# Deep Learning for Peak Load Duration Curve Forecasting

**George Morris William Tangka*[1], Lidya Chitra Laoh[2]**
[1]Sistem Informasi, [2]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Klabat
e-mail: *[1]gtangka@unklab.ac.id, [2]lidya.laoh@unklab.ac.id

***Abstract***

*As the energy landscape changes towards renewable energy sources and smart grid technologies, accurate prediction of peak load duration curve (PLDC) becomes crucial to ensure power system stability. The background to this research is the urgent need for more effective prediction methods to manage increasingly complex energy loads. This research presents a leading-edge approach to PLDC prediction, leveraging Deep Learning, a subsection of artificial intelligence. Focusing on data from the Taiwan State Electric Company, this study uses a Long Short-Term Memory (LSTM) network to capture complex load patterns. The LSTM model, consisting of two layers and trained on 2019-2020 data, demonstrated excellent accuracy with a Mean Absolute Percentage Error (MAPE) as low as 0.03%. These results confirm the potential of Deep Learning to revolutionize PLDC predictions in complex energy systems. These research recommendations involve exploring diverse datasets, integrating real-time data streams, and conducting comparative analyses for more reliable prediction methodologies. The benefits of this research include providing relevant insights for sustainable energy resource management amidst a dynamic energy landscape.*

***Keywords***—Energy Forecasting, Peak Load Duration Curve (PLDC), Long Short-Term Memory (LSTM), Electric Power System, Energy Resource Management

## 1. INTRODUCTION

The energy landscape is undergoing a profound transformation with the increasing integration of renewable energy sources and the evolution of smart grid technologies [1, 2, 3]. In this context, the accurate forecasting of the Peak Load Duration Curve (PLDC) plays a pivotal role in ensuring the stability and efficiency of power systems. One key challenge in this realm is the dynamic nature of load patterns, necessitating advanced forecasting methodologies [4, 5].

This paper focuses on the critical task of forecasting PLDC using a cutting-edge Deep Learning approach. The PLDC, representing the duration and magnitude of electricity demand at various levels, is a fundamental tool for grid operators and policymakers in managing energy resources effectively. Accurate energy forecasting models can significantly aid in formulating energy policies and designing plans for the development of alternative energy sources [6]. Moreover, deep learning techniques are effective in learning patterns from large datasets generated by smart grids, thereby improving the accuracy of demand forecasting [7, 8]. Furthermore, the integration of AI-driven forecasting models enhances the ability to balance supply and demand, which is critical for maintaining grid stability and preventing outages [9]. By leveraging these advanced methods, grid operators can better predict power load changes, ensuring a stable and efficient energy supply [3]. Despite the recognized importance of PLDC forecasting, existing methodologies often struggle to capture the intricate patterns inherent in today's complex energy systems [10, 11].

This research leverages the power of Deep Learning, a subset of artificial intelligence, to address the shortcomings of traditional forecasting techniques. Deep Learning models, with their ability to discern intricate patterns in large datasets, hold promise for enhancing the accuracy and reliability of PLDC forecasts [12]. Through the utilization of advanced neural networks and datadriven methodologies, this study aims to push the boundaries of peak load forecasting.

The motivation behind this research stems from the pressing need for more precise and adaptable forecasting tools in the face of the evolving energy landscape [13]. By harnessing the capabilities of Deep Learning, we endeavor to contribute to advancing peak load forecasting, offering insights that are timely and essential for the sustainable management of energy resources.

In the subsequent sections, this study provides an in-depth exploration of the methodology used, detailing the intricacies of the Deep Learning approach employed. This study extends beyond the theoretical framework by presenting empirical results and discussing their implications. Additionally, this study acknowledges the challenges inherent in this approach. Moreover, it also delineates avenues for future research, aiming to catalyze further advancements in the energy forecasting field.

## 2. METHODOLOGY

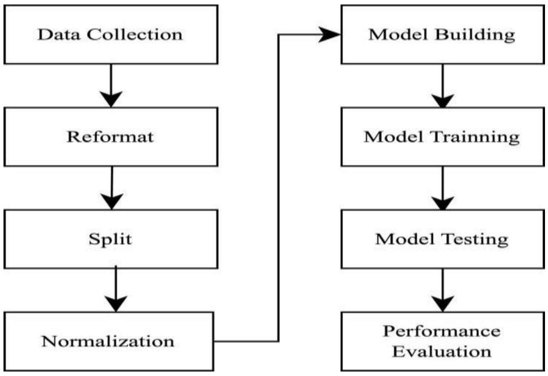


*Figure 1. Research Flow*

### *2.1. Data Collection*

The study collects data from Taiwan Power Company (TPC), a state-owned electric power company in Taiwan with the tasks of production, conveyance, and dissemination of electrical energy [14]. TPC is accountable for overseeing the complete electricity infrastructure of Taiwan, with three distinct segment regions, namely, Northern, Central, and Southern. The data collection process retrieves historical data from TPC's database, which encompasses the period between January 2019 to December 2021, with an hourly frequency as shown in Table 1. Subsequently, the gathered data undergoes preprocessing procedures. Furthermore, it is utilized to construct and assess the suggested deep-learning models.

*Table 1. Dataset*

| Year | Month | Day | 1st Hour Power Supply (MWH) | 2nd Hour Power Supply (MWH) | 3rd Hour Power Supply (MWH)) | … | 21st Hour Power Supply (MWH) | 22nd Hour Power Supply (MWH) | 23rd Hour Power Supply (MWH) | 24th Hour Power Supply (MWH) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2019 | Jan | 1 | 19764.906 | 19350.184 | 18716.08 | … | 23814.71 | 23444.212 | 22341.632 | 21042.128 |
| 2019 | Jan | 2 | 19833.707 | 18994.797 | 18565.55 | … | 26923.41 | 26231.613 | 25031.397 | 23730.813 |
| 2019 | Jan | 3 | 22094.691 | 21201.898 | 20597.482 | … | 27129.491 | 26329.096 | 25146.668 | 23718.793 |
| 2019 | Jan | 4 | 22160.877 | 21246.498 | 20700.795 | … | 26708.344 | 25946.904 | 25057.622 | 23846.548 |
| 2019 | Jan | 5 | 22486.723 | 21431.509 | 20829.685 | … | 25050.314 | 24525.529 | 23766.962 | 22585.972 |
| 2019 | Jan | 6 | 21310.137 | 20543.001 | 19883.334 | … | 24282.512 | 23921.833 | 22850.434 | 21394.803 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | ⋮ | ⋮ | ⋮ | ⋮ |
| 2021 | Dec | 27 | 22215.698 | 21362.598 | 20743.839 | … | 29168.957 | 28382.25 | 27028.268 | 25649.713 |
| 2021 | Dec | 28 | 24030.195 | 23014.854 | 22525.574 | … | 28725.84 | 28056.749 | 26899.774 | 25588.192 |
| 2021 | Dec | 29 | 24190.063 | 23169.69 | 22535.087 | … | 28370.999 | 27634.987 | 26535.703 | 25292.028 |
| 2021 | Dec | 30 | 23816.493 | 22805.723 | 22211.047 | … | 27571.436 | 27007.935 | 25950.608 | 24743.781 |
| 2021 | Dec | 31 | 23309.605 | 22340.873 | 21608.246 | … | 25484.809 | 25074.627 | 24211.809 | 23477.303 |

### 2.2. Data Pre-Processing

In ensuring the accurate results of the models, data pre-processing is essential to remove any anomalies or inconsistencies. This subsection outlines the procedures undertaken during the pre-processing of data. Reformatting is necessary to meet the requirements of this research. This process entails altering the data's structure by converting it from one format to another. The initial three columns denote the chronological sequence of the year, month, and day. The fourth column up to the sixth column denotes the maximum load experienced during the initial hour of the day. To facilitate the research endeavor, it is necessary to reformat the data into a distinct configuration comprising three distinct columns. The initial column of the dataset comprises the temporal information for the entire annual duration. The remaining columns display the hourly electricity peak load for each year in descending order.

Subsequently, the researcher examined the data to identify any instances of missing values or outliers. The identification of missing values was carried out through the utilization of the Pandas library in the Python programming language, and the subsequent filling was performed via the implementation of forward-fill and backward-fill techniques. To identify outliers, the researcher computed the z-scores for every value in the column and subsequently determined the indices of values that exceed 3 standard deviations from the mean (i.e., values with a z-score greater than 3), which are classified as outliers.

#### 2.2.1. Reformat

To facilitate the research endeavor, it is necessary to reformat the data into a distinct configuration comprising four distinct columns. The initial column of the dataset comprises the temporal information for the entire annual duration. The remaining columns display the hourly electricity peak load for each year in descending order. The reformatted data set can be seen in Table 2.

*Table 2. Reformatted Data Set*

| Duration $n^{th}$ hour | 2019 | 2020 | 2021 |
|---|---|---|---|
| 1st | 37067.066 | 37714.613 | 38606.512 |
| 2nd | 37021.664 | 37644.028 | 38585.844 |
| 3rd | 36988.158 | 37565.054 | 38485.668 |
| 4th | 36682.192 | 37561.483 | 38438.33 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 8756th | 15609.531 | 17547.768 | 17192.031 |
| 8757th | 15575.42 | 17489.571 | 17124.601 |
| 8758th | 15573.419 | 17485.715 | 17089.818 |
| 8759th | 15506.682 | 17466.267 | 17055.424 |
| 8760th | 15422.519 | 17432.848 | 16967.943 |

#### 2.2.2. Split

Following that, data sets for training and testing should be separated. The training set is utilized to train the machine learning model, allowing it to learn patterns and relationships within the data. On the other hand, the testing set serves as an independent dataset that the trained model has not encountered during the learning phase. This separation is vital for evaluating the model's ability to make accurate predictions on unseen data, providing insights into its generalization performance and potential to handle new, previously unseen instances. The model was trained using the initial two years of data spanning from 2019 and 2020, while the data in 2021 was reserved for testing purposes.

#### 2.2.3. Normalization

Within the field of machine learning, normalization stands as a fundamental preprocessing technique crucial to the effective preparation of data. The utilization of the

MinMaxScaler from the scikit-learn library in the presented code exemplifies a strategic approach to standardizing numerical features. By transforming both training and testing data, this scaler ensures a consistent scale, typically ranging between 0 and 1, for numerical variables. This normalization process is particularly pertinent in the context of neural network models, such as Long Short-Term Memory (LSTM) networks, where disparate feature scales can impede convergence and model performance. The mathematical equation for normalization can be expressed using the Min-Max Scaling method, which transforms each data point into its normalized counterpart. $X_i'$:

$$X_i' = \frac{X_i - min\ (X)}{(X) - min\ (X)} \tag{1}$$

Where:
$X_i$ = an individual data point,
$(X)$ = minimum value of the feature in the dataset
$(X)$ = maximum value.

### *2.3. Long-Short Term Memory (LSTM)*

The LSTM model, which was introduced by Hochreiter and Schmidhuber 1997, has gained significant attention in the field of time series data analysis and prediction [15]. It is a type of Recurrent Neural Network (RNN) that has been widely utilized for this purpose. The conventional RNN is recognized to exhibit a notable reduction in gradient during backpropagation when the gap between the pertinent data and the point of its utilization is extensive [16, 17].

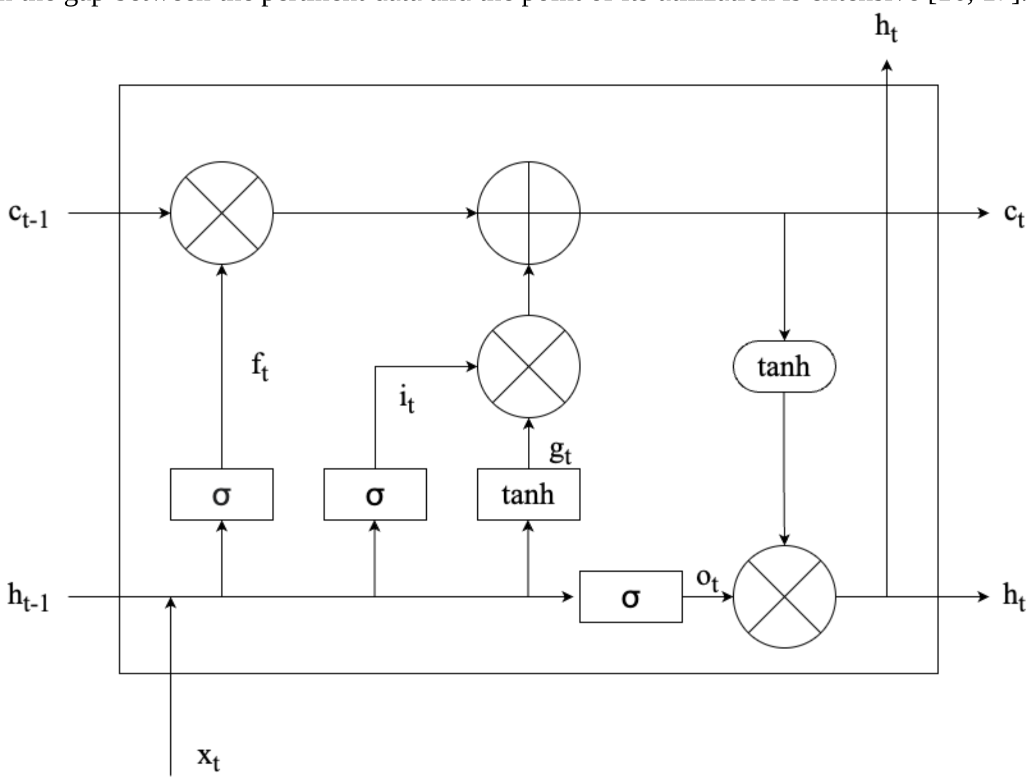


*Figure 2. LSTM Architecture*

The LSTM model is comprised of memory blocks that possess a distinctive characteristic referred to as the cell state. $c_{t-1}$, hidden state $h_{t-1}$, and input state $x_t$ Engage in interaction with the LSTM gates: input gate $x_t$, input node $g_t$, forget gate $f_t$, and output gate $o_t$. The first gate in the LSTM that determines whether data to preserve or discard from the previous cell state is the forget gate, defined as:

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{2}$$

The output of the forget gate is denoted by $f_t$, while the parameters (weights and bias) are represented by $W_{fh}$, $W_{fx}$, and $b_f$. The function of sigmoid activation is denoted by the symbol σ. The input gate is responsible for determining the specific information that will be incorporated into the cell state. The process is executed through a dual-step approach, wherein each step employs a distinct activation function. Consequently, the entries undergo the sigmoid function described as:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{3}$$

Resulting in output values ranging from 0 to 1, where $i_t$ Represents the output of the input gate. The parameters of the input gate include $W_{ix}$, $W_{ih}$, and $b_i$. The entries undergo the tanh function as in:

$$g_t = tan\,h\left(W_{gx}x_t + W_{gh}h_{t-1} + b_g\right) \tag{4}$$

Where the new candidate's values are denoted as $g_t$, while the activation function used is $tan\,h$, thereby generating a vector of fresh candidate values to facilitate the update of the cell state. The updated cell state is described as:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{5}$$

Updates the cell state by adding the product of the results obtained from both activation functions. Finally, the output gate ascertains the information of the cell state that is to be produced, as described in:

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{6}$$
$$h_t = o_t \odot tanh\,(c_t) \tag{7}$$

where the output gate weights and bias, denoted as $W_{oh}$, $W_{ox}$, and $b_o$, respectively, are utilized in conjunction with the memory cell's output vector, represented as $h_t$.

### *2.4. Performance Evaluation*

The inclusion of a quantitative measure for performance evaluation is imperative for this research, as its objective is to assess and contrast the prognostic capabilities of diverse predictive models [13, 18, 19, 20]. Assessing the accuracy of untrained data is commonly conveyed by utilizing the Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE); respectively are expressed in:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \tilde{y}_t| \tag{8}$$

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(y_t - \tilde{y}_t)^2 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\tilde{y}_t - y_t)^2} \quad (10)$$

$$MAPE(\%) = \left(\frac{1}{N}\sum_{t=1}^{N}\left(\frac{|\tilde{y}_t - y_t|}{y_t}\right)\right) \cdot 100 \quad (11)$$

Where $N$ signifies the total count of data points, $y_t$ corresponds to the observed peak load at a given time $t$, and $\tilde{y}_t$ stands for the predicted peak load at the same time $t$. This is done to avoid any potential instances of content duplication.

## 3. RESULTS AND DISCUSSION

The experiments were conducted using historical peak load data from TPC, which consists of hourly power consumption values from January 1st, 2019, until December 31st, 2021. The dataset underwent partitioning into distinct training and testing sets. The data from 2019 and 2020 were utilized as the training set, while the test set consisted of 2021 data. Figure 3 displays the distribution of testing and training sets.

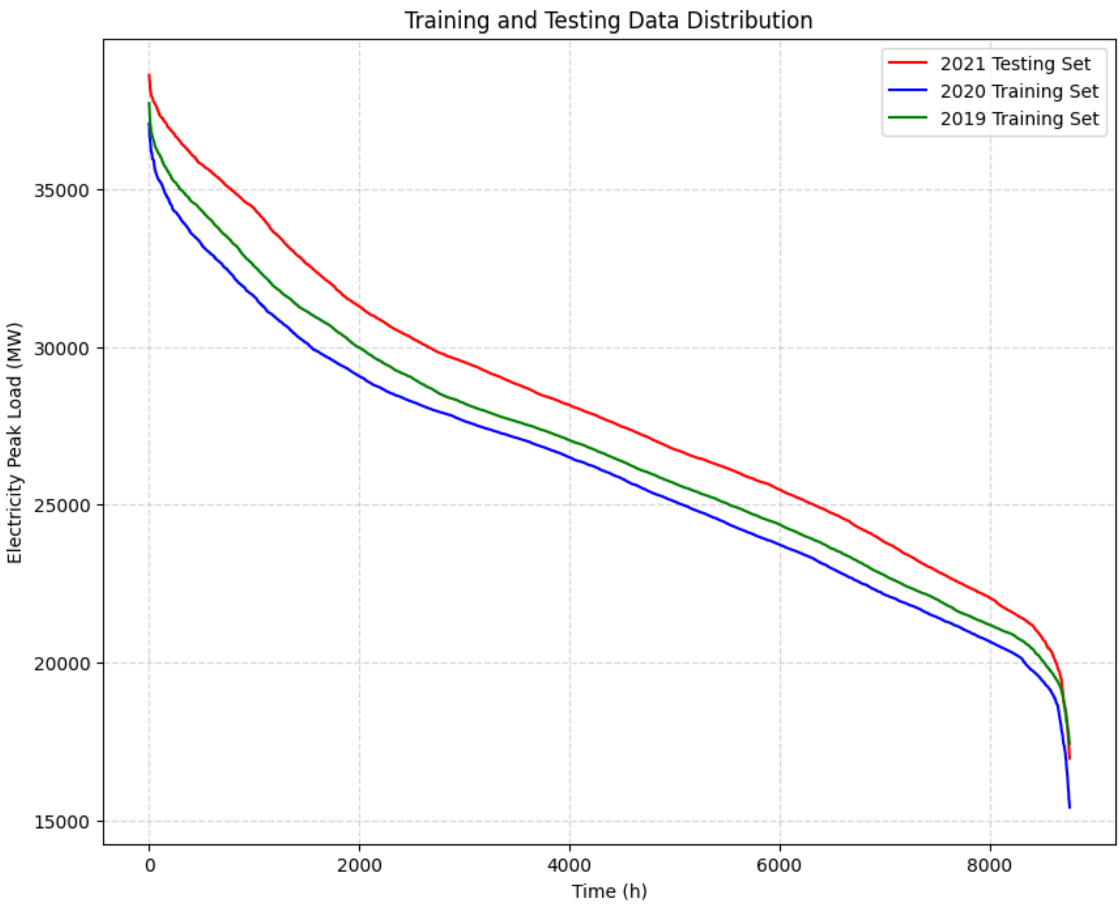


*Figure 3. Data Distribution*

The LSTM model necessitates time series data preparation, accounting for the temporal dependence in load data. A data structure was constructed to predict the current load value based on the preceding load value, utilizing a custom function. The LSTM model was established with the Keras library, comprising two LSTM layers and a Dense layer. The initial LSTM layer incorporated 100 neurons and implemented a 'return_sequences' attribute to enable multiple LSTM layers. The subsequent LSTM layer incorporated 50 neurons. The final Dense layer acted as the output layer, producing the predicted load values.

The model was compiled utilizing MSE as the loss function and the Adam optimizer; the LSTM model underwent training on the preprocessed training dataset for a total of 50 epochs, utilizing a batch size of 32. After the completion of the training phase, the model made prognostications regarding the load values for the testing dataset. These predictions were then subjected to an inverse transformation, thereby restoring them to their original scale for the sake of lucidity. The model's performance was evaluated with MAE, MSE, RMSE, and MAPE metrics. Figure 4 displays the forecasted outcome of the PLDC using the LSTM model.

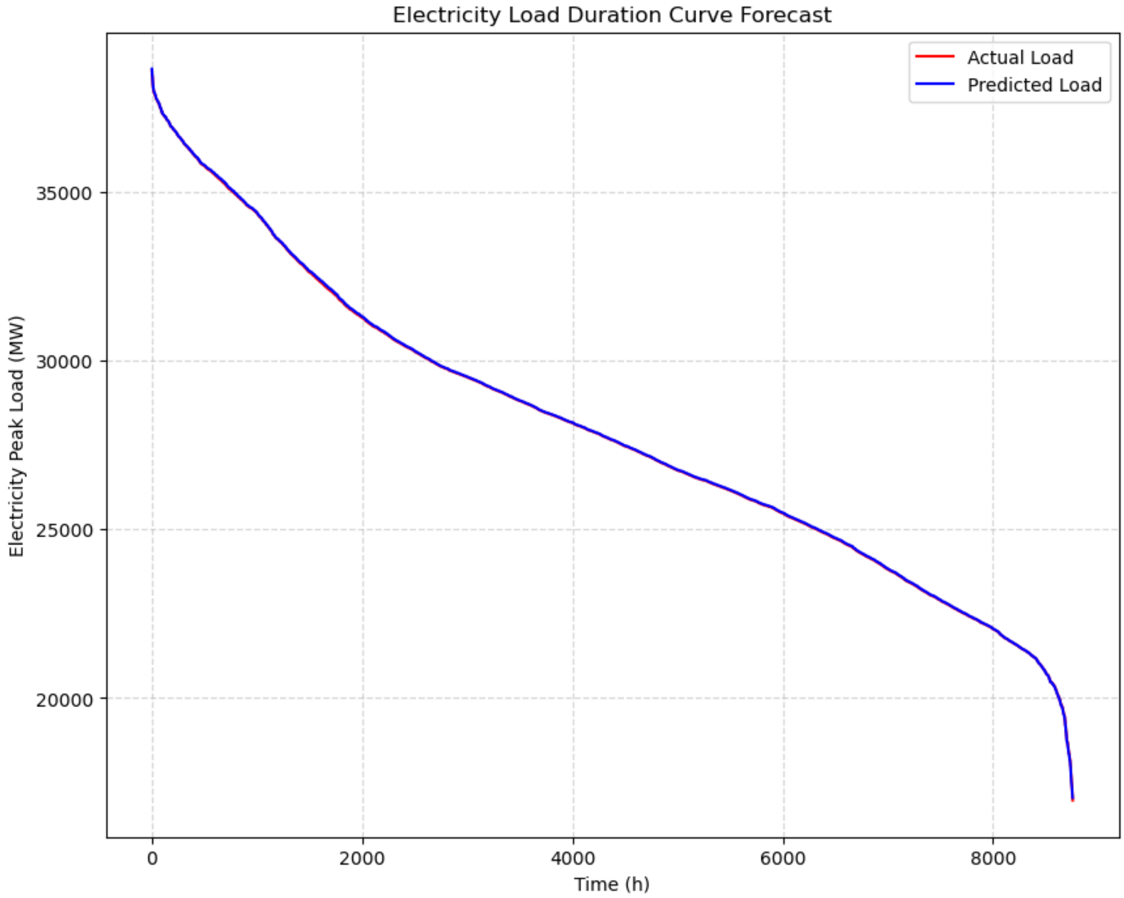


*Figure 4. Forecasting Result*

The model exhibited remarkable accuracy with an MAE of approximately 8.84, MSE of approximately 108.57, RMSE of 10.42, and MAPE of 0.03 %.

## 4. CONCLUSION

The remarkable forecasting accuracy achieved by the LSTM model holds profound implications in the context of the evolving energy landscape outlined in the background. As the energy sector undergoes a transformative shift towards increased reliance on renewable sources and the integration of smart grid technologies, accurate peak load demand forecasting emerges as a linchpin for ensuring the stability and efficiency of power systems. Traditional forecasting methods have often struggled to capture the intricate patterns inherent in today's complex energy systems, emphasizing the need for advanced methodologies.

This study specifically addresses the challenge of forecasting Peak Load Duration Curves (PLDC) using Deep Learning, a subset of artificial intelligence. The PLDC, representing the duration and magnitude of electricity demand at various levels, is a fundamental tool for grid operators and policymakers in managing energy resources effectively. The success of the LSTM model in achieving a Mean Absolute Percentage Error (MAPE) as low as 0.03% underscores its potential to revolutionize PLDC forecasting.

## 5. RESEARCH LIMITATION AND RECOMMENDATION

While this study contributes valuable insights into the realm of peak load forecasting using LSTM-based Deep Learning models, it is imperative to acknowledge certain limitations that warrant consideration. Firstly, the generalizability of the findings may be influenced by the specific characteristics of the dataset used in this research. Future studies should explore diverse datasets representing varied energy landscapes to enhance the external validity of the proposed model.

Secondly, the dynamic nature of energy systems introduces complexities that might not be fully encapsulated by the LSTM model. Incorporating additional contextual factors, such as policy changes, economic shifts, or unforeseen external events, could further enrich the forecasting model. The present study assumes a static scenario, and future research could benefit from a more dynamic and adaptive approach.

To advance the field of peak load forecasting and address the identified limitations above, several avenues for future research are proposed. Firstly, researchers should explore the integration of real-time data streams and continuous model updating to enhance the adaptability of forecasting models to changing energy dynamics. This could involve the incorporation of IoT devices and advanced sensor technologies.

Furthermore, investigating the transferability of the LSTM model across different geographical regions and energy infrastructures is essential. Each region has unique energy consumption patterns, and developing models that can be easily adapted to diverse contexts would contribute to the robustness of forecasting methodologies.

Additionally, a comparative analysis between LSTM-based models and other advanced forecasting techniques, such as hybrid models or ensemble methods, could provide a comprehensive understanding of the strengths and weaknesses of each approach. This comparative assessment would guide practitioners and policymakers in selecting the most suitable forecasting methodology based on their specific requirements.

## REFERENCES

[1] United Nations Department of Economic and Social Affairs, *The Sustainable Development Goals Report 2023: Special Edition. in The Sustainable Development Goals Report*. United Nations, 2023. doi: 10.18356/9789210024914.

[2] S. Cevik, "Climate Change and Energy Security: The Dilemma or Opportunity of the Century?" *IMF Working Paper*, no. 2022/174, International Monetary Fund, Washington, DC, Sep. 2022. [Online]. Available: https://www.imf.org/-/media/Files/Publications/WP/2022/English/wpiea2022174-print-pdf.ashx.

[3] International Energy Agency, "*World Energy Outlook 2020*," IEA. Accessed: Jun. 01, 2023. [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2020

[4] J. A. Dowling et al., "Role of Long-Duration Energy Storage in Variable Renewable Electricity Systems," *Joule*, vol. 4, no. 9, pp. 1907–1928, Sep. 2020, doi: 10.1016/j.joule.2020.07.007.

[5] Y. Liu, R. Zheng, and J. Yuan, “The economics of peaking power resources in China: Screening curve analysis and policy implications,” *Resources, Conservation and Recycling*, vol. 158, p. 104826, Jul. 2020, doi: 10.1016/j.resconrec.2020.104826.

[6] K.-C. Chen and L.-Y. Chuang, "Deep Learning for Forecasting Electricity Demand in Taiwan," *Mathematics*, MDPI, vol. 10, no. 14, p. 2547, Jul. 2022. [Online]. Available: https://www.mdpi.com/2227-7390/10/14/2547.

[7] J. M. Aguiar-Pérez and M. Á. Pérez-Juárez, "An Insight of Deep Learning Based Demand Forecasting in Smart Grids," *Sensors*, MDPI, vol. 23, no. 3, p. 1467, Jan. 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/3/1467.

[8] "Optimization and Research of Smart Grid Load Forecasting Model Based on Deep Learning," *International Journal of Low-Carbon Technologies*, Oxford Academic, 2023.

[9] V. Franki, D. Majnarić, and A. Višković, “A Comprehensive Review of Artificial Intelligence (AI) Companies in the Power Sector,” *Energies*, vol. 16, no. 3, Art. no. 3, Jan. 2023, doi: 10.3390/en16031077.

[10] A. Azeem, I. Ismail, S. M. Jameel, and V. R. Harindran, “Electrical Load Forecasting Models for Different Generation Modalities: A Review,” *IEEE Access*, vol. 9, pp. 142239–142263, 2021, doi: 10.1109/ACCESS.2021.3120731.

[11] Y. Amara-Ouali, M. Fasiolo, Y. Goude, and H. Yan, “Daily peak electrical load forecasting with a multi-resolution approach,” *International Journal of Forecasting*, Jul. 2022, doi: 10.1016/j.ijforecast.2022.06.001.

[12] E. J. A. Nainoon and I. O. Habiballah, “Load Forecasting using Machine Learning Methods: Review,” *International Journal of Engineering Research & Technology*, vol. 11, no. 12, Dec. 2022, doi: 10.17577/IJERTV11IS120094.

[13] A. Hamdan et al., “AI in renewable energy: A review of predictive maintenance and energy optimization,” *International Journal of Science and Research Archive*, vol. 11, no. 1, Art. no. 1, 2024, doi: 10.30574/ijsra.2024.11.1.0112.

[14] Taiwan Power Company, *“Sustainability Report.”* 2022. [Online]. Available: https://csr.taipower.com.tw/en/download/sustainables/Aw/2022_csr.pdf

[15] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[16] A. Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network,” *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[17] G. Van Houdt, C. Mosquera, and G. Nápoles, “A review on the long short-term memory model,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[18] S. D. Haleema, “Short-Term Load Forecasting using Statistical Methods: A Case Study on Load Data,” *International Journal of Engineering Research & Technology*, vol. 9, no. 8, Aug. 2020, doi: 10.17577/IJERTV9IS080182.

[19] G. Nalcaci, A. Özmen, and G. W. Weber, "Long-term load forecasting: models based on MARS, ANN and LR methods," *Central European Journal of Operations Research*, vol. 27, no. 4, pp. 1033–1049, Dec. 2019, doi: 10.1007/s10100-018-0531-1.

[20] P. Koponen, J. Ikäheimo, J. Koskela, C. Brester, and H. Niska, "Assessing and Comparing Short Term Load Forecasting Performance," *Energies*, vol. 13, no. 8, Art. no. 8, Jan. 2020, doi: 10.3390/en13082054.