Electric Load Forecasting using Machine Learning for Peak Demand Management in Smart Grids

Sajeh Zairi
Department of Management Information System, College of Business and Economics, Qassim University, Qassim, Saudi Arabia
s.zairi@qu.edu.sa

Mushira Freihat
Department of Management Information System, College of Business and Economics, Qassim University, Qassim, Saudi Arabia
mm.freiat@qu.edu.sa (corresponding author)

Received: 22 February 2025 | Revised: 25 March 2025 and 5 April 2025 | Accepted: 7 April 2025

 DOI: https://doi.org/10.48084/etasr.10687

ABSTRACT

This paper presents an innovative hybrid Deep Learning (DL) approach to address the challenges of electric load forecasting in smart grids, particularly focusing on scenarios involving missing or noisy input data. The study utilizes real-world hourly load data from Qassim city, Saudi Arabia, to develop and validate the proposed model. The hybrid approach combines a Convolutional Neural Network (CNN) with various sequence-learning models, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional Long Short-Term Memory (BiLSTM), to capture both spatial and temporal dependencies in the dataset. A multi-step preprocessing pipeline ensures data normalization and outlier handling, whereas advanced DL architectures extract and analyze features to improve prediction accuracy. The experimental results demonstrate the hybrid CNN-GRU model's superior performance, achieving lower error rates (e.g., RMSE, NRMSE, and MAPE) compared to traditional methods and standalone DL algorithms. This work underscores the potential of hybrid DL models for enhancing load forecasting accuracy, providing critical insights for optimizing smart grid operations and ensuring sustainable energy management.

Keywords-energy consumption; machine-learning; electrical consumption rationalization

I. INTRODUCTION

The trend of industrialization is increasing at a tremendous rate, and concurrently, the global population and economic activity are on the rise, which has the potential to raise energy demand steeply in the coming decades. Electricity, a pivotal form of energy, can be produced from a variety of sources, including thermal energy, solar energy, hydropower, nuclear energy, wind energy, and fossil fuels. The dynamics of population growth, economic development, and the resulting surge in electricity demand impose considerable strain on energy production systems. The management of electricity is a complex process that encompasses the generation, transmission, and distribution of electricity in order to ensure consistency in power supply [1, 2].

Electric utilities are defined as transmission systems that deliver electricity from generating plants to consumers, comprising a collection of networks. Key elements of this system include power stations that produce electricity, voltage upgrading stations that adjust voltage in accordance with consumer demand, and transmission and distribution networks that facilitate energy supply to end users. However, traditional electrical systems are experiencing challenges, such as higher utilization, which leads in high operating costs and deterioration in power quality. The following problems can be identified in this context, necessitating the construction of new plants: conventional grids lack the capacity to accurately predict power failures, estimate their consequences, or plan resource distribution and energy storage [3].

It has also been observed that contemporary electrical networks are virtually the same as those utilized over the past century, despite increasing demands due to population growth. Traditional systems have multiple drawbacks, such as low transparency, slow mechanical switching devices, poor supervision, and ineffective control of the energy flows. Conventional electric grids are characterized by one-way transmission and are vulnerable to climate variations, fossil fuel consumption, increased population density, failure of a particular component, and low capacity of energy storage. Therefore, a transition to novel grid technologies is necessary to address these challenges [4].

The concept of the smart grid is a more advanced development concept that mirrors the upgrade of the conventional electrical system. The smart grid is an enhanced form of networks that incorporates data control and management of devices, innovative technologies, information communication equipment, and field-based devices. all coordinated to enable network physical control. These improvements facilitate the observation of grid processes, the transmission of information to control centers, and the grid's capacity for adaptation to change. Energy management and monitoring, fault diagnosis, load profiling, and security improvements are feasible for smart grids, which rely on the acquisition and information analysis of large amounts of high-dimensional data [3]. Accurate forecasting of electrical load can effectively establish a power system capable of delivering an uninterrupted power supply, as illustrated in Figure 1.

Smart grids allow energy providers and consumers to communicate and share a significant amount of information, thereby increasing the reliability and safety of electricity. This interactive system has been developed to empower consumers to make the right decisions about their energy consumption in order to reduce electricity costs and mitigate the impact of natural disasters and cyber-attacks. Smart grids are also used to support the principles of sustainable development through the use of renewable energy, energy saving, and decrease of greenhouse gas emission. With the help of automated fault detection and analysis, smart grids also provide a better response than conventional power systems.

Smart grid solution elements include communication infrastructures, smart sensors, and interfaces of residential, commercial, and industrial use. Wireless and wired communication technologies are employed to facilitate the transfer of data between smart meters and other electronic devices. The main disadvantages of wireless systems include financial costs and signal strength in the remote areas, as well as the dependency on battery power. In contrast, wired technologies provide reliable, battery-free connections. On the other hand, wireless technologies include communication technologies based on specific application needs.

Smart grids offer numerous benefits, including rate management, real-time analysis of electricity consumption, rapid power supply restoration, and efficient energy control. They also support the Internet of Things (IoT), facilitating the development of smart homes, buildings, and cities, as well as cloud computing for energy and data management. Nevertheless, there are still obstacles to overcome, including technical and social challenges, such as security risks, backward storage, a lack of legislation, and high capital costs. In response to these challenges, the incorporation of Machine Learning (ML) algorithms for grid forecasting and management emerges as a promising solution. For example, ML techniques have the capacity to collect data from power generation facilities, distribution centers, and any associated facilities such as industries and housing facilities. The utilization of ML techniques facilitates the development of smart grids, which are capable of predicting consumer behavior, optimizing expenses, and reducing emissions [5, 6].

This study identifies the critical issue of missing input data due to sensor or system failures in smart grids, which impacts the reliability of load forecasting. The research seeks to bridge this gap by proposing a robust forecasting model capable of handling such challenges. Furthermore, ensuring an uninterrupted and stable power supply is a fundamental objective for modern energy systems. Consequently, the research is driven by the need to improve grid stability through precise load forecasting, which is vital for the effective management and optimization of smart grids. In addition, the study emphasizes the potential of advanced ML techniques, such as neural networks, to enhance the prediction accuracy and reliability of load forecasting in smart grids, addressing the shortcomings of traditional methods.

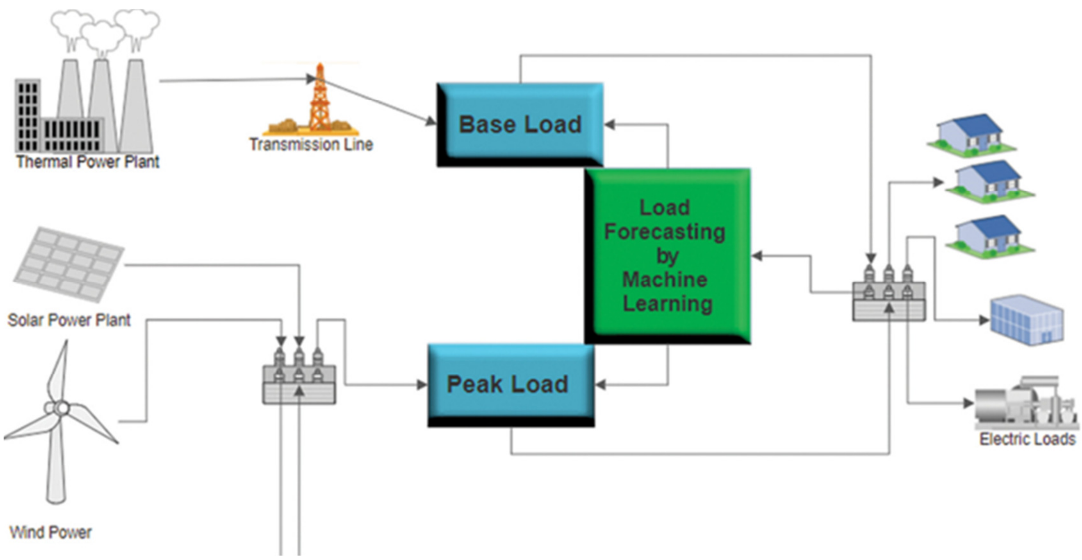


Fig. 1. Plan for managing peak load control.

A central focus of this study is the utilization of contemporary forecasting methods, as they are better suited for today's complex energy systems. Traditional electric load forecasting approaches, such as ARIMA and regression, depend on statistical models that assume fixed patterns but struggle with handling complex variations. In contrast, modern techniques leverage AI and ML such as Long Short-Term Memory (LSTM) and Artificial Neural Networks (ANNs), to process big data and dynamically adapt to real-time fluctuations. As a result, traditional models tend to be rigid and less accurate, whereas modern approaches offer greater flexibility and precision. Therefore, modern forecasting methods are more effective for managing evolving energy demands and smart grids.

The primary challenges we address include handling missing or noisy data, capturing both spatial and temporal dependencies, and improving forecasting accuracy with limited data. Many smart grid forecasting models struggle with incomplete or unreliable data due to sensor malfunctions or communication failures. To mitigate this, our model incorporates preprocessing techniques and Deep Learning (DL) patterns and long-term dependencies in load data. Our hybrid Convolutional Neural Network-Recurrent Neural Network (CNN-RNN) approach improves prediction accuracy by leveraging convolutional feature extraction and recurrent sequence modeling. Furthermore, many DL models require large datasets for optimal performance, which can be a limitation in real-world scenarios. Our approach is optimized to provide reliable predictions even with constrained training data, making it practical for smart grid applications.

II. RELATED WORK

Load forecasting is a distinct problem between conventional and smart grids due to their different operational characteristics. Traditional grid systems rely on basic statistical analysis combined with a study of historical energy consumption data to predict energy demand. These forecasting models have difficulty adapting to changing energy consumption patterns because they were designed to analyze static datasets. In addition, traditional grids offer limited real-time management capabilities because they lack the monitoring and response systems to effectively integrate renewable energy and control demand response programs. Load forecasting in smart grids utilizes advanced technology through sensors and smart meters with communication networks to collect information on electricity consumption, system generation, and conditions. Complex energy consumption trends become visible through ML and DL algorithm processing of the collected data. The detailed power consumption forecasts of smart grids lead to enhanced power grid management through improved optimization.

Additionally, smart grids support dynamic pricing strategies and demand-response programs, encouraging consumers to adjust their energy usage based on real-time grid conditions. The dynamic, adaptive, and data-driven nature of load forecasting in smart grids distinguishes them from conventional systems, enabling improved efficiency, reliability, and sustainability in energy management. Several studies focusing on this topic are summarized in the remainder of this section.

Among all the different ML load forecasting methods, ANNs have advantages as they can learn features on data for precise regression, making them one of the most widely used for load forecasting [7, 8]. In the preliminary study of the current work, authors in [8] proposed a method to construct electric load forecast based on ML and evaluate its effectiveness with environmental data compared to other ML methods. Their study compared models including Support Vector Machines (SVMs), Multiple Linear Regression (MLR), a neuro-fuzzy inference system, neuro-ANNs for both long-term and short-term load forecasting. The study showed that the accuracy of the existing models of the power usage was high; SVM responded best to the long-term basis, whereas ANN responded to the short-term basis. A similar comparison by the authors in [9] explored genetic algorithms, SVM, and Deep Neural Networks (DNNs), showing that ANN outperformed these models while maintaining manageable complexity.

Experts have conducted research to optimize ANN prediction models by improving input variable selection methods. Authors in [10] integrated K-Nearest Neighbors (KNN) with ANNs to develop their load prediction systems. The research employed environmental variables while clustering input data using KNN, which led to better ANN outcomes during training on preprocessed variables. In addition, they investigated which variables affect cooling load accuracy predictions in an office building setting and found that clustering input data improved prediction performance.

The smart grid sector uses both ANNs and Extreme Gradient Boosting (XGBoost) along with other ML models for load forecasting. Residents can benefit from the residential load forecasting model using XGBoost technology, which was found to be successful in preventing overfitting according to the authors in [11]. The research demonstrated that XGBoost successfully solves overfitting problems better than other predictive approaches. Authors in [12] showed that Random Forest Regression (RFR) and XGBoost achieved the lowest computational complexity levels, followed by Support Vector Regression (SVR) and ANN. Flexible Neural Trees (FNTs) implemented a complex computational framework that achieved inferior results than RFR as well as XGBoost. The research by the authors in [13] developed a load forecasting system that integrated the XGBoost one-step-ahead forecaster with quantile regression. The dynamic adjustment of quantile regression model parameters through recent prediction data improved the final output results. The proposed model framework achieved the highest accuracy level and real-time prediction reliability among all tested models from other ML and DL approaches. Authors in [14] designed a cloud-based load prediction model that integrates Grey Wolf Optimizer (GWO) optimization and ELM machine through an improved integration technique. The exploitation and optimization performed by GWO on ELM parameters led to an upgraded model that outperformed its base version.

Author in [15] designed an ensemble load forecasting method for one-day-ahead predictions by integrating data mining techniques for training, feature selection, and elimination. The study utilized eight different ML models to enhance the ensemble approach and concluded that the proposed comprehensive framework outperformed the individual models and delivered improved forecasting accuracy.

Authors in [16] discuss significant advancements in household load forecasting by introducing a DL model enhanced through hyperparameter optimization. Their main contributions include the development of a robust forecasting framework that improves the accuracy of predicting household energy consumption. The study emphasizes the effectiveness of DL techniques in capturing complex patterns in load data, whereas also highlighting the importance of optimizing hyperparameters to achieve better model performance. Overall, it provides valuable insights and methodologies for improving energy management in smart grids. Authors in [17] present a novel approach to electricity load forecasting by integrating Discrete Wavelet Transform (DWT) with LSTM networks. Their main contributions include demonstrating how DWT effectively decomposes load data into different frequency components, thereby enhancing the LSTM's ability to capture temporal dependencies. This combination significantly improves forecasting accuracy compared to traditional methods. The study provides a comprehensive framework that shows the benefits of using advanced DL techniques for more precise electricity load predictions. Authors in [18] introduce an adaptive deep-learning framework for load forecasting that combines transformer models with domain knowledge. Their main contributions include the development of a hybrid approach that leverages the strengths of transformers in capturing complex temporal patterns while incorporating expert insights for enhanced contextual understanding. The study highlights the effectiveness of combining advanced ML techniques with domain expertise to optimize energy management and forecasting in smart grids. Authors in [19] present a DL-driven hybrid model for short-term load forecasting, emphasizing its application in smart grid information management. Their main contributions include the integration of multiple DL techniques to enhance forecasting accuracy and the ability to effectively manage and analyze data within smart grid systems. The model demonstrates improved performance in predicting short-term energy demand, while facilitating better decision-making and resource allocation in smart grids. Overall, the study underscores the potential of hybrid DL approaches to optimize load forecasting and enhance smart grid operations. Furthermore, a comprehensive overview of the recent developments in DL methods applied to short-term energy load forecasting may be found in [20].

A. Discussion of the Literature Review in Relation to the Goal of the Paper

The literature review provides a comprehensive analysis of existing load forecasting methods and their limitations, forming the basis for the goals of this study. Traditional forecasting approaches, often reliant on basic statistical models and historical data, fail to capture the dynamic and non-linear patterns of modern energy consumption. These methods also lack the robustness required for real-time applications, particularly in scenarios with missing data, which is a critical challenge in smart grids. ML-based methods, including ANNs, SVMs, and gradient boosting techniques, have been extensively applied in load forecasting. ANNs, for instance, are praised for their ability to extract features and perform accurate regression. However, their performance can vary depending on the quality of the input data and the feature selection processes. Studies integrating ANNs with techniques like KNN have demonstrated improved prediction accuracy by clustering input variables. Similarly, advanced methods like XGBoost have shown strong performance in addressing overfitting issues and achieving high prediction accuracy.

Despite these advancements, significant gaps remain in the literature, particularly in the areas of missing data handling, feature selection optimization, and model adaptability in dynamic smart grid environments. These gaps are directly aligned with the goals of this paper, which aims to address the issue of missing input data by developing a robust neural network-based forecasting model. By leveraging ML techniques and applying methods such as Principal Component Analysis (PCA) for feature extraction, the proposed approach seeks to improve grid stability and prediction accuracy.

In summary, the literature review underscores the need for innovative forecasting models that can address real-world challenges, such as missing data and dynamic energy consumption patterns. This is consistent with the objective of this study to develop and validate a model that addresses these limitations while advancing the capabilities of smart grid systems.

B. Machine Learning in Power Systems

Electrical loads are categorized into two types: base load, which remains relatively constant, and peak load, which fluctuates. Predicting the peak load using ML techniques can improve the stability of the power system. This study explores ML approaches and discusses existing methods based on their models, which are classified as linear and non-linear. Linear models, such as the least-squares method, are simple and fit a best-fit line to the data. While widely used, they may produce large errors when data points are sparsely spaced. Non-linear models, though more complex and computationally intensive, provide better accuracy, especially for the highly non-linear nature of electrical loads. Examples include SVMs and Gaussian Process Regression (GPR). SVMs combined with Ant Colony (AC) optimization have proven effective for load prediction. Key terms include prediction variables (features used to train the model) and response variables (values predicted based on the features). Non-linear models are preferred when prediction variables exhibit non-linear relationships.

Load prediction can be short, medium, or long term, depending on the prediction timeframe and training data. Accurate predictions require careful feature selection to avoid redundancy and overfitting. Historical load data play a crucial role in model training as they provide the foundation for accurate predictions.

In this study, the electrical demand load is predicted for specific hours of the day based on various prediction variables, framing the task as a regression problem. Regression involves predicting a response variable based on its relationship with one or more prediction variables. It can be classified as:

- Linear regression: A single response variable and a single prediction variable.
- Multiple regression: A single response variable and multiple prediction variables.
- Multivariate regression: Multiple response variables and multiple prediction variables.

The choice between linear or polynomial regression depends on the curve used for data fitting, and the selection of the optimal model is guided by modeling and simulation results. The range of regression approaches includes KNN, neural networks, SVMs, ANNs, and GPR. Determining the correct regression technique requires evaluation of three performance aspects, including computational speed, algorithm complexity, and hyperparameter requirements. The model assessment incorporates several key metrics, including R-Square (R^2), adjusted R^2 along with RMSE, MSE and MAE.

Accurate predictions necessitate high-quality data. Noise, which is often introduced during data collection, must be filtered using statistical or mathematical methods. Large, error-free datasets with high and independent covariance between prediction and response variables are critical for effective ML models. The classification of the load forecasting techniques is presented in Figure 2.

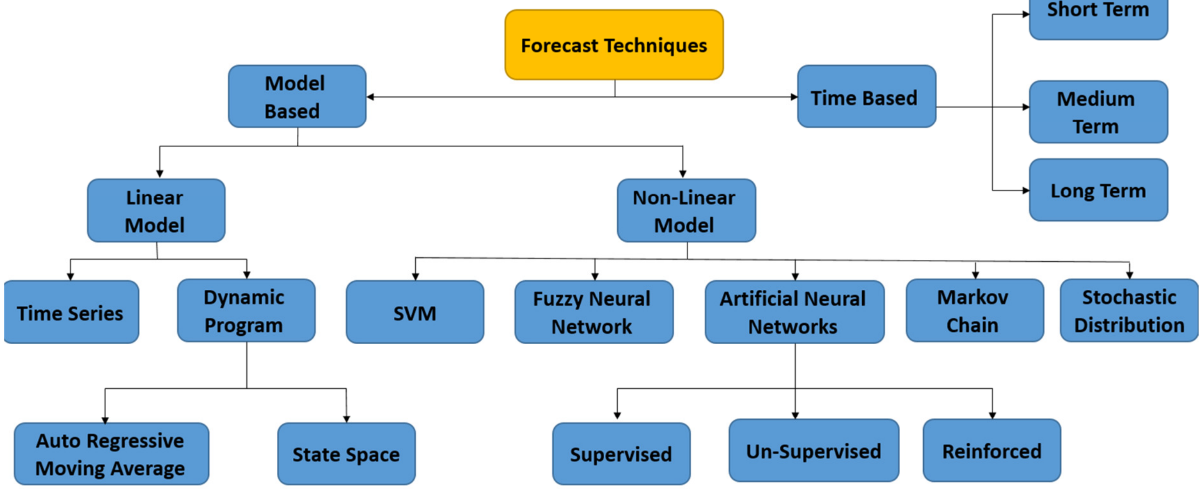


Fig. 2. Classification of load forecasting techniques.

III. METHODOLOGY

A. Forecasting Model

The present research incorporates real-world data from Saudi smart grids containing hourly load information for the city of Qassim during the year 2021. A total of 4876 records, representing two seasonal periods, are included in this dataset along with a resolution of one hour. The detailed load data for March to September 2021 shown in Figure 3 focuses on Qassim, which is considered the largest metropolitan area in Saudi Arabia in terms of population and industrial and commercial business operations. The proposed DL models utilize the 4876 records within each dataset to generate their basis for load forecasting.

The prediction model takes historical energy consumption data from Qassim smart meters as input. This model addresses data inconsistencies, missing entries, outliers, and other challenges present in energy consumption datasets and provides precise predictions for short-term energy consumption. By leveraging DL methods, optimal performance can be achieved across evaluation metrics such as RMSE, NRMSE, and MAPE. These predictive insights play a crucial role in improving Saudi Arabia's energy infrastructure by enabling the Ministry of Energy to make data-driven adjustments to physical systems based on the forecasted outcomes. As we handle private data in smart grid forecasting, prioritizing data privacy and ethical considerations is crucial. To address this challenge, the proposed model should follow several key steps. First, it is essential to anonymize user data to safeguard individual identities. Second, ensuring transparency in data collection and usage fosters consumer trust. Third, implementing strong security measures is vital to protect against data breaches. Lastly, ethical guidelines must regulate the use of AI algorithms to promote fairness and prevent bias in energy distribution.

B. Hybrid Approach

Sound prediction of customer energy consumption within smart grid environments is a complicated time-series predictive modeling task. Our approach relies on advanced preprocessing techniques with a hybrid DL framework due to the complex prediction requirements. The predictive accuracy of time-series data is highly dependent on reliable noise reduction techniques, as such datasets typically have significant noise levels. The Savitzky–Golay (SG) filter serves as a reliable noise reduction approach according to [21]. Future research will include implementation of the described technique for our studies.

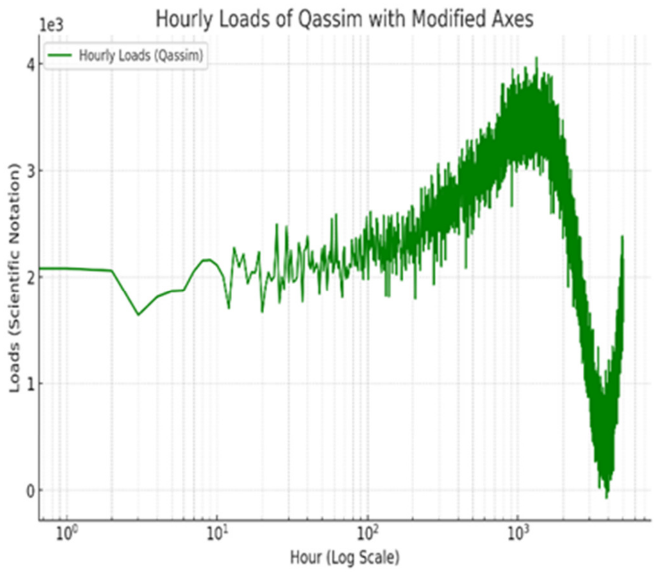


Fig. 3. Hourly load data for Qassim city (March 2021–September 2021) with a one-hour time resolution.

Preprocessing steps are essential for improving model performance and maintaining data integrity. Data normalization scaled the input features to a uniform range, promoting faster convergence during training. Data augmentation was utilized to introduce variability through synthetic data generation, which helped mitigate overfitting and enhanced the model's generalization capabilities. Additionally, missing values were addressed to prevent incomplete data from biasing the results; techniques such as mean imputation and forward filling preserved continuity in the time series. By systematically implementing these procedures, the overall quality and robustness of the dataset were significantly enhanced, resulting in more accurate load forecasts.

The methodology includes an innovative single-layer model followed by a DL hybrid framework as its secondary component. The proposed DL models utilize CNN–LSTM and CNN–Gated Recurrent Unit (GRU) and CNN–Bidirectional Long Short-Term Memory (BiLSTM) architectural components as shown in Figure 4. The feature extraction function is performed within the convolutional layer, yet the pooling layers take on the important task of minimizing feature map invariance effects. Additionally, the activation function facilitates the learning of complex patterns throughout the training process.

The study integrated a dropout layer after the CNN feature extraction block in order to feed information to the LSTM, GRU and BiLSTM layers. The dropout layer is activated to deactivate specific neurons according to a random procedure, which reduces overfitting and improves the generalization capability of the model.

A CNN design typically adopts an approach that builds up targets from broad to specific details. Using this architecture design requires a high computational complexity due to the numerous trainable parameters it contains. A kernel size of 32 served to prevent overfitting during the convolutional layer. The sequence-learning block included one layer each of LSTM, GRU, and BiLSTM models with 200 hidden units configured per layer.

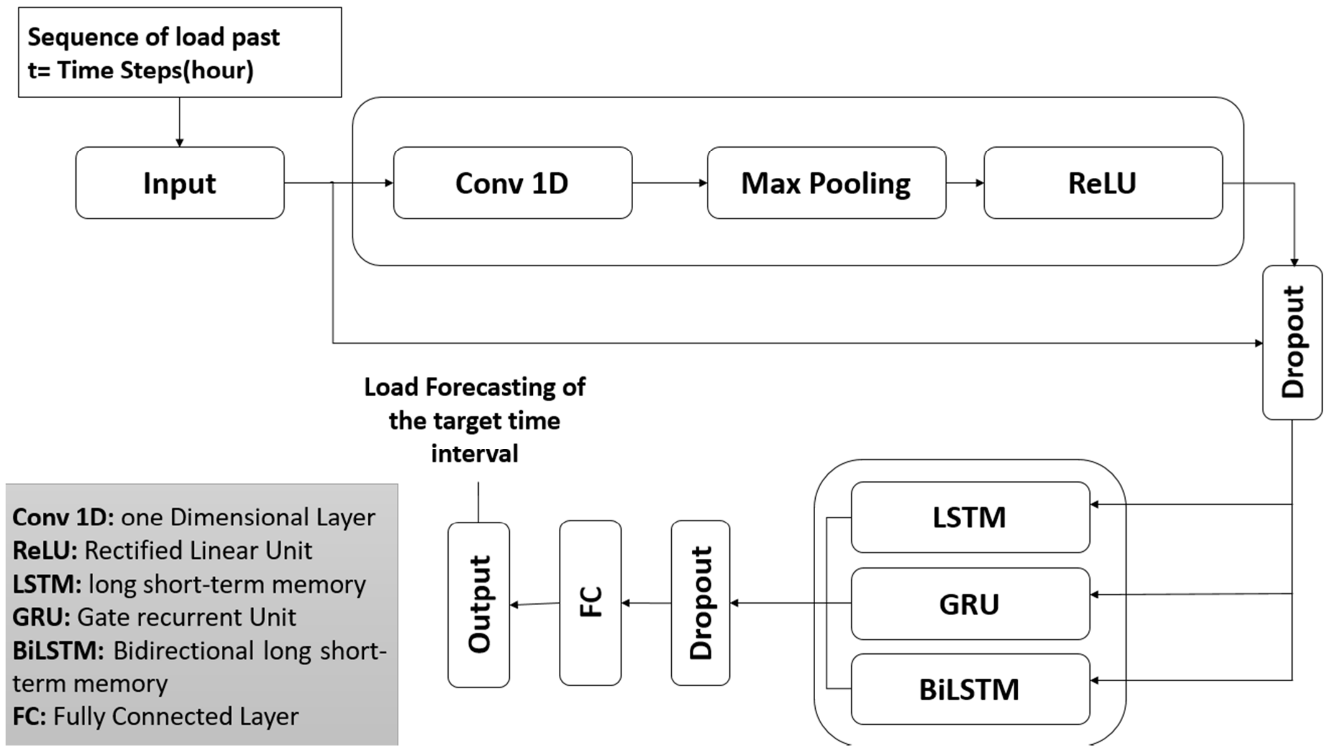


Fig. 4. Structure of the proposed hybrid deep learning algorithm, integrating 1-D CNN with multiple deep learning models.

The return sequence parameter was given a value of false in the final layer to output only the last hidden state, although the parameter value was true throughout the other sequence-learning blocks to maintain full state sequences during processing. A dropout layer was added prior to the fully connected layer as an overfitting reduction strategy. Our hyperparameter tuning was conducted using a trial-and-error approach, testing multiple configurations to identify the optimal settings. Key hyperparameters and their selection criteria included the CNN kernel size and filters, whith a kernel size of 3 and 32 filters were chosen based on empirical testing to balance feature extraction and computational efficiency. For GRU hidden units, different values (100, 200, 400) were tested, with 200 units providing the best tradeoff between model complexity and performance. A dropout rate of 0.2 was selected to reduce overfitting while maintaining model stability. The batch size and learning rate were set to 64 and 0.001, respectively, using the Adam optimizer, based on convergence behavior and validation performance. Finally, the model was trained for 2000 epochs, with early stopping applied when the validation loss did not improve for 50 consecutive epochs.

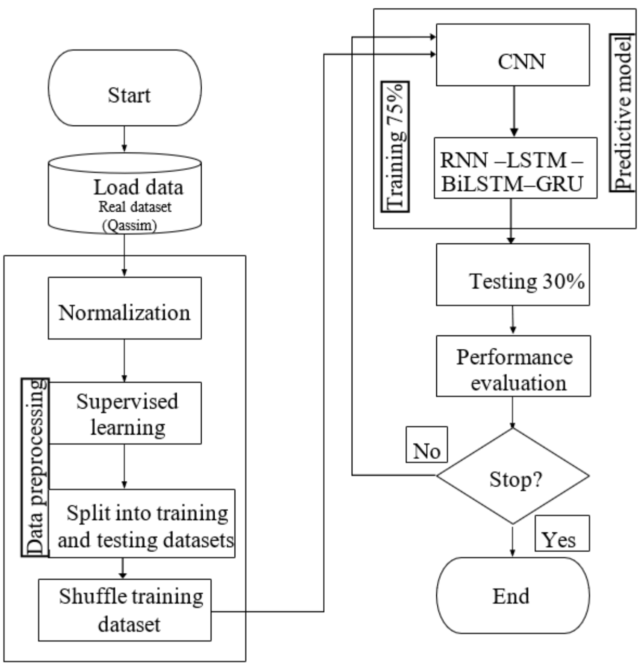


Fig. 5. Flowchart illustrating the innovative hybrid DL algorithms for enhanced load forecasting.

Figure 5 shows the sequence of operations for the proposed hybrid DL algorithm. The analysis of Qassim load authentic data proceeds by accepting information from its real-world dataset. The input dataset undergoes a preprocessing treatment that normalizes it before the researchers partition it into 75% training sets and 25% testing subsets. The predictive model consists of two parts, using CNN in the first stage and RNN-LSTM-BiLSTM-GRU in the second stage to process the training data. Every experiment implementing the hybrid DL model starts with CNN as the first layer before moving to a traditional DL algorithm in the second layer. The performance of the model is evaluated by analyzing the remaining 25% of the full dataset. The model moves toward finalization when successful predictions are made with acceptable accuracy levels; however, if the model prediction results do not meet the requirements, the researchers test hybrid DL model combinations through retraining sessions using current training data.

Table I summarizes the parameters implemented in the developed DL system. The research adopted Adam optimizer for optimization while relying on MAE as the selected loss function. The proposed model uses 1-D CNN to extract time-series features and then applies three different analysis methods through LSTM, GRU, and BiLSTM models. Feature extraction involved the model working in forward direction until the first stage of analysis, which could use LSTM or GRU or BiLSTM features. The process required backward training in the second feature analysis stage to improve the learning process.

TABLE I. HYPERPARAMETER SETTINGS OF THE PROPOSED HYBRID DEEP LEARNING MODEL

Parameter	Setting
Network architecture	LSTM, GRU, BiLSTM, RNN, ANN CNN-LSTM, CNN-GRU, CNN-BiLSTM
Optimizer	Adam
Loss function	MAE
Learning rate	{0.0005}
Adjustment	learning rate = 1×10^{-6}
Lag	1:50 h look back
TrainFcn	gradient descent momentum (traingdx)
LearnRateDropPeriod	400
LearnRateDropFactor	0.2
CNN, ANN	32 hidden units
LSTM, GRU, BiLSTM layers	400 hidden units

The two stages of 1-D CNN worked together utilizing a filter span of 5 with 32 hidden units to boost input feature recognition. Two successive layers of LSTM along with GRU or BiLSTM components using 400 hidden units existed in the model structure for feature analysis and output prediction. A ReLU activation function was used in the model structure, which was trained over 2000 epochs. The process of DL model optimization and complex stochastic process management required the optimization of network design parameters in addition to activation functions, loss functions, and hidden unit numbers and training epochs. An NP-hard optimization problem exists as part of this process.

The hyperparameters in this study were selected through a trial-and-error approach, with 400 hidden units identified as the optimal configuration after extensive experimentation. While different DL algorithms have unique computational methods, selecting the best hyperparameters for all models is crucial for efficient modeling, given the high computational cost and the need for accurate predictions.

In MATLAB environment, we applied the CNN–RNN hybrid model design, which originated from the study in [14]. The basic CNN–RNN hybridization algorithm served as a starting point to design models combining CNN with RNN as well as with LSTM, BiLSTM, and GRU components. The models were developed and tested on a MacBook Pro (13-inch

M1 2020) with 8 GB of memory partitioned into 4 GB of performance memory and 4 GB of efficiency memory. The implementation was done within the MATLAB 64-Bit (maci64) development environment using the DL toolbox [22], and the pseudocode used to select the appropriate hybrid DL model is as follows.

```
#Decision pseudocode for selecting the
appropriate hybrid DL model
data = xlsread("data.xlsx");
mu = mean(dataTrain);
sig = std(dataTrain);
if strcmp(NetOption,"CNN-RNN")
elseif strcmp(NetOption,"CNN-LSTM")
elseif strcmp(NetOption,"CNN-BiLSTM")
elseif strcmp(NetOption,"CNN-GRU")
elseif strcmp(NetOption,"RNN")
elseif strcmp(NetOption,"GRU")
elseif strcmp(NetOption,"LSTM")
elseif strcmp(NetOption,"BiLSTM")
elseif strcmp(NetOption,"FeedForwardNet")
end
YPred = sig.*YPred + mu;
YTest = sig.*YTest + mu;
```

IV. RESULTS AND DISCUSSION

Standard evaluation metrics were used to assess the accuracy of predictions derived from experimental procedures. Specifically, the performance was evaluated using RMSE along with its derivative Coefficient of Variation of RMSE (CV or NRMSE), MAE, and MAPE:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - p_i)^2}{n}} \tag{1}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\bar{y}} \times 100\% \tag{2}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{n}|y_i - p_i| \tag{3}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i - p_i}{y_i} * 100 \tag{4}$$

Performance evaluation of forecasts was performed by training on 75% of the data and testing on 25% of the data, using previously unseen input. Model evaluation was used to compare the proposed hybrid DL models against traditional DL algorithms such as ANN, RNN, LSTM, GRU, and BiLSTM. The single-model BiLSTM achieved 0.9869 R, 80.5873 RMSE, 1.526 NRMSE, and 0.9991 MAPE values. The average MAPE value of this model exceeded that of the other models. The results of CNN combined with BiLSTM showed R, RMSE, NRMSE, and MAPE values of 0.9872, 78.085, 1.4786, and 0.9497, respectively. The CNN-GRU showed the best prediction accuracy as it achieved R of 0.9875, RMSE of 77.4877, NRMSE of 1.4673, and MAPE of 0.9505. To further illustrate and clarify the performance advantages of the proposed hybrid CNN-GRU model, Table II provides a comparative summary of the prediction accuracy against several benchmark models. Table II reveals that the hybrid CNN-GRU model provides lower RMSE, NRMSE and MAPE values compared to all other ANN, BiLSTM, and CNN-BiLSTM approaches tested, thus demonstrating its excellent predictive power. This study presents CNN-GRU as the main focus because it balances operational efficiency with high accuracy levels, making it suitable for real-world smart grid applications despite the superior RMSE and MAPE performance of CNN-RNN.

TABLE II. COMPARATIVE PERFORMANCE OF THE PROPOSED MODEL WITH BENCHMARK MODELS

Model	RMSE	NRMSE	MAPE (%)
ANN	80.5873	1.526	0.9991
BiLSTM	80.5873	1.526	0.9991
CNN-BiLSTM	78.085	1.4786	0.9497
Hybrid CNN-GRU (Proposed)	77.4877	1.4673	0.95.5
CNN-RNN	20.7501	1.2227	0.7591

State-of-the-art DL methods LSTM and GRU achieved lower prediction accuracy compared to the proposed hybrid models. RNN-based methods also underperformed. With a small dataset, ANN produced superior results compared to RNN, LSTM and BiLSTM, GRU. However, the effectiveness of advanced DL models would increase with larger data input. The CNN–RNN model provided exceptional accuracy in Qassim load forecasting as it showed R of 0.9927, RMSE of 20.7501, NRMSE of 1.2227, and MAPE of 0.7591. The research supports hybrid models as the most effective option in short-term load forecasting applications due to their CNN-based configuration.

The proposed hybrid CNN–GRU model was evaluated using MAPE and RMSE measurements along with a regression plot analysis. A comparative evaluation of the forecasting performance used predicted sequences against actual observations to provide both visual and analytical assessments. The researchers examined the predicted load over different hourly periods during each iteration and Figure 6 shows the comparison results.

Furthermore, an ANN architecture was developed using MATLAB by testing different numbers of hidden neurons within the hidden layer. The network was evaluated through RMSE measurements and R statistical values within its different configurations. In the selection process, an architectural configuration was chosen that provided the highest R value and the lowest RMSE value. The optimal model with 32 hidden units was evaluated as presented in Table III.

TABLE III. UPDATED PREDICTION ERROR TABLE FOR QASSIM RESULTS

Model	R	RMSE	NRMSE (%)	MAPE (%)
ANN	0.9881	75.3124	1.4721	0.9453
RNN	0.9862	92.1653	1.7984	1.1523
LSTM	0.987	78.4567	1.5238	0.9934
GRU	0.9872	77.9981	1.5124	0.9871
BİLSTM	0.9873	76.8532	1.4947	0.9765
CNN-RNN	0.9825	94.6782	1.8215	1.2074
CNN-LSTM	0.988	74.9821	1.4623	0.9405
CNN-BiLSTM	0.9878	75.1278	1.4698	0.9423
CNN-GRU	0.9883	74.4523	1.4512	0.9392

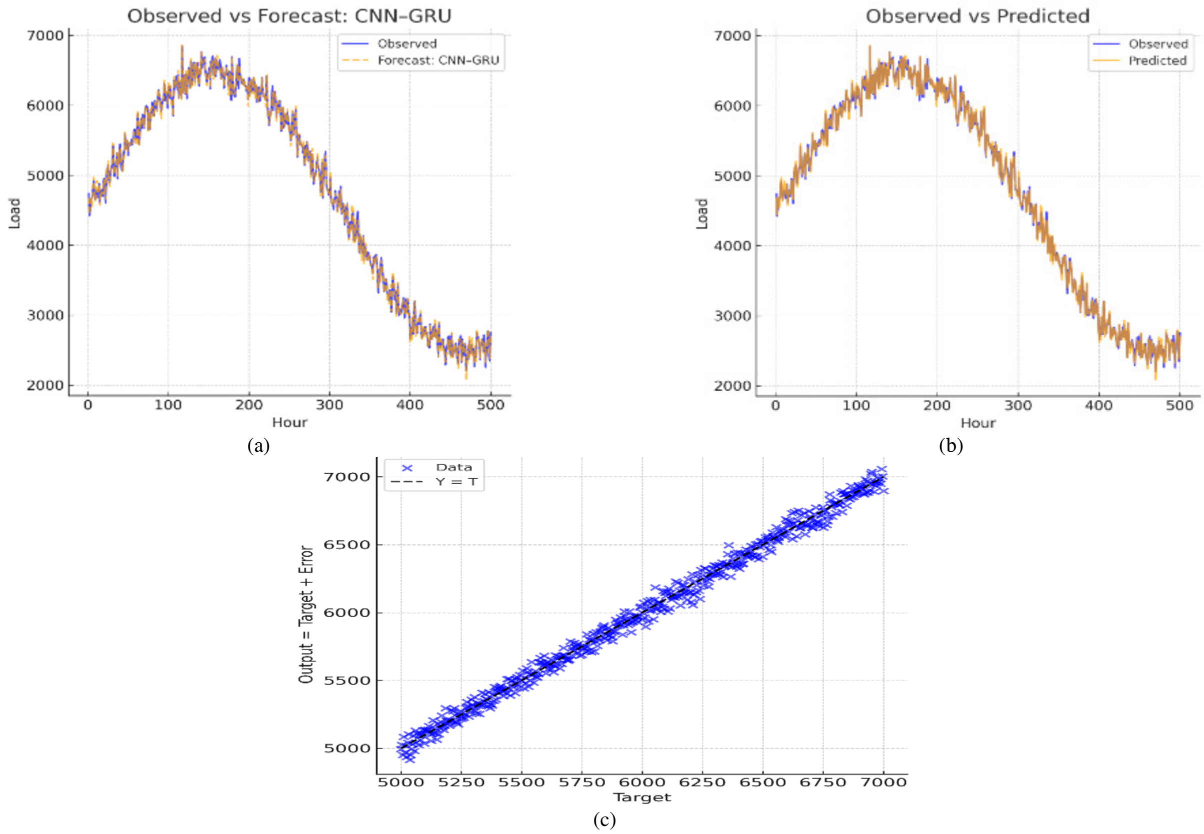


Fig. 6. Prediction performance of the hybrid CNN–GRU model for Qassim load: (a) training and testing phase results, (b) testing phase results, and (c) regression analysis plot.

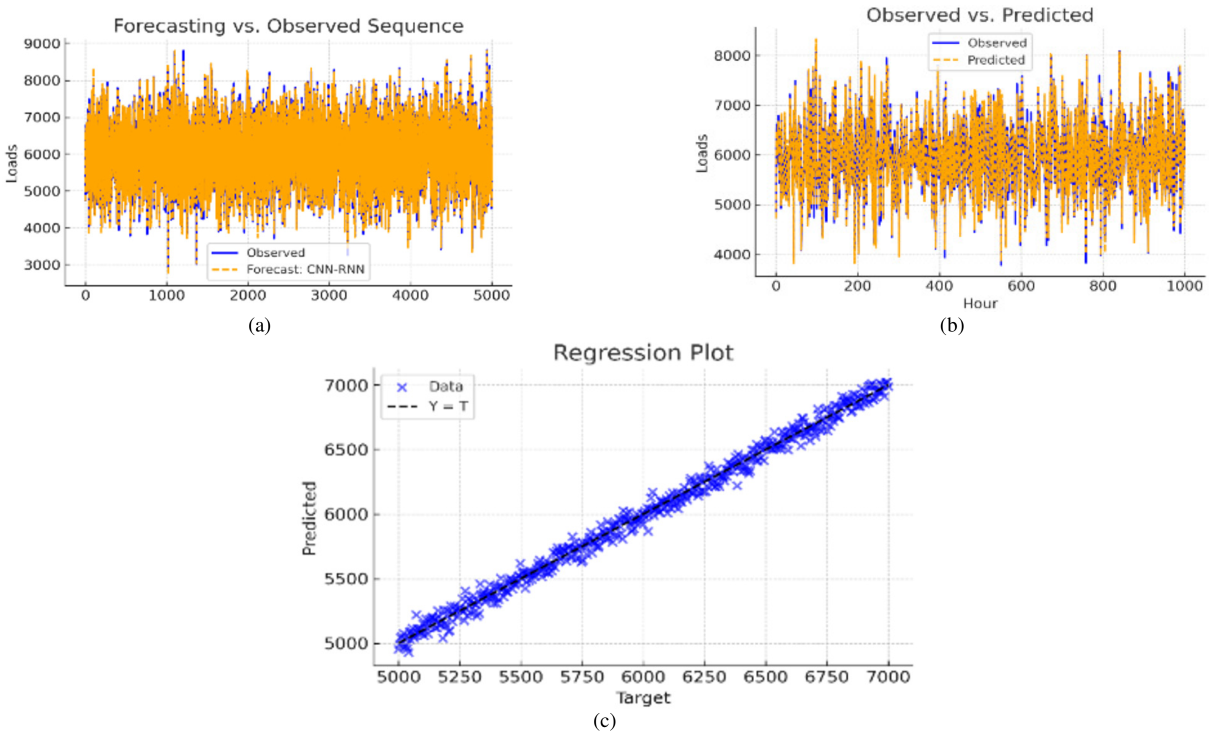
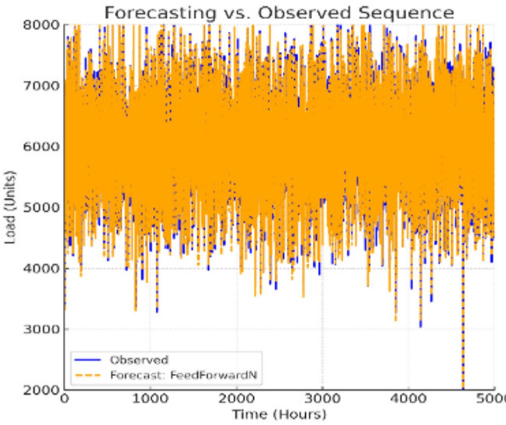
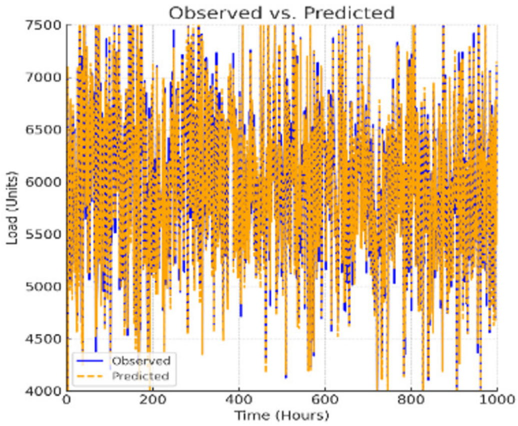


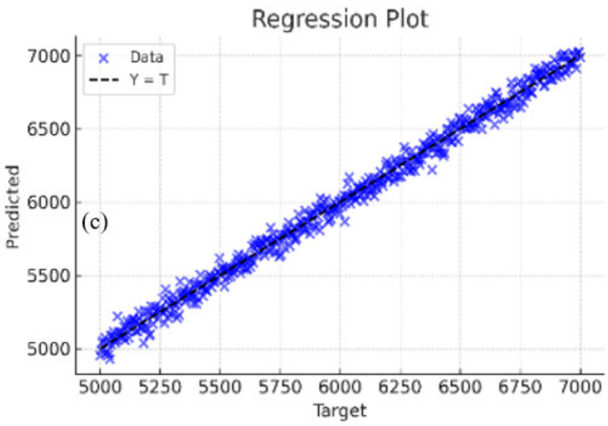
Fig. 7. Performance evaluation of the CNN–RNN model for Qassim load: (a) training and testing phase results, (b) testing phase results, and (c) regression analysis plot.



(a)



(b)



(c)

Fig. 8. Performance evaluation of the ANN model for Qassim load: (a) training and testing phase results, (b) testing phase results, and (c) regression analysis plot.

The performance plots for the CNN–RNN models along with the ANN models are shown in Figures 7 and 8. When analyzing the training and testing phases, a regression coefficient of 0.9881 was obtained from the regression plot, which shows high accuracy as it approaches the value of one. The error distribution of the training and testing data appears normal based on the histogram presented in Figure 7(c).

The results of the combination of the hybrid CNN–RNN model are presented in Figure 8, which demonstrates actual and predicted Qassim loads. The hybrid DL model in Figure 8(c) predicts Qassim load data with an R = 0.9883 regression coefficient that proves its excellent predictive power.

The proposed hybrid DL model demonstrated exceptional empirical performance, outperforming traditional prediction models in forecasting accuracy. The integration of the two models significantly improved the short-term load forecasting accuracy and enabled the extraction of complex spatiotemporal dependencies within the Qassim smart grid datasets.

It can be concluded that the hybrid approach effectively improved the predictive capabilities of the models. These hybrid models were specifically designed to capture complex relationships and patterns in the load data, providing a robust solution for precise energy demand forecasting. The results of this approach hold practical significance and can serve as a valuable tool for strategic power system planning, optimization, and operational decision-making in smart grids. As shown earlier, the performance evaluation results confirm that CNN-GRU outperforms other models. By integrating CNN's spatial feature extraction with GRU's efficient sequence learning, the hybrid model significantly enhances load forecasting accuracy. With fewer computations than LSTM, GRU enables faster convergence and reduces training costs. Additionally, CNN minimizes input dimensionality, improving generalization and resilience to noisy or missing data. Performance metrics (RMSE, MAPE, R) validate the accuracy and stability of CNN-GRU, whereas visualizations and case studies highlight its practical effectiveness in smart grid applications.

It is also important to remember that the proposed model exhibits strong scalability, as it can be quickly trained on new datasets to adapt to diverse consumption patterns influenced by climate, population density, and industrial activity. Its ability to perform effectively across different regions with minimal retraining or fine-tuning highlights its transferability, a key factor in ensuring adaptability to new cities or energy markets. Moreover, the model is designed to handle variations in input features such as weather conditions, socioeconomic factors, and grid infrastructure, allowing it to seamlessly adjust to evolving scenarios. Scalability is also reinforced by the model's capacity to efficiently process large datasets, making it well-suited for deployment in extensive and complex energy markets. This capability is a defining strength of the proposed CNN-RNN model, ensuring both robustness and flexibility in real-world applications.

V. CONCLUSION

Efficient load forecasting is crucial for managing energy consumption and optimizing costs in smart grids. This study focuses on improving load forecasting for Qassim city using a hybrid Deep Learning (DL) approach. The proposed model combines two DL algorithms, such as Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional Long Short-Term Memory (BiLSTM), or Recurrent Neural Network (RNN), to enhance the prediction accuracy and network architecture.

Our evaluation showed that hybrid models such as CNN–GRU and CNN–RNN outperformed other methods in predicting energy consumption, achieving lower RMSE, NRMSE, and MAPE values for both single-step and multi-step forecasts. The hybrid approach leverages Qassim's smart meter data to improve scalability and optimize grid operations.

The proposed model can significantly influence policy-making by optimizing energy distribution strategies, leading to more efficient resource allocation and reduced waste. It also supports the development of smart grid regulations by providing insights that enhance the integration of renewable energy sources. By analyzing data and predicting demand patterns, the model facilitates informed decision-making. Ultimately, these capabilities contribute to a more sustainable and resilient energy landscape.

Although the hybrid DL framework is computationally intensive compared to conventional methods, it provides more accurate forecasting results. Future work could incorporate metaheuristic algorithms such as genetic optimization to refine DL hyperparameters and reduce computational cost. Additionally, noise reduction techniques such as the Savitzky–Golay (SG) filter and wavelet decomposition can further improve prediction accuracy. Real-time forecasting could also be achieved by combining offline training on historical data with online testing on recent inputs.

Future research should also focus on addressing the challenges of real-time implementation of the proposed model. It is essential to evaluate computational requirements and hardware constraints, as these factors play a critical role in the deployment of CNN-GRU in practical smart grid applications. Investigating key aspects such as processing speed, memory consumption, and hardware acceleration techniques (e.g., GPUs, edge computing) would offer valuable insights into the feasibility of the model for real-world use. Furthermore, assessing the trade-off between accuracy and computational efficiency could help optimize the model for real-time forecasting applications.

ACKNOWLEDGMENT

The authors gratefully acknowledge Qassim University, represented by the Deanship of Scientific Research, on the financial support for this research under the number (2023-SDG- 1 -HSRC-36957) during the academic year 1445 AH / 2023 AD.

REFERENCES

[1] M. Espinoza, J. A. K. Suykens, R. Belmans, and B. De Moor, "Electric Load Forecasting," *IEEE Control Systems Magazine*, vol. 27, no. 5, pp. 43–57, Oct. 2007, https://doi.org/10.1109/MCS.2007.904656.

[2] A. Abessi, V. Vahidinasab, and M. S. Ghazizadeh, "Centralized Support Distributed Voltage Control by Using End-Users as Reactive Power Support," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 178–188, Jan. 2016, https://doi.org/10.1109/TSG.2015.2410780.

[3] K. Mahmud, J. Ravishankar, M. J. Hossain, and Z. Y. Dong, "The Impact of Prediction Errors in the Domestic Peak Power Demand Management," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4567–4579, Jul. 2020, https://doi.org/10.1109/TII.2019.2946292.

[4] N. T. Dung and N. T. Phuong, "Short-Term Electric Load Forecasting Using Standardized Load Profile (SLP) And Support Vector Regression (SVR)," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4548–4553, Aug. 2019, https://doi.org/10.48084/etasr.2929.

[5] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-Term Load Forecasting in Smart Grids: An Intelligent Modular Approach," *Energies*, vol. 12, no. 1, Jan. 2019, Art. no. 164, https://doi.org/10.3390/en12010164.

[6] M. Jawad *et al.*, "Machine Learning Based Cost Effective Electricity Load Forecasting Model Using Correlated Meteorological Parameters," *IEEE Access*, vol. 8, pp. 146847–146864, 2020, https://doi.org/10.1109/ACCESS.2020.3014086.

[7] D. Solyali, "A Comparative Analysis of Machine Learning Approaches for Short-/Long-Term Electricity Load Forecasting in Cyprus," *Sustainability*, vol. 12, no. 9, May 2020, Art. no. 3612, https://doi.org/10.3390/su12093612.

[8] K. P. Amber, R. Ahmad, M. W. Aslam, A. Kousar, M. Usman, and M. S. Khan, "Intelligent techniques for forecasting electricity consumption of buildings," *Energy*, vol. 157, pp. 886–893, Aug. 2018, https://doi.org/10.1016/j.energy.2018.05.155.

[9] H. Sangrody, N. Zhou, S. Tutun, B. Khorramdel, M. Motalleb, and M. Sarailoo, "Long term forecasting using machine learning methods," in *2018 IEEE Power and Energy Conference at Illinois*, Champaign, IL, USA, 2018, pp. 1–5, https://doi.org/10.1109/PECI.2018.8334980.

[10] Y. Ding, Q. Zhang, T. Yuan, and F. Yang, "Effect of input variables on cooling load prediction accuracy of an office building," *Applied Thermal Engineering*, vol. 128, pp. 225–234, Jan. 2018, https://doi.org/10.1016/j.applthermaleng.2017.09.007.

[11] M. A. Kabir and S. Anowarul Fattah, "Machine Learning based Load Forecasting using Efficient Feature Extraction Scheme in Smart Grid," in *2024 IEEE 1st International Conference on Green Industrial Electronics and Sustainable Technologies*, Imphal, India, 2024, pp. 1–5, https://doi.org/10.1109/GIEST62955.2024.10959795.

[12] M. Nasseri, T. Falatouri, P. Brandtner, and F. Darbanian, "Applying Machine Learning in Retail Demand Prediction—A Comparison of Tree-Based Ensembles and Long Short-Term Memory-Based Deep Learning," *Applied Sciences*, vol. 13, no. 19, Oct. 2023, Art. no. 11112, https://doi.org/10.3390/app131911112.

[13] X. Dong, L. Qian, and L. Huang, "Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Jeju, South Korea, 2017, pp. 119–125, https://doi.org/10.1109/BIGCOMP.2017.7881726.

[14] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, Oct. 2015, https://doi.org/10.1016/j.rser.2015.04.065.

[15] H. Sanchez. "Time Series Forecasting Using Hybrid CNN – RNN." Mathworks. https://www.mathworks.com/matlabcentral/fileexchange/91360-time-series-forecasting-using-hybrid-cnn-rnn.

[16] H. A. Al-Jamimi, G. M. BinMakhashen, M. Y. Worku, and M. A. Hassan, "Advancements in Household Load Forecasting: Deep Learning Model with Hyperparameter Optimization," *Electronics*, vol. 12, no. 24, Dec. 2023, Art. no. 4909, https://doi.org/10.3390/electronics12244909.

[17] S. A. Nabavi, S. Mohammadi, N. H. Motlagh, S. Tarkoma, and P. Geyer, "Deep learning modeling in electricity load forecasting: Improved accuracy by combining DWT and LSTM," *Energy Reports*, vol. 12, pp. 2873–2900, Dec. 2024, https://doi.org/10.1016/j.egyr.2024.08.070.

[18] J. Gao, Y. Chen, W. Hu, and D. Zhang, "An adaptive deep-learning load forecasting framework by integrating transformer and domain knowledge," *Advances in Applied Energy*, vol. 10, Jun. 2023, Art. no. 100142, https://doi.org/10.1016/j.adapen.2023.100142.

[19] X. Wen, J. Liao, Q. Niu, N. Shen, and Y. Bao, "Deep learning-driven hybrid model for short-term load forecasting and smart grid information management," *Scientific Reports*, vol. 14, no. 1, Jun. 2024, Art. no. 13720, https://doi.org/10.1038/s41598-024-63262-x.

[20] R. Chandrasekaran and S. K. Paramasivan, "Advances in Deep Learning Techniques for Short-term Energy Load Forecasting Applications: A Review," *Archives of Computational Methods in Engineering*, vol. 32, no. 2, pp. 663–692, Mar. 2025, https://doi.org/10.1007/s11831-024-10155-x.

[21] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network," in *2017 51st Annual Conference on Information Sciences and Systems*, Baltimore, MD, USA, 2017, pp. 1–6, https://doi.org/10.1109/CISS.2017.7926112.

[22] "Deep Learning Toolbox." Mathworks. https://www.mathworks.com/products/deep-learning.html.