiong, L. Han, and X. Qu, "Bus Load Forecasting Based on Maximum Information Coefficient and CNN-LSTM Model," *2023 IEEE Int. Conf. Image Process. Comput. Appl. ICIPCA 2023*, pp. 659–663, 2023, doi: 10.1109/ICIPCA59209.2023.10257944.

[34] O. Rubasinghe *et al*., "A Novel Sequence to Sequence Data Modelling Based CNN-LSTM Algorithm for Three Years Ahead Monthly Peak Load Forecasting," *IEEE Trans. Power Syst.*, vol. 39, no. 2, pp. 1932–1947, 2024, doi: 10.1109/TPWRS.2023.3271325.

[35] L. Wenya, "Cooling, Heating and Electric Load Forecasting for Integrated Energy Systems Based on CNN-LSTM," *2021 6th Int. Conf. Power Renew. Energy, ICPRE 2021*, no. 1, pp. 808–812, 2021, doi: 10.1109/ICPRE52634.2021.9635244.

[36] S. M. H. Rizvi, "Time Series Deep learning for Robust Steady-State Load Parameter Estimation using 1D-CNN," *Arab. J. Sci. Eng.*, 2021, doi: 10.1007/s13369-021-05782-6.

[37] S. Bian, Z. Wang, W. Song, and X. Zhou, "Feature extraction and classification of time-varying power load characteristics based on PCANet and CNN+Bi-LSTM algorithms," *Electr. Power Syst. Res.*, vol. 217, 2023, doi: 10.1016/j.epsr.2023.109149.

[38] C. Li, R. Hu, C. Y. Hsu, and Y. Han, "Short-term Power Load Forecasting based on Feature Fusion of Parallel LSTM-CNN," *2022 IEEE 4th Int. Conf. Power, Intell. Comput. Syst. ICPICS 2022*, no. 4, pp. 448–452, 2022, doi: 10.1109/ICPICS55264.2022.9873566.

[39] T. H. Bao Huy, D. N. Vo, K. P. Nguyen, V. Q. Huynh, M. Q. Huynh, and K. H. Truong, "Short-Term Load Forecasting in Power System Using CNN-LSTM Neural Network," *Conf. Proc. - 2023 IEEE Asia Meet. Environ. Electr. Eng. EEE-AM 2023*, pp. 1–6, 2023, doi: 10.1109/EEE-AM58328.2023.10395221.

[40] F. Xu, G. Weng, Q. Ye, and Q. Xia, "Research on Load Forecasting Based on CNN-LSTM Hybrid Deep Learning Model," *2022 IEEE 5th Int. Conf. Electron. Technol. ICET 2022*, pp. 1332–1336, 2022, doi: 10.1109/ICET55676.2022.9824615.

[41] S. Luo, Z. Ni, X. Zhu, P. Xia, and H. Wu, "A Novel Methanol Futures Price Prediction Method Based on Multicycle CNN-GRU and Attention Mechanism," *Arab. J. Sci. Eng.*, vol. 48, no. 2, pp. 1487–1501, 2023, doi:10.1007/s13369-022-06902-6.

[42] S. S. Subbiah and J. Chinnappan, "Deep learning based short term load forecasting with hybrid feature selection," *Electr. Power Syst. Res.*, vol. 210, p. 108065, 2022, doi: 10.1016/j.epsr.2022.108065.

[43] H. Shahinzadeh, H. Sadrarhami, M. M. Hayati, H. Majidi-Gharehnaz, M. Abapour, and G. B. Gharehpetian, "Review and Comparative Analysis of Deep Learning Techniques for Smart Grid Load Forecasting," *2024 20th CSI Int. Symp. Artif. Intell. Signal Process. AISP 2024*, pp. 1–9, 2024, doi: 10.1109/AISP61396.2024.10475303.

[44] S. Singh and M. M. Tripathi, "A Comparative Analysis of Extreme Gradient Boosting Technique with Long Short-Term Memory and

Layered Recurrent Neural Network for Electricity Demand Forecast," *2021 6th Int. Conf. Recent Trends Electron. Information, Commun. Technol. RTEICT 2021*, pp. 297–302, 2021, doi: 10.1109/RTEICT52294.2021.9573988.

[45] O. A. Lawal and J. Teh, "Assessment of dynamic line rating forecasting methods," *Electr. Power Syst. Res.*, vol. 214, p. 108807, 2023, doi: 10.1016/j.epsr.2022.108807.

[46] A. K. Mishra, P. Mishra, and H. D. Mathur, "A deep learning assisted adaptive nonlinear deloading strategy for wind turbine generator integrated with an interconnected power system for enhanced load frequency control," *Electr. Power Syst. Res.*, vol. 214, 2023, doi: 10.1016/j.epsr.2022.108960.

[47] Y. Liu, X. Wang, S. Wang, and Z. Xu, "Short-term Power Load Forecasting Based on Temporal Convolutional Network," *2022 Int. Conf. Information, Control. Common. Technol.*, pp. 1–4, doi: 10.1109/ICCT56057.2022.9976543.

[48] G. Gürses-Tran, T. A. Körner, and A. Monti, "Introducing explainability in sequence-to-sequence learning for short-term load forecasting," *Electr. Power Syst. Res.*, vol. 212, p. 108366, 2022, doi: 10.1016/j.epsr.2022.108366.

[49] S. Luo, Y. Rao, J. Chen, H. Wang, and Z. Wang, "Short-Term Load Forecasting Model of Distribution Transformer Based on CNN and LSTM," *7th IEEE Int. Conf. High Volt. Eng. Appl. ICHVE 2020 - Proc.*, pp. 1–4, 2020, doi: 10.1109/ECV49031.2020.9279813.

[50] A. K. A. Penaloza, A. Balbinot, and R. C. Leborgne, "Review of deep learning application for short-term household load forecasting," *2020 IEEE PES Transm. Distrib. Conf. Exhib. - Lat. Am. T D LA 2020*, pp. 1–6, 2020, doi: 10.1109/TDLA47668.2020.9326148.

[51] Z. Tian, W. Liu, W. Jiang, and C. Wu, "CNNs-Transformer based day-ahead probabilistic load forecasting for weekends with limited data availability," *Energy*, vol. 293, p. 130666, 2024, doi: 10.1016/j.energy.2024.130666.

[52] X. Qu *et al.*, "Personalized Federated Learning for Heterogeneous Residential Load Forecasting," *Big Data Min. Anal.*, vol. 6, no. 4, pp. 421–432, 2023, doi: 10.26599/BDMA.2022.9020043.

[53] R. Liu, T. Chen, G. Sun, S. M. Muyeen, S. Lin, and Y. Mi, "Short-term probabilistic building load forecasting based on feature integrated artificial intelligent approach," *Electron. Power Syst. Res.*, vol. 206, p. 107802, 2022, doi: 10.1016/j.epsr.2022.107802.

[54] M. Zhang, Z. Yu, and Z. Xu, "Short-term load forecasting using recurrent neural networks with input attention mechanism and hidden connection mechanism," *IEEE Access*, vol. 8, pp. 186514–186529, 2020, doi: 10.1109/ACCESS.2020.3029224.

[55] W. Zeng *et al.*, "Hybrid CEEMDAN-DBN-ELM for online DGA serials and transformer status forecasting," *Electr. Power Syst. Res.*, vol. 217, 2023, doi: 10.1016/j.epsr.2023.109176.

[56] S. K. Mishra, V. K. Gupta, R. Kumar, S. K. Swain, and D. K. Mohanta, "Multi-objective optimization of economic emission load dispatch incorporating load forecasting and solar photovoltaic sources for carbon neutrality," *Electr. Power Syst. Res.*, vol. 223, p. 109700, 2023, doi: 10.1016/j.epsr.2023.109700.

[57] Y. Wang, J. Chen, X. Chen, X. Zeng, and S. Member, "Short-Term Load Forecasting for Industrial Customers Based on TCN-LightGBM," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1984–1997, 2021, doi: 10.1109/TPWRS.2020.3028133.

[58] S. F. Stefenon *et al.*, "Time series forecasting using ensemble learning methods for emergency prevention in hydroelectric power plants with dam," *Electr. Power Syst. Res.*, vol. 202, 2022, doi: 10.1016/J.EPSR.2021.107584.

[59] M. Aouad, H. Hajj, K. Shaban, R. A. Jabr, and W. El-Hajj, "A CNN-Sequence-to-Sequence network with attention for residential short-term load forecasting," *Electr. Power Syst. Res.*, vol. 211, p. 108152, 2022, doi: 10.1016/j.epsr.2022.108152.

[60] A. Huang, J. Zhou, T. Cheng, X. He, J. Lv, and M. Ding, "Short-term load forecasting for holidays based on similar days selecting and XGBoost model," *2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems (ICPS)*, pp. 1-6, 2023, doi: 10.1109/ICPS58381.2023.10128055.

[61] A. Banga and R. Ahuja, "Accurate Detection of Electricity Theft Using Classification Algorithms and Internet of Things in Smart Grid," *Arab. J. Sci. Eng.*, vol. 47, no. 8, pp. 9583–9599, 2022, doi: 10.1007/s13369-021-06313-z.

[62] L. Zhang, C. Zheng, Z. Zhao, and D. Zhang, "Short-term Load Forecasting based on Kalman Filter and Nonlinear Autoregressive Neural Network," *Proc. 33rd Chinese Control Decis. Conf. CCDC 2021*, pp. 3747–3751, 2021, doi: 10.1109/CCDC52312.2021.9602793.

[63] H. Min, F. Lin, K. Wu, J. Lu, Z. Hou, and C. Zhan, "Broad learning system based on Savitzky – Golay filter and variational mode decomposition for short-term load forecasting," *2022 IEEE Int. Symp. Prod. Compliance Eng. - Asia*, pp. 1–6, 2022, doi: 10.1109/ISPCE-ASIA57917.2022.9970794.

[64] M. Sharma, N. Mittal, A. Mishra, and A. Gupta, "An efficient approach for load forecasting in agricultural sector using machine learning," *e-Prime - Adv. Electr. Eng. Electron. Energy*, vol. 6, p. 100337, 2023, doi: 10.1016/j.prime.2023.100337.

[65] U. Sencan and G. S. N. Arica, "Forecasting of Day-Ahead Electricity Price Using Long Short-Term Memory-Based Deep Learning Method," *Arab. J. Sci. Eng.*, vol. 47, no. 11, pp. 14025–14036, 2022, doi: 10.1007/s13369-022-06632-9.

[66] S. M. J. Jalali *et al.*, "An advanced deep neuroevolution model for probabilistic load forecasting," *Electr. Power Syst. Res.*, vol. 211, p. 108351, 2022, doi: 10.1016/j.epsr.2022.108351.

[67] Y. Ma, L. Yu, G. Zhang, and Z. Lu, "Design of a multi-energy complementary scheduling scheme with uncertainty analysis of the source-load prediction," *Electr. Power Syst. Res.*, vol. 220, p. 109268, 2023, doi: 10.1016/j.epsr.2023.109268.

[68] J. Shi, W. Zhang, Y. Bao, D. W. Gao, and Z. Wang, "Load Forecasting of Electric Vehicle Charging Stations: Attention Based Spatiotemporal Multi-Graph Convolutional Networks," *IEEE Trans. Smart Grid*, no. 8, p. 1, 2023, doi: 10.1109/TSG.2023.3321116.

[69] J. Zhu, R. Xu, J. Gu, X. Zhang, and C. Sun, "Wind / Storage Power Scheduling Based on Time – Sequence Rolling Optimization," *Arab. J. Sci. Eng.*, vol. 48, no. 5, pp. 6219–6236, 2023, doi: 10.1007/s13369-022-07220-7.

[70] J. C. del-Pino-López, P. Cruz-Romero, and J. C. Bravo-Rodríguez, "Evaluation of the power frequency magnetic field generated by three-core armored cables through 3D finite element simulations," *Electr. Power Syst. Res.*, vol. 213, p. 108701, 2022, doi: 10.1016/j.epsr.2022.108701.

[71] N. Rezaei, R. Rajabi, and A. Estebsari, "Electricity Price Forecasting Model based on Gated Recurrent Units," 2*022 IEEE Int. Conf. Environ. Electr. Eng. 2022 IEEE Ind. Commer. Power Syst. Eur. EEEIC / I CPS Eur. 2022*, pp. 1–5, 2022, doi: 10.1109/EEEIC/ICPSEurope54979.2022.9854746.

[72] Z. Sun, L. Von Krannichfeldt, and Y. Wang, "Trading and Valuation of Day-Ahead Load Forecasts in an Ensemble Model," *IEEE Trans. Ind. Appl.*, vol. 59, no. 3, pp. 2686–2695, 2023, doi: 10.1109/TIA.2023.324171.

[73] R. Gao, P. N. Suganthan, Q. Zhou, K. Fai Yuen, and M. Tanveer, "Echo state neural network based ensemble deep learning for short-term load forecasting," *Proc. 2022 IEEE Symp. Ser. Comput. Intell. SSCI 2022*, pp. 277–284, 2022, doi: 10.1109/SSCI51031.2022.10022067.

[74] A. Valencia, O. Gorbatov, and L. Eleftheriadis, "Hardware alarms reduction in Radio Base Stations by forecasting Power Supply Units headroom," *Electr. Power Syst. Res.*, vol. 213, p. 108519, 2022, doi: 10.1016/j.epsr.2022.108519.

[75] B. Li, Y. Mo, F. Gao, and X. Bai, "Short-term probabilistic load forecasting method based on uncertainty estimation and deep learning model considering meteorological factors," *Electr. Power Syst. Res.*, vol. 225, p. 109804, 2023, doi: 10.1016/j.epsr.2023.109804.

[76] V. K. Saini, A. S. Al-Sumaiti, and R. Kumar, "Data driven net load uncertainty quantification for cloud energy storage management in residential microgrid," *Electr. Power Syst. Res.*, vol. 226, p. 109920, 2024, doi: 10.1016/j.epsr.2023.109920.

[77] H. Shi, L. Wang, R. Scherer, M. Wozniak, P. Zhang, and W. Wei, "Short-Term Load Forecasting Based on Adabelief Optimized Temporal Convolutional Network and Gated Recurrent Unit Hybrid Neural Network," *IEEE Access*, vol. 9, pp. 66965–66981, 2021, doi: 10.1109/ACCESS.2021.3076313.

[78] H. Hu and B. Zheng, "Short-term electricity load forecasting based on CEEMDAN-FE-BiGRU-Attention model," *Int. J. Low-Carbon Technol.*, vol. 19, pp. 988–995, 2024, doi: 10.1093/ijlct/ctae040.

[79] S. Cen and C. G. Lim, "Multi-Task Learning of the PatchTCN-TST Model for Short-Term Multi-Load Energy Forecasting Considering

Indoor Environments in a Smart Building," *IEEE Access*, vol. 12, pp. 19553–19568, 2024, doi: 10.1109/ACCESS.2024.3355448.

[80] A. Irankhah, S. R. Saatlou, M. H. Yaghmaee, S. Ershadi-Nasab, and M. Alishahi, "A parallel CNN-BiGRU network for short-term load forecasting in demand-side management," *2022 12th Int. Conf. Comput. Knowl. Eng. ICCKE 2022*, pp. 511–516, 2022, doi: 10.1109/ICCKE57176.2022.9960036.

[81] A. Parizad and C. J. Hatziadoniu, "A Real-Time Multistage False Data Detection Method Based on Deep Learning and Semisupervised Scoring Algorithms," *IEEE Syst. J.*, vol. 17, no. 2, pp. 1753–1764, 2023, doi: 10.1109/JSYST.2023.3265021.

[82] J. Gan, "A Load Forecasting Approach Based on Graph Convolution Neural Network," 2022 *IEEE Intl Conf Dependable, Auton. Secur. Comput. Intl Conf Pervasive Intell. Comput. Intl Conf Cloud Big Data Comput. Intl Conf Cyber Sci. Technol. Congr*., pp. 1–3, 2022, doi: 10.1109/DASC/PiCom/CBDCom/Cy55231.2022.9927829.

[83] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1701–1708, 2014, doi: 10.1109/CVPR.2014.220.

[84] N. L. W. Keijsers, "Neural Networks," *Encycl. Mov. Disord. Three-Volume Set*, pp. 257-259, 2010.

[85] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[86] L. Vinet and A. Zhedanov, "Deep Learning," *J. Phys. A Math. Theor.*, vol. 44, no. 8, pp. 1689–1699, 2011.

[87] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

[88] V. Nair and G. E. Hinton, "Rectified linear units improve Restricted Boltzmann machines," *ICML 2010 - Proceedings, 27th Int. Conf. Mach. Learn.*, no. 3, pp. 807–814, 2010.

[89] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8689, pp. 818–833, 2014, doi: 10.1007/978-3-319-10590-1_53.

[90] S. Schreck, I. Prieur, D. Comble, S. Thiem, and S. Niessen, "A Methodological Framework to support Load Forecast Error Assessment in Local Energy Markets," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3212–3220, 2020

[91] Based deep neural network for monocular depth prediction," *SN Appl. Sci.*, vol. 5, no. 1, 2023, doi: 10.1007/s42452-022-05235-1.

[92] A. P. Wibawa, A. B. P. Utama, H. Elmunsyah, U. Pujianto, F. A. Dwiyanto, and L. Hernandez, "Time-series analysis with smoothed Convolutional Neural Network," *J. Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00599-y.

[93] W. Li and K. L. E. Law, "Deep Learning Models for Time Series Forecasting: A Review," *IEEE Access*, vol. 12, pp. 92306–92327, 2024, doi: 10.1109/ACCESS.2024.3422528.

[94] A. Mahmoud and A. Mohammed, "Leveraging Hybrid Deep Learning Models for Enhanced Multivariate Time Series Forecasting," *Neural Process. Lett.*, vol. 56, no. 5, pp. 1–25, 2024, doi: 10.1007/s11063-024-11656-3.

[95] J. Mo, R. Wang, M. Cao, K. Yang, X. Yang, and T. Zhang, "A hybrid temporal convolutional network and Prophet model for power load forecasting," *Complex Intell. Syst.*, vol. 9, no. 4, pp. 4249–4261, 2023, doi: 10.1007/s40747-022-00952-x.