SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I5P113

Volume 11 Issue 5, 138-149, May 2024

Original Article

Short Term Load Forecasting for Smart Distribution System Planning Using Deep Neural Networks: A Hybrid Approach

Katkar Siddhant Satyapal[1], Arunkumar Patil[2], Kunal Samad[3], Santosh Diggikar[4]

[1,2,3,4]*Department of Electrical Engineering, Central University of Karnataka, Karnataka, India.*

[2]*Corressponding Author : arunkumarpatil@cuk.ac.in*

Received: 08 March 2024 Revised: 09 April 2024 Accepted: 07 May 2024 Published: 29 May 2024

***Abstract** - Accurate load forecasting plays a crucial role in the management and control of electrical power in distribution systems. Short-Term Load Forecasting (STLF) is particularly vital for distribution planning, as it provides precise load predictions for the immediate future. This paper introduces an innovative hybrid deep-learning model specifically designed for STLF systems. The proposed hybrid model combines the strengths of Bidirectional Long Short-Term Memory (Bi-LSTM) and Gated Recurrent Unit (GRU) networks. The study utilizes a high-resolution real-world dataset, consisting of historical load consumption and weather-related features, with 30-minute intervals from the period of January 1, 2006, to December 31, 2010. This model is benchmarked against prominent standalone models such as Bi-LSTM, GRU, LSTM, and CNN, and hybrid models like CNN-LSTM and ConvLSTM-GRU. The model's performance is evaluated using various validation metrics, including R-squared error, Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The results show that the proposed model outperforms all conventional models, offering significant improvements in forecast accuracy. Thus, the study highlights the potential of hybrid models in revolutionizing forecasting methodologies, paving the way for a smart distribution system.*

***Keywords** - Short Term Load Forecasting (STLF), Smart distribution system, High-resolution dataset, Bi-LSTM, GRU, Validation metrics, Hybrid Model.*

1. Introduction

A power system is a network of electrical components and devices that work together to generate, transmit, and distribute electricity. It encompasses power generation plants, transmission lines, substations, and a distribution network. The distribution system is a vital component that delivers electricity from substations to end-users, ensuring a reliable and efficient supply. Its importance lies in bridging the gap between large-scale power generation and individual consumers, enhancing the resilience and flexibility of the overall power grid.

Effective management of the distribution system involves proper distribution planning, a process that strategically designs and optimizes the layout of electrical infrastructure. Distribution planning is crucial for meeting the rising demand for electricity, ensuring that system expansions are both technically sound and economically viable. This process is influenced by various factors, including load forecasting, substation expansion, site selection, load assignment to substations, and considerations for primary voltage, feeder routes, number of feeders, conductor size, and overall cost [1].

By comprehensively addressing these factors, the planning ensures the quality and safe delivery of electricity. Load forecasting is one of the significant components of the effective management of power distribution planning. It categorizes itself based on the forecasted periods, as outlined in Table 1. The accuracy of load forecasting across various time frames profoundly influences the stability and operational costs of the power grid.

A case study from a utility company in the United Kingdom highlights the financial implications of load forecasting. The study revealed that a mere 1% decrease in load forecasting error led to an impressive annual operating cost reduction of 10 million pounds [2] . This underscores the importance of precise load forecasting in power distribution systems, emphasizing its role in ensuring grid stability and contributing to substantial cost savings. The roots of load forecasting trace back to the insights of Samuel Insull, a pioneer in the electric utility industry, in 1894. Insull's analysis of load usage patterns revealed distinct peak consumption periods during the day for domestic users and at night for industrial consumers [3].

Table 1. Types of load forecasting and their applications

Type of Load Forecasting	Time Period	Application Areas						
		Real-Time System Operation	Automatic Generation Control	Security Assessment	Maintenance Scheduling	Fuel Reserve Management	Capacity Expansion	Budgeting
Short-Term	(Several min up to one Week)	✓	✓	✓				
Medium-Term	(Few weeks to months)				✓	✓		
Long-Term	(Several months to years)						✓	✓

This early understanding laid the groundwork for modern load forecasting, enabling utilities to make efficient decisions on unit commitment, spinning reserve minimization, and maintenance scheduling. Moreover, it optimizes power flow in transmission networks, mitigating the risks of underloads and overloads. Precise load forecasting goes beyond immediate operational benefits, resulting in significant cost savings in operations and maintenance. However, the landscape of load forecasting is evolving in tandem with the modernization of power grids.

This modernization is happening through the integration of Renewable Energy Resources (RER), implementation of Demand Response Mechanisms (DRM), and increasing the prevalence of electric vehicles. These factors introduce heightened uncertainty, posing a significant challenge to load forecasting. All these changes are steering the development of Smart Distribution Systems. Such systems may encounter more unpredictable load deviations due to end-user behaviour, necessitating a shift towards more precise and high-resolution load forecasting techniques.

Addressing these challenges necessitates the integration of advanced data acquisition systems, such as Advanced Metering Infrastructure, which are capable of providing high-resolution datasets at intervals ranging from a minute to several hours. The vast amount of data generated by these systems presents a significant challenge. To overcome these hurdles, efficient data handling solutions through the integration of data analytics and artificial intelligence become essential. This fusion of technology and energy systems has sparked a surge of interest in load forecasting in both academic and industrial domains. These technologies not only streamline data processing but also contribute to enhancing the accuracy of load forecasting within the context of energy systems.

Short Term Load Forecasting provides forecasts for up to a few hours ahead and plays a crucial role in maintaining system balance and security. It assists in automatic generation control, ensuring a consistent match between electricity supply and demand, thereby preventing power outages and maintaining grid stability. Furthermore, STLF supports real-time market operations, aiding utilities in optimizing their energy trading strategies. Despite the existence of medium-term and long-term load forecasting, the significance of STLF in maintaining real-time operational efficiency and system stability underscores its critical role in distribution planning.

Traditional statistical techniques for Short-Term Load Forecasting (STLF), also known as parametric methods, face several challenges that can affect their accuracy and efficiency. One major obstacle is the limited availability of historical data, especially in rapidly changing or newly developed areas, making it challenging to make precise predictions about future loads. Seasonal variations and sensitivity to weather conditions present another set of challenges, as conventional methods may struggle to capture the impact of extreme weather events on energy consumption patterns.

Another difficulty in STLF lies in the seemingly random nature of load movements, posing a significant barrier to accurate predictions. To overcome this challenge, leveraging the success of Artificial Neural Networks (ANN), particularly Deep Neural Networks (DNN) based models, has gained popularity in recent years [4]. Deep neural networks are preferred for STLF due to their proficiency in identifying intricate patterns in data. They are particularly effective at discerning spatial attributes and time-dependent trends in load data, which greatly improves the accuracy of forecasts.

Furthermore, DNNs have demonstrated superior precision and robustness in their predictions compared to other forecasting models. Hybrid models for STLF utilize a variety of deep neural network architectures to boost the accuracy of predictions compared to standalone models. They excel in detecting spatial characteristics and time-related patterns in load data, which significantly enhances the forecasting performance. This hybrid approach often treats

STLF as a regression-based problem, utilizing regression analysis as a crucial mathematical tool. Regression analysis offers distinct advantages in load forecasting. Firstly, it offers an understanding of how variables relate in terms of strength and direction, allowing for the use of multiple predictors. This enables the prediction of outcomes even when these predictors exhibit correlations among themselves. Secondly, regression analysis serves as a corrective mechanism for errors rooted in prior assumptions. Lastly, remarkable results can be achieved with a relatively modest dataset, making it a pragmatic choice for short-term load forecasting [5].

Hence, the proposed research work introduces an innovative approach-a regression model of hybrid DNN architecture. This hybrid model combines Bi-directional LSTM and GRU, aiming to provide a more accurate solution for STLF. This approach addresses the limitations of traditional Recurrent Neural Networks (RNN), especially in handling long input dependencies and processing large datasets. Extensive experiments conducted as part of the research showcase the effectiveness of the hybrid model's architecture in training, testing, and validating the dataset.

The results present a promising advancement in short-term load forecasting methodology, offering a potential solution to the challenges posed by the modernization of power grids. As the energy landscape continues to evolve, accurate and efficient load forecasting remains paramount for ensuring the reliability and sustainability of power systems. The proposed hybrid DNN approach represents a step forward in meeting these evolving challenges, contributing to the ongoing optimization of distribution systems operations and planning.

The major contributions of this study are:

1. An innovative hybrid approach is introduced, combining two potent deep learning techniques, Bi-LSTM and GRU. This unique fusion leverages the individual strengths of these two standalone models, thereby enhancing the accuracy of load predictions.
2. The research establishes that this deep learning network achieves superior accuracy when compared to conventional methods, presenting a substantial enhancement in the efficiency of power grid operations.
3. The proposed method exhibits versatility, capable of being applied to diverse types of load data, rendering it a valuable tool for various applications within the energy sector. These contributions have the potential to propel advancements in load forecasting and deliver significant advantages to the power and energy sector.

The rest of the paper is organized as follows: Section 2 delves into an extensive review of the existing literature on short-term load forecasting. The spotlight is on hybrid models that employ Artificial Neural Networks (ANN) and Deep Neural Networks (DNN). These sophisticated Machine learning techniques have demonstrated promising outcomes in the realm of forecasting applications. Moving on to section 3, the paper presents the dataset used for the study and provides a detailed description of the developed models. In section 4, the paper outlines the experimental framework, detailing how the model was trained and evaluated. It also presents the results derived from these experiments, offering valuable insights into the performance and effectiveness of the proposed model. Finally, section 5 concludes the paper.

2. Review of Literature

Deep Learning (DL) has become a formidable tool in Short-Term Load Forecasting because of its capacity to manage vast datasets and generate precise predictions. In contrast to conventional Machine Learning approaches, Deep Learning models possess the capability to autonomously acquire and enhance their predictive abilities through experience, which eliminates the need for explicit programming.

They are capable of handling complex, high-dimensional datasets, making them more suitable for forecasting. Machine Learning, while effective in many scenarios, may not always be the best choice for predictive models. This is primarily due to its limitations in handling large-scale, non-linear data and its reliance on hand-engineered features. Deep Learning, on the other hand, can automatically extract useful features from raw data, making it a more robust choice for STLF. In recent years, various Deep Learning models have been employed for STLF which includes CNN, RNN, LSTM, GRU, Bi-LSTM and so on. Convolutional Neural Networks (CNNs) offer good results for both one-quarter and 24-hour-ahead forecasts [6].

Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM), have also been utilized due to their ability to capture temporal dependencies in time-series data [7, 8]. All these methods have shown to be more suitable due to the non-linear nature of electric load data. They can extract dynamic features from the data as well as make accurate predictions. However, standalone models have their limitations. For instance, CNNs are primarily designed for spatial data and may not perform well on temporal data.

On the other hand, RNNs are designed for temporal data and may not capture spatial features effectively. This led to the development of hybrid models, which combine the strengths of multiple models to overcome these hurdles. Hybrid models, such as those combining LSTM networks and CNN models, have shown promising results in STLF. They leverage the strengths of both models, with CNNs capturing spatial features and LSTMs capturing temporal features. This results in a more robust and accurate forecasting model. Likewise, there are various hybrid models developed to minimize the error and make it robust, which are discussed below in Table 2.

Table 2. Hybrid models for STLF based on deep neural networks

Reference	Publish Year	Model	Validation Metrics	Outcome
[16]	2021	CNN-STACK-BiLSTM	RMSE, MAE, MSE and MAPE	The proposed CNN-STACK-Bi-LSTM hybrid model provides more accuracy than CNN-LSTM, ANN, DNN and CNN-Bi-LSTM.
[17]	2022	ConvLSTM-GRU	MAE, MSE,RSME and MAPE	The proposed ConvLSTM-GRU model has achieved greater precision than FCBRM, BPTT, CNN-LSTM, and CNN-M-BLSTM.
[18]	2022	LSTM-RNN	MSE, RSME and MAPE	The proposed LSTM-RNN model provides less error and more accuracy than standalone LSTM and ANN.
[19]	2022	VMD and GRU-TCN	RMSE & MAPE	The proposed VMD and GRU-TCN model provides more accuracy than compared to GRU, TCN, LSTM, Prophet, XG Boost VMD-GRU, VMD-TCN and VMD-GRU-TCN.
[20]	2022	RMR-HFS-LSTM	MAPE & RSME	The performance of the proposed RMR-HFS-LSTM model outperforms other models, such as MLP and RNN.
[21]	2022	MTMV-CNN-LSTM	R^2, RMSE, MAE, and MAPE	The hybrid model proposed, based on MTMV-CNN-LSTM, effectively tackles challenges related to excessive repetitive data and suboptimal convolution effects. This approach notably improves the overall generalizability and accuracy of the model.
[22]	2022	GA-BiGRU	RMSE & MAPE	The results of the GA-BiGRU evolutionary deep learning technique suggest superior forecasting accuracy compared to other methods.
[23]	2022	TCN-DNN	RMSE & MAE	Proposed TCN-DNN model has achieved more accurate short-term load forecasting than standalone LSTM, TCN and DNN.
[24]	2022	CNN–OTHER	R^2, RMSE, NRMSE, and MAPE	The predictions indicate that combining the CNN model with RNN, LSTM, Bi-LSTM, or GRU can enhance accuracy.
[25]	2022	EEMD-GRU	RMSE, MAE, and MAPE	The experimental findings indicate that the proposed EBGRU model outperforms other deep learning models, such as EBRNN and EBLSTM.
[26]	2023	RESNET-LSTM	RMSE, MAE, and MAPE	Proposed ResNet-LSTM model provides more accuracy than compared to MLR, LSTM, CNN, ResNet, and CNN-LSTM.
[27]	2023	LSTM-Split Convolution	RMSE, MAE, and MAPE	Proposed LSTM-SC provides more accuracy compared to CNN-LSTM, LSTM, CNN-LSTM, DNN, M-BDLSTM, ResNet-LSTM and ANN-LSTM.
[28]	2023	Bi-LSTM-Random Forest	MAPE& RSME	The hybrid deep learning model combining Bi-LSTM and random forest proves to be more effective in enhancing forecasting accuracy and achieving superior forecasting results compared to individual models such as DBN, Bi-LSTM, RF, and CNN-LSTM.
[29]	2023	Bi-LSTM-AWDO	RMSE & MAPE	The ensemble model proposed, based on Bi-LSTM-AWDO, demonstrates superior accuracy compared to other models.
[30]	2023	VMD–CNN–LSTM and VMD–CNN–GRU	RMSE, MAE, and MAPE	The VMD–CNN–LSTM and VMD–CNN–GRU models proposed here exhibit a more comprehensive utilization of data, resulting in accurate forecasting compared to CNN-LSTM, GRU, MLP, LSTM, CNN, and CNN-GRU.

This rigorous literature review has been carried out to identify unexplored dimensions within hybrid algorithms that have not yet been applied to Short Term Load Forecasting. After an exhaustive examination of various hybrid models, in this paper, an innovative approach is proposed that combines the strengths of Bi-LSTM and GRU networks. This hybrid model capitalizes on the benefits of both networks, culminating in a robust and precise short-term load forecasting system.

The Bi-LSTM network, renowned for its ability to capture dependencies in sequential data from both past and future, is employed to model the temporal characteristics of the load data. The GRU network, recognized for its capability to learn long-term dependencies, supplements the Bi-LSTM by identifying complex patterns in the load data. This pioneering approach signifies a substantial advancement in the field of forecasting, with the potential to surpass existing models in terms of prediction accuracy.

3. Dataset and Proposed Model

3.1. Brief about the Dataset

The dataset is collected from [31] in order to analyze the system load behaviour. It consists of a half-hourly record dated from 1 January 2006 to 31 December 2010. The load behaviour is influenced by various factors, which can be broadly categorized into economic, temporal elements, weather conditions, and random disturbances [32]. The 'Dew Point' signifies the temperature at which the air becomes saturated, unable to retain all moisture in gas form, causing some to condense into liquid water droplets.

The 'Wet Bulb Temperature' is the minimum temperature achievable through evaporative cooling, always between the dry bulb temperature and the dew point. 'Relative Humidity,' on the other hand, measures the moisture in the air relative to the maximum it could hold at that temperature. Random disturbances, although not included in the current model, could potentially represent unexpected events or anomalies not captured by the other factors.

The dataset also includes parameters derived from the 'SYSLoad' variable, such as 'SYSLoad_prev_hr_avg' (the previous 24-hour average load), 'SYSLoad_lag24' (the 24-hour lagged load), and 'SYSLoad_lag168' (the 168-hour or previous week lagged load). These derived parameters enhance the model by capturing short-term trends and daily and weekly cyclical patterns in the system load. The dataset uses all these factors as predictors, with 'SYSLoad' serving as a response, providing a robust framework for understanding and predicting system load behaviour. The ratio for training and testing of a dataset is 80:20.

3.2. Deep Learning Networks

In the field of STLF, deep neural networks have proven to be powerful tools. In this, the discussion is focused on three specific models: Bi-LSTM, GRU and the Proposed Hybrid Model. The selection of Bi-LSTM and GRU is justified by their exclusive incorporation into the hybrid model, which omits other alternatives. Figure 1 illustrates the operational flow encompassing all the developed models. The essential requirement for successful STLF models lies in their ability to adeptly capture temporal dependencies within the dataset for enhancement in forecasting accuracy.

While each model has its strengths and drawbacks, the collective utilization of them in a hybrid model consistently produces superior results. This strategic combination not only boosts robustness but also ensures reliability, providing an effective solution for Short-Term Load Forecasting.

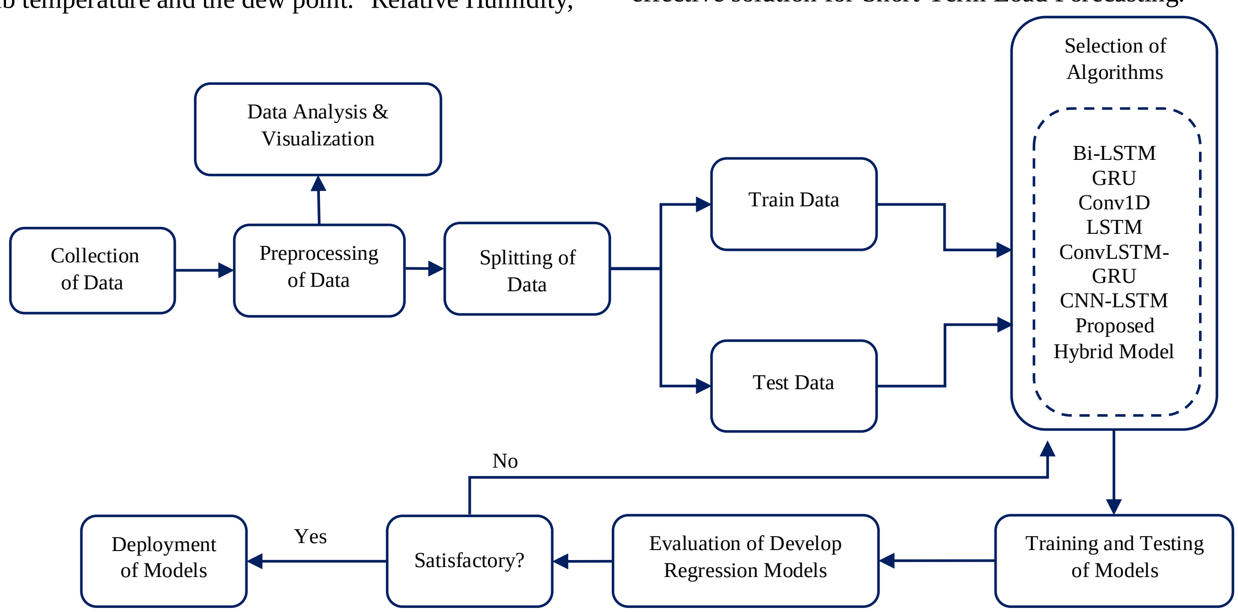

Fig. 1 Procedural flow of developed models

3.2.1. Bidirectional-LSTM

LSTM is a type of RNN architecture designed by 'Sepp Hochreiter and Jurgen Schmidhuber' in 1997 to capture and learn dependencies in sequential data [33]. Its unique memory cell structure enables it to retain and utilize information over extended periods, making it particularly effective for tasks such as time series prediction.

The Bidirectional LSTM is an extension of traditional LSTMs that can enhance model performance on sequence classification or regression problems. In scenarios where all time steps of the input sequence are available, Bi-LSTMs train two LSTMs on the input sequence. The first LSTM is trained on the input sequence as-is, and the second is on a reversed copy of the input sequence. Outputs at the same step are typically concatenated. This approach provides additional context to the network, resulting in faster and more comprehensive Learning of the problem. The equations representing the computation of Bi-LSTM are described below:

Forget Gate

The forget gate decides what information should be thrown away or kept. Input data from the current and previous hidden state is passed through this gate. The equation is:

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{1}$$

Input Gate

The input gate updates the cell state with new information. The equations are:

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = tanh(W_C.[h_{t-1}, x_t] + b_C) \tag{3}$$

Cell State

The cell state is updated in this step. The old cell state is multiplied by the forget vector, which forgets the things decided to forget earlier. Then, it adds the scaled input to the state.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

Output Gate

The output gate uses the current input and the previous hidden state to decide what information the hidden state should carry. The output gate's output is then multiplied by the cell state, and this product forms the current hidden state.

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * tanh(C_t) \tag{6}$$

Where, $x_t$is the input at time step (t), h_{t-1} is the hidden state at the previous time step(f_t ,i_t,$\tilde{C}_t$ C_t and o_t) are the forget gate, input gate, cell input, cell state, and output gate values, respectively, at time step (t), W_o and b_o are the weights & biases, σ is the sigmoid function and $tanh$ is the hyperbolic tangent function.

3.2.2. GRU

It is also a variant of recurrent neural networks designed by 'Kyunghyun Cho and others for sequence modelling in various domains, including text, speech, and time-series data [34]. Like LSTM, it utilizes gating mechanisms to control the flow of information through the network, enabling efficient learning of long-range dependencies in sequential data while simplifying the architecture by combining memory and hidden states. It introduces the concept of a gating mechanism that modulates the flow of information inside the unit without using a memory unit. The equations representing the computation of the GRU network are described below:

Update Gate

Decides what information to throw away and what new information to add.

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{7}$$

Reset Gate

Decides how much past information to forget.

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{8}$$

Current Memory

Content stores the relevant information from the past.

$$\tilde{h}_t = tanh(W.[r_t * h_{t-1}, x_t]) \tag{9}$$

Final Memory at the Current Time Step

It is a combination of the current memory content and the memory modulated by the update gate.

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{10}$$

Where, x_t is the input at the current time, σ is the sigmoid function, $W_z$weight matrix for the update gate. h_{t-1} is the hidden state at the previous time step, W_r is the weight matrix for the reset gate, and $tanh$ is the hyperbolic tangent function.

3.2.3. Proposed Hybrid Model

The proposed model effectively combines two powerful deep learning techniques: Bi-LSTM and GRU networks. During the experimental phase, it was observed that increasing the number of layers or units led to overfitting. To address this issue, the model was optimized to include three distinct layers.

The first layer is a Bi-LSTM layer equipped with 100 units, which processes the input data by capturing dependencies from both past and future sequences. This layer is configured with 'return_sequences=True', allowing it to provide a full sequence of outputs to the subsequent layer for comprehensive processing.

Following the Bi-LSTM, the second layer is a GRU layer with 50 units. This layer refines the understanding of the data by capturing intricate patterns that the Bi-LSTM might have missed. It is set with 'return_sequences=False', meaning it only outputs the last element of its processed sequence, thus focusing on the most relevant information for prediction.

The final component of the model is a Dense layer with a single unit, which consolidates the outputs from the previous layers to generate the final load forecast. The model utilizes Mean Squared Error (MSE) as the loss function, chosen for its effectiveness in highlighting larger discrepancies between the predicted values and the actual values. The Adam optimizer is employed to minimize this error, optimizing the model for better accuracy. Training is executed with a batch size of 16 for up to 50 epochs.

To prevent overfitting, an early stopping mechanism is implemented, monitoring the validation loss during training and stopping the process if no improvement is observed for three consecutive epochs. This structured approach ensures the model is both efficient and robust, capable of achieving high accuracy without overfitting. The overall specification of the proposed optimized hybrid model is summarized in Table 3.

Table 3. Specifications of optimized proposed hybrid model

Specification	Proposed Hybrid Model
Model Type	Sequential
Number of Layers	3
Layer 1	Bidirectional LSTM (100 units, return sequences)
Layer 2	GRU (50 units, no return sequences)
Layer 3	Dense (1 unit)
Loss Function	Mean Squared Error
Optimizer	Adam
Call-backs	Early Stopping (monitoring 'val_loss', patience=3, verbose=1)
Max Training Epochs	50
Batch Size	16
Validation Split	0.1
Shuffle	False
Verbose	1
Learning Rate	0.001

4. Results and Discussions

4.1. Details of Implementation

The models were developed on the Google Colab platform using a T4 GPU hardware accelerator and the Keras library. Leveraging the default implementation of deep neural networks in Keras provided significant flexibility for customizing network architecture. This adaptability facilitated the integration of diverse layers, activation functions, and optimization techniques.

4.2. Performance Evaluation Plots

4.2.1. Train & Loss Plot

Throughout the training of the model, losses for each input were meticulously recorded. Figure 2 illustrates the training and loss curve for the proposed hybrid network. In line with the experimental setup, early stopping was employed with a maximum of 50 epochs, and it was determined that 14 epochs were sufficient to identify the optimal model snapshot.

The strategic choice of this maximum span, validated by the early stopping mechanism, underscores a commitment to efficiency without compromising the model's Learning and convergence. This disciplined approach to training, even within a constrained epoch range, contributes to the overall stability and effectiveness of the proposed hybrid network in capturing intricate patterns and relationships within the data.

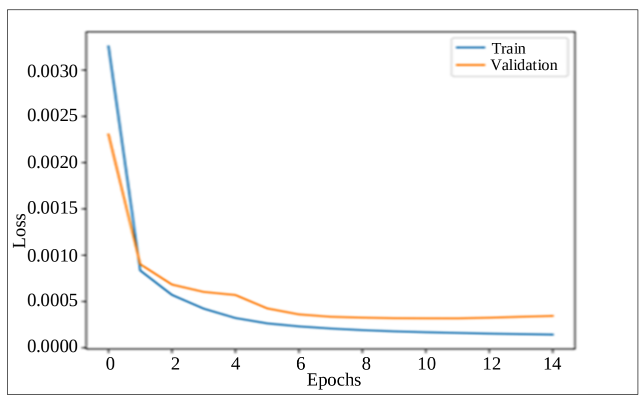

Fig. 2 Train & loss curve

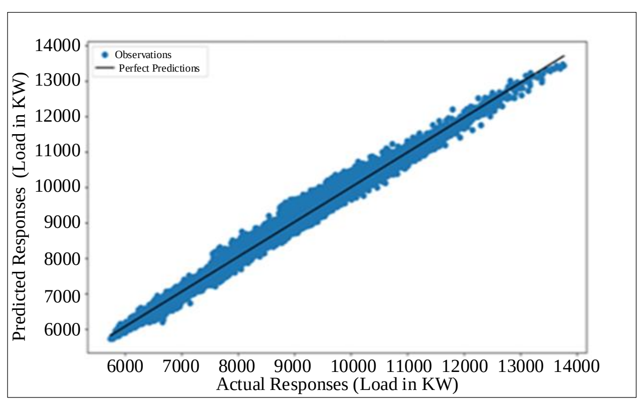

Fig. 3 R-squared plot of proposed hybrid model

4.2.2. R-Squared Plot

This plot serves as a tool to assess the effectiveness of the trained model, providing insights into how accurately the regression model predicts various response values. This procedure involves graphing the model's anticipated outcomes against the observed outcomes. In an optimal situation, an impeccable regression model would produce predicted responses that exactly match the true responses, causing all points to align along a diagonal line, as shown in Figure 3.

The assessment of various regression models relies on their R^2 values, which reflect the variability of the data. Specifically, the Bi-LSTM model exhibits an R^2 value of 0.9842, the GRU model attains 0.9814, the Conv1D model achieves 0.9686, the LSTM model reaches 0.9800, the ConvLSTM-GRU model records 0.9846, and the CNN-LSTM model registers 0.9844.

Notably, the proposed hybrid model surpasses all others, boasting the highest R^2 value of 0.9899, indicative of its remarkably close predictions to actual values. Consequently, the proposed hybrid model stands out as superior based on R^2 values, providing the optimal fit to the data. It is crucial to recognize that while a higher R^2 value generally suggests a better fit, it may not always represent the superiority of a model. Therefore, for a more comprehensive evaluation of the proposed model's performance, additional validation metrics such as MAE, MSE, RMSE, and MAPE, as shown in Table 4, have been considered.

4.2.3. Responses Plot

The responses plot provides a visual contrast between the predicted and actual responses, serving as a tool for assessing model effectiveness. A seamless alignment between the model's predicted response and the actual response indicates a high level of accuracy. In this instance, the proposed hybrid model and other benchmark models are assessed for their performance using response plots in the Figures 4(a)-(f). The response plot clearly depicts the superior predictive capabilities of the proposed model when forecasting responses on specific days. The deliberate choice of these specific days for evaluating the model stems from the intention to gauge its effectiveness in capturing diverse seasonal patterns and addressing potential challenges encountered throughout different months.

The selection of representative days spanning various seasons serves as a strategic approach to provide a thorough comprehension of the model's robustness and accuracy in Short-Term Load Forecasting across a spectrum of temporal contexts. This targeted analysis allows for an in-depth examination of how well the proposed model adapts to fluctuations in demand, providing valuable insights into its reliability and performance under diverse conditions.

4.3. Validation Metrics

The evaluation of the developed models is carried out using various metrics such as r-square error, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE) metrics and Mean Absolute Percentage Error (MAPE), as outlined in [5]. The criteria used to evaluate the proposed model using the validation metric are discussed below:

- R-squared, ranging from 0 to 1, indicates a strong model when closer to 1.
- A good model is characterized by a smaller Root Mean Square Error (RMSE), which is always positive.
- A smaller Mean Absolute Error (MAE) value signifies a successful model.
- Mean Squared Error (MSE) should be as low as possible for a successful model.
- MAPE error percentage having values very near to zero indicates a high level of accuracy.

Table 4. Short term load forecasting models performance validation

Serial No.	Regression Models	R-Square	MAE (kW)	RMSE (kW)	MSE (kW^2)	MAPE (%)
1	Bi-LSTM	0.9842	137.16	172.72	29347.24	1.59
2	GRU	0.9814	147.16	187.28	35077.21	1.75
3	Conv1D	0.9686	184.89	243.38	59238.06	2.15
4	LSTM	0.9800	146.38	193.85	37581.68	1.69
5	ConvLSTM-GRU	0.9846	131.34	169.12	28602.90	1.54
6	CNN-LSTM	0.9844	133.75	169.89	29013.90	1.56
7	Proposed Hybrid Model	0.9899	107.09	137.95	19032.79	1.23

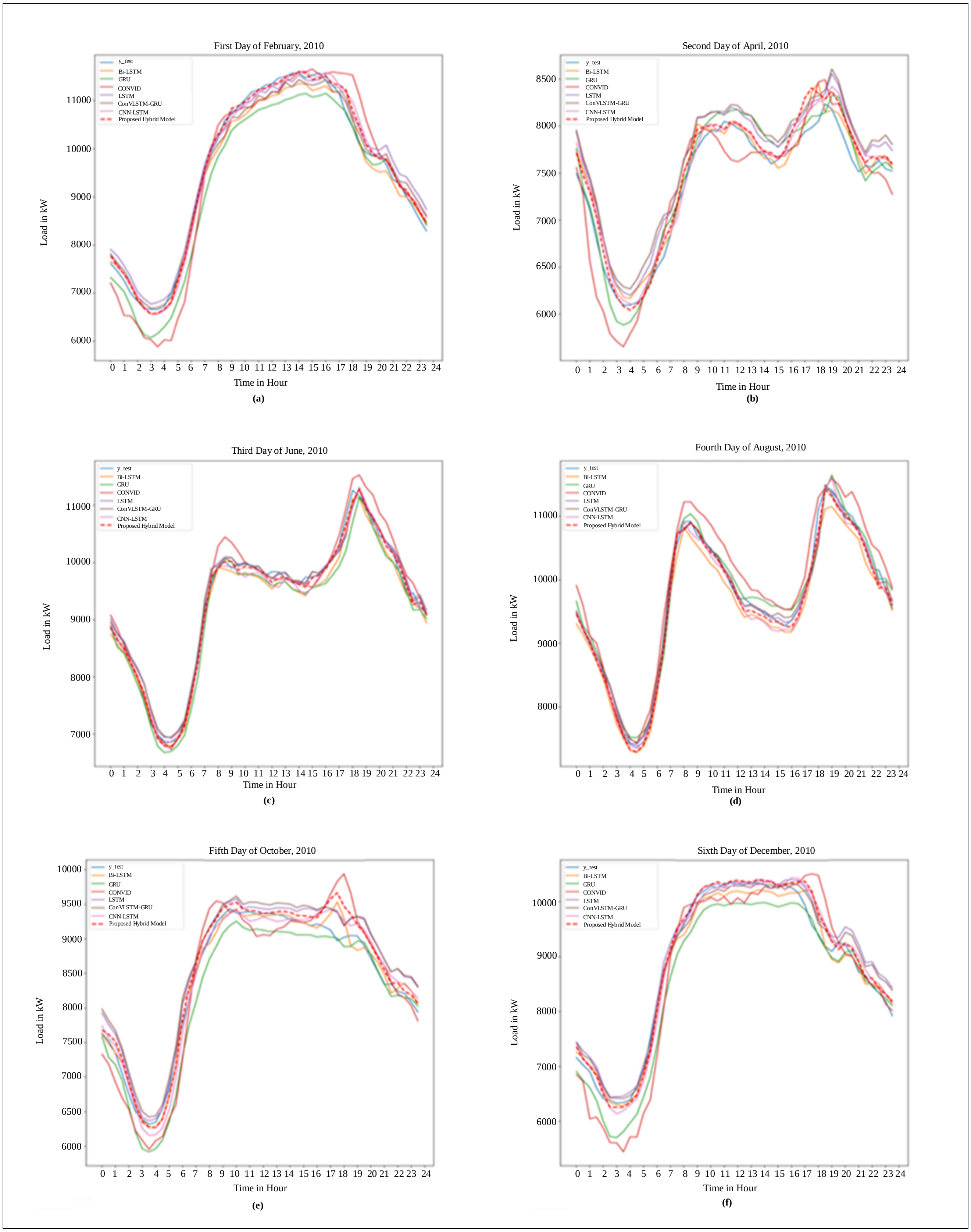

Fig. 4 Response plot of a) First day of Feb 2010, b) Second day of April 2010, c) Third day of June 2010, d) Fourth day of Aug 2010, e) Fifth day of Oct 2010, and f) Sixth day of Dec 2010.

The above results collectively highlight the strong performances exhibited by various Deep Neural Network (DNN) models. However, the proposed hybrid model stands out prominently, showcasing exceptional capabilities in comparison to its counterparts. Notably, the response plot visually reinforces this distinction, revealing a near-perfect alignment between predicted and actual values. This exceptional accuracy suggests that the proposed model excels at capturing the intricacies of the forecasting task.

The remarkable overlap in the response plot underscores the model's ability to match its predictions with actual load values closely. This near-perfect correspondence not only accentuates the precision of the proposed model but also signifies its capability to navigate the complexities inherent in load forecasting accurately. The distinct performance of the proposed hybrid model, as illustrated by both quantitative metrics and visual representation, positions it as a superior choice for precise and reliable load prediction compared to other DNN models evaluated in the study.

4.4. Discussion on Findings and Interpretation of Results

This analysis provides an in-depth evaluation of the empirical performance enhancements achieved by the proposed model compared to other benchmark algorithms across various validation metrics, as based on Figure 5.

The findings consistently highlight the superiority of the proposed hybrid model in augmenting predictive accuracy. The hybrid demonstrates a statistically significant improvement relative to the standalone models across all metrics considered in the study. Comparisons with the ConvLSTM-GRU and CNN-LSTM hybrid models reveal considerable yet moderate enhancements across all metrics.

These findings underscore the superior predictive performance of the proposed model, affirming its potential for precise and reliable short-term load forecasting. This outcome serves as a testament to the efficacy of integrating a hybrid approach that capitalizes on the strengths of diverse algorithms, thereby advancing forecasting methodologies.

As discussed earlier, even a modest 1% decrease in load forecasting error can result in a remarkable annual operating cost reduction of £10 million. In light of this financial implication, the findings further emphasize the critical importance of the proposed hybrid model, which consistently outperforms baseline architectures across all metrics. Collectively, these empirical results underscore the efficacy of the proposed hybrid model, positioning it as a promising advancement over existing architectures in the domain under consideration.

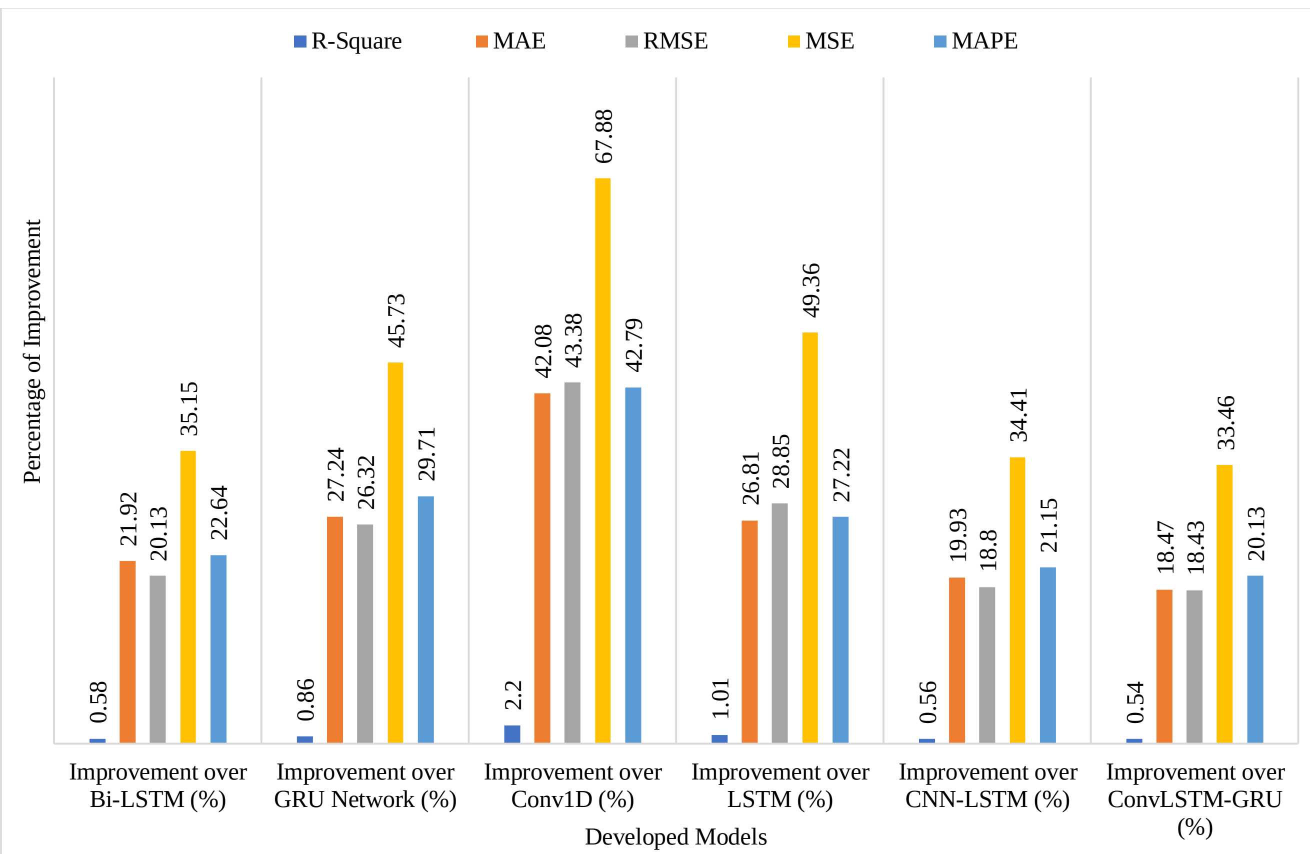

Fig. 5 Percentage of improvement gained by the proposed hybrid model compared to others in terms of validation metrics

5. Conclusion

This study introduces a groundbreaking approach to short-term load forecasting by integrating Bi-LSTM and GRU networks within a hybrid model. The synergistic combination of these algorithms demonstrates substantial improvements in forecast precision, highlighting the potential of hybrid deep learning methodologies in advancing Short-Term Load Forecasting (STLF) techniques.

To further enhance the proposed model's capabilities, future refinements could involve the inclusion of additional meteorological inputs, such as wind speed and precipitation, to boost overall performance. Moreover, dynamic adaptation mechanisms should be integrated to accommodate the inherent variability in load patterns arising from both random and planned events. Advanced feature selection techniques, such as Minimum Redundancy Maximum Relevance (MRMR), R-Relief, and Random Forest, could be employed to contribute to ongoing improvements in the model's accuracy.

The outcomes of this research significantly contribute to the development of more reliable and efficient short-term forecasting techniques for smart distribution system planning. It is crucial to acknowledge that while the proposed model demonstrates an overall enhancement, the extent of this improvement may vary across different datasets and problem domains.

Moving forward, the intention is to deploy the proposed hybrid model for short-term load forecasting in the context of creating a smart and sustainable university campus. The campus, functioning as its power distribution system, will benefit from the model's capabilities in optimizing distribution planning processes.

This application not only serves as a practical extension of the research findings but also aligns with the broader goal of integrating advanced artificial intelligence into power systems to create more robust and accurate load forecasting models.

References

[1] Turan Gonen, *Electric Power Distribution System Engineering*, 2nd ed., CRC Press, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[2] John Sharp, "Comparative Models for Electrical Load Forecasting, D.H. Bunn and E.D. Farmer, eds. (Wiley, New York, 1985) £24.95, pp. 232," *International Journal of Forecasting*, vol. 2, no. 2, pp. 241-242, 1986. [CrossRef] [Google Scholar] [Publisher Link]

[3] A.E. Okoye, and T.C. Madueme, "A Theoretical Framework for Enhanced Forecasting of Electrical Loads," *International Journal of Scientific and Research Publications*, vol. 6, no. 6, pp. 554-560, 2016. [Google Scholar] [Publisher Link]

[4] Moises Cordeiro Costas et al., "Load Forecasting with Machine Learning and Deep Learning Methods," *Applied Sciences*, vol. 13, no. 13, pp. 1-25, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Mithun Madhukumar et al., "Regression Model-Based Short-Term Load Forecasting for University Campus Load," *IEEE Access*, vol. 10, pp. 8891-8905, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Nadjib Mohamed Mehdi Bendaoud, and Nadir Farah, "Using Deep Learning for Short-Term Load Forecasting," *Neural Computing and Applications*, vol. 32, pp. 15029-15041, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shahzad Muzaffar, and Afshin Afshari, "Short-Term Load Forecasts Using LSTM Networks," *Energy Procedia*, vol. 158, pp. 2922-2927, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Changchun Cai et al., "Short-Term Load Forecasting Based on Deep Learning Bidirectional LSTM Neural Network," *Applied Sciences*, vol. 11, no. 17, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Behnam Farsi et al., "On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach," *IEEE Access*, vol. 9, pp. 31191-31212, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Dabeeruddin Syed et al., "Deep Learning-Based Short-Term Load Forecasting Approach in Smart Grid with Clustering and Consumption Pattern Recognition," *IEEE Access*, vol. 9, pp. 54992-55008, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Shafiul Hasan Rafi et al., "A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network," *IEEE Access*, vol. 9, pp. 32436-32448, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Tasarruf Bashir et al., "Short-Term Electricity Load Forecasting Using Hybrid Prophet-LSTM Model Optimized by BPNN," *Energy Reports*, vol. 8, pp. 1678-1686, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Hui Hwang Goh et al., "Multi-Convolution Feature Extraction and Recurrent Neural Network Dependent Model for Short-Term Load Forecasting," *IEEE Access*, vol. 9, pp. 118528-118540, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Manh-Hai Pham, Minh-Ngoc Nguyen, and Yuan-Kang Wu, "A Novel Short-Term Load Forecasting Method by Combining the Deep Learning with Singular Spectrum Analysis," *IEEE Access*, vol. 9, pp. 73736-73746, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Yuhong Xie, Yuzuru Ueda, and Masakazu Sugiyama, "A Two-Stage Short-Term Load Forecasting Method Using Long Short-Term Memory and Multilayer Perceptron," *Energies*, vol. 14, no. 18, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Mustaqeem, Muhammad Ishaq, and Soonil Kwon, "Short-Term Energy Forecasting Framework Using an Ensemble Deep Learning Approach," *IEEE Access*, vol. 9, pp. 94262-94271, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Fath U. Min Ullah et al., "Deep Learning-Assisted Short-Term Power Load Forecasting Using Deep Convolutional LSTM and Stacked GRU," *Complexity*, vol. 2022, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] B.U. Islam, and S.F. Ahmed, "Short-Term Electrical Load Demand Forecasting Based on LSTM and RNN Deep Neural Networks," *Mathematical Problems in Engineering*, vol. 2022, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Changchun Cai et al., "Short-Term Electrical Load Forecasting Based on VMD and GRU-TCN Hybrid Network," *Applied Sciences*, vol. 12, no. 13, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Siva Sankari Subbiah, and Jayakumar Chinnappan, "Deep Learning Based Short Term Load Forecasting with Hybrid Feature Selection," *Electric Power Systems Research*, vol. 210, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Shiyun Zhang et al., "A CNN and LSTM-Based Multi-Task Learning Architecture for Short and Medium-Term Electricity Load Forecasting," *Electric Power Systems Research*, vol. 222, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Azfar Inteha et al., "A Data-Driven Approach for Day Ahead Short Term Load Forecasting," *IEEE Access*, vol. 10, pp. 84227-84243, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Haihong Bian et al., "Load Forecasting of Hybrid Deep Learning Model Considering Accumulated Temperature Effect," *Energy Reports*, vol. 8, pp. 205-215, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Abdullah Alrasheedi, and Abdulaziz Almalaq, "Hybrid Deep Learning Applied on Saudi Smart Grids for Short-Term Load Forecasting," *Mathematics*, vol. 10, no. 15, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Siva Sankari Subbiah, and Paramasivan Senthil Kumar, "Deep Learning Based Load Forecasting with Decomposition and Feature Selection Techniques," *Journal of Scientific & Industrial Research*, vol. 81, no. 5, pp. 505-517, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Xinfang Chen et al., "Short-Term Load Forecasting and Associated Weather Variables Prediction Using ResNet-LSTM Based Deep Learning," *IEEE Access*, vol. 11, pp. 5393-5405, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] Irshad Ullah et al., "Multi-Horizon Short-Term Load Forecasting Using a Hybrid of LSTM and Modified Split Convolution," *PeerJ Computer Science*, vol. 9, pp. 1-27, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Hongsheng Xu et al., "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. [CrossRef] [Google Scholar] [Publisher Link]

[29] Ibrahim Anwar Ibrahim, and M.J. Hossain, "Short-Term Multivariate Time Series Load Data Forecasting at a Low-Voltage Level Using Optimized Deep-Ensemble Learning-Based Models," *Energy Conversion and Management*, vol. 296, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] Fachrizal Aksan et al., "Load Forecasting for the Laser Metal Processing Industry Using VMD and Hybrid Deep Learning Models," *Energies*, vol. 16, no. 14, pp. 1-24, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[31] David Willingham, Electricity Load Forecasting for the Australian Market Case Study, Mathworks. [Online]. Available: https://in.mathworks.com/matlabcentral/fileexchange/31877-electricity-load-forecasting-for-the-australian-market-case-study

[32] G. Gross, and F.D. Galiana, "Short-Term Load Forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558-1573, 1987. [CrossRef] [Google Scholar] [Publisher Link]

[33] Sepp Hochreiter, and Jurgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[34] Kyunghyun Cho et al., "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches," *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, USA, pp. 103-111, 2014. [CrossRef] [Google Scholar] [Publisher Link]