This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/upec65436.2025.11279758, 2025 60th International Universities Power Engineering Conference (UPEC)

GridNet: A Hybrid LSTM-Based Deep Learning Approach for Accurate Electricity Consumption Forecasting in Smart Grid Systems

Talal Alazemi
Department of Laboratory Technology, College of Technological Studies, Public Authority for Applied Education and Training (PAAET)
Shuwaikh, Kuwait
ts.alazmi@Paaet.edu.kw

Mohamed Darwish
Department of Electronics and electrical engineering, Brunel University of London
Uxbridge, UK
mohamed.darwish@brunel.ac.uk

Muhammad Zubair
Department of Computer and Information Sciences Universiti Teknologi PETRONAS,
Seri Iskandar 32610, Malaysia
Muhammad_22000228@utp.edu.my

***Abstract*— Accurate forecasting of power production and consumption is crucial for optimizing smart grid operations, especially with the growing integration of renewable energy sources, and minimizing CO_2 emissions. This study develops GridNet, a deep learning-based model for forecasting next-hour electricity consumption using the past 24 hours of electricity and time features. GridNet utilizes a two-layer stacked Long Short-Term Memory (LSTM) network with dropout layers to prevent overfitting and improve generalization, followed by a dense output layer. Early stopping and learning rate reduction callbacks were applied to enhance convergence efficiency. The model was optimized using the Adam optimizer and a hybrid loss function combining Mean Squared Error (MSE) with L1 regularization for improved prediction accuracy and robustness. GridNet was trained on a six-year hourly dataset from Romania, covering diverse energy sources like nuclear, wind, solar, for 30 epochs. The performance of the proposed model GridNet was evaluated using several standard metrics, including the Coefficient of Determination (R^2), Mean Absolute Error (MAE), MSE, and Root Mean Squared Error (RMSE). Additionally, GridNet's effectiveness was compared with state-of-the-art algorithms, including XGBoost, KNN, Random Forest, and SVM. GridNet outperforms all competing algorithms by up to 59.9% in R^2, 72.0% in MAE, 89.8% in MSE, and 68.1% in RMSE, demonstrating its superior accuracy in forecasting electricity consumption trends in smart grids.**

Keyword—deep learning, electricity consumption forecasting, LSTM, renewable energy sources, smart grids

I. Introduction

Generating electricity often involves burning fossil fuels like coal, oil, and natural gas. This process releases carbon dioxide (CO_2) into the atmosphere, contributing to the greenhouse effect and global climate change. In fact, burning fossil fuels for power generation accounts for over 40% of energy-related CO_2 emissions. As the demand for electricity rises, especially with the increasing use of air conditioning during extreme weather events, the reliance on fossil fuels may also increase, leading to higher emissions [1], [2]. To address this, accurately forecasting building energy consumption is crucial. Such predictions help engineers balance supply and demand, optimize energy use, and make informed decisions [3]–[5]. Smart grid systems are emerging as a transformative solution for achieving greater efficiency, stability, and sustainability in energy distribution and management. By integrating advanced technologies, smart grids facilitate real-time monitoring and control, enabling more efficient energy use and reducing the carbon footprint of electricity generation [1], [6]. However, accurately predicting energy consumption within smart grids remains challenging due to the inherent randomness influenced by factors such as consumer behavior and weather conditions. The growing complexity of modern smart grids, driven by renewable energy integration and dynamic consumption patterns, has heightened the need for accurate electricity consumption forecasting [3], [7].

Likewise, Velayutham et al. [8] introduced a Quantum Boltzmann Machine (QBM) for smart grid energy forecasting, integrating variables like reaction time, power balance, and price elasticity via smart meters. The QBM models empirical probabilities to predict future consumption and identify local extremes. However, its reliance on historical data may not fully capture future changes in consumption patterns or external factors. Further, Aljarrah et al. [9] proposed a load forecasting using an LSTM-RNN model enhanced by the Improved Sparrow Search Algorithm (ISSA), while real-time energy trading and load balancing are managed through the BSET-AVVO algorithm. However the network was very complex. Furthermore, Aguiar et al. [10] proposed deep learning techniques, particularly LSTM networks based on RNNs, to forecast smart grid energy demand by learning consumption patterns from customer data, with an emphasis on short-term forecasting for real-time demand response and grid balance.

Additionally, Softah et al. [11] proposed a DVQR framework combined with K-means clustering to forecast smart meter demand by capturing variability and dependencies. However, it may face high computational costs with large datasets. Also, Lilhore et al. [12] proposed a hybrid CNN-BiLSTM model with two-way attention and MPSO optimization for short-term load prediction, using t-Nearest Neighbors for preprocessing and validated on the UCI Electrical Grid Stability dataset. However, the model's complexity may increase training time and computational requirements. Further, Nowak et al. [13] proposed decentralized algorithms leveraging smart meter and IoT infrastructure to compute and forecast network power flows in near real-time without centralizing consumer data. Field trials in Rolle VD validate the method, showing it can detect congestion issues effectively while preserving data privacy and matching centralized method's performance. Furthermore, Olumba et al. [14] proposed GradientSHAP, combining gradient boosting algorithms with SHAP values to predict energy demand while enhancing interpretability. Performance is validated using the European energy demand dataset against linear regression and SVR models. Additionally, Syu et al. [15] introduced distributed Multi-Head (DMH) learning systems for power consumption prediction in smart factories. DMH employs multi-head learning mechanisms to reduce noise interference and enhance accuracy. It utilizes split learning, where head networks reside on clients (AGVs) and the prediction network on the server, minimizing client-to-server data transmission and enhancing privacy.

Similarly, Khan et al. [16] proposed methodology involves a dual-stream deep learning architecture for power forecasting. The first stream utilizes a CNN to capture spatial features, while the second stream employs a LSTM network to extract temporal patterns. These features are then fused and processed through a self-attention mechanism to enhance feature selection. Finally, the integrated features undergo Principal Component Analysis (PCA) for dimensionality reduction before being input into fully connected layers for forecasting, however the system was complex. Also, Liu et al. [17] proposed methodology for Short-Term Load Forecasting (STLF) integrates a pattern extraction mechanism with an attention-based framework to enhance forecasting accuracy. Initially, the model employs a LSTM network to capture temporal dependencies from historical load data. Subsequently, a Time Pattern Attention (TPA) mechanism is applied to discern significant temporal features, effectively filtering out irrelevant information and emphasizing critical patterns.

Recent advancements in LSTM networks have significantly enhanced energy forecasting accuracy by effectively capturing complex temporal dependencies in sequential data. Studies have demonstrated that LSTM models can achieve low Mean Absolute Percentage Error (MAPE) and RMSE values, indicating their proficiency in predicting energy demand and production [18]. However, challenges such as computational complexity and the need for large datasets remain, which can impact the scalability and efficiency of these models in real-world applications. Motivated by these capabilities, this study proposes GridNet, a deep learning framework designed to enhance electricity consumption forecasting in smart grid systems. GridNet leverages a two-layer stacked LSTM architecture combined with a hybrid loss function to improve predictive accuracy and generalization. The main contribution of this paper is given below.

- Developed GridNet, a deep learning-based predictive algorithm for smart grid electricity consumption forecasting, featuring a two-layer stacked LSTM architecture interleaved with dropout layers to prevent overfitting and enhance generalization.
- Introduced a hybrid loss function that combines MSE with L1 regularization to enhance prediction accuracy and reduce model complexity.
- Incorporated early stopping and learning rate reduction callbacks to optimize training efficiency and convergence.
- Forecasting performance are compared to traditional machine learning models like XGBoost, KNN, Random Forest, and SVM, based on R^2, MAE, MSE, and RMSE metrics.

The rest of this paper is organized as follows: Section II details the proposed methodology, including the hybrid loss function, GridNet architecture, and evaluation metrics. Section III presents the experimental results, discusses the findings, and compares the performance with competing algorithms. Finally, Section IV concludes the study with key insights and outlines directions for future research.

II. Methodology

A. Hybrid Loss function

To improve both predictive accuracy and generalization capability, a hybrid loss function was adopted in proposed GridNet. This hybrid loss integrates the traditional MSE and L1 regularization [19]. The MSE loss is crucial for accurately predicting energy production and consumption. It measures the average squared difference between actual and predicted values, ensuring the GridNet learns to minimize prediction errors. In energy forecasting, precise prediction of parameters like consumption, production, and renewable contributions (wind, solar, etc.) is essential to maintain grid stability, optimize energy resources, and reduce operational costs. Mathematically the MSE loss function can be represented as given in Equation (1).

$$L_{MSE} = \frac{1}{N}\sum_{J=1}^{N}(y_j - \hat{y}_j)^2 \qquad (1)$$

Where N is number of samples, y_j is the ground truth value for sample and $\hat{y}_j$ indicate predicted value for sample j. While minimizing prediction errors (via MSE) is important. complex deep learning models like LSTMs can easily overfit the training data. L1 regularization addresses this problem by promoting sparsity , making some weights exactly zero, simplifying the GridNet to select feature automatically improving generalization to unseen data [20]. Mathematically L1 regularization can be represented as given in Equation (2).

$$L_{L_1} = \sum_{J=1}^{N}|\omega_i| \qquad (2)$$

Where ω_i are the trainable model parameters i.e., *weights*, and L1 controls the complexity of GridNet by pushing unnecessary weights towards zero. The hybrid loss function is given in Equation (3).

$$L_{Hybrid} = L_{MSE} + \lambda \sum_{J=1}^{\cdot}|\omega_i| \qquad (3)$$

λ is the regularization coefficient (e.g., 0.001), balancing the importance of sparsity against prediction accuracy. The hybrid loss function combines MSE for accurate energy prediction and L1 regularization to promote model sparsity and generalization, ensuring reliable and robust energy forecasting. The combined effect, as depicted by the red curve in the plot as shown in Fig 1, results in a loss function that is sensitive to both prediction errors and model complexity. This dual consideration leads to improved generalization performance, especially in high-dimensional or noisy datasets. By integrating MSE and L1 regularization, the hybrid loss function harnesses the strengths of both approaches, making it well-suited for time series forecasting and other regression tasks in smart grid applications.

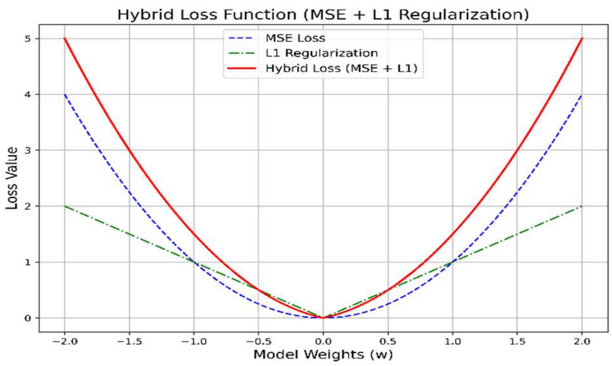

Fig 1: Hybrid Loss: MSE and L1 Regularization

The goal of incorporating L1 regularization is to encourage sparsity in the model parameters, reduce overfitting, and promote more robust learning in complex energy forecasting tasks.

B. GridNet Architecture

The proposed model architecture is composed of sequential layers tailored to capture temporal dependencies and enhance prediction robustness. GridNet is designed to accept input data in the form of sequences, where each sequence consists of 24 consecutive time steps. At each step, the model receives 11 distinct features representing relevant information such as electricity production sources, consumption, and time-based attributes. The input is structured as a three-dimensional tensor with shape (B, 24, 11), where B denotes the batch size. This format enables GridNet to learn temporal patterns across multiple features over a 24-hour period.

The first LSTM layer in the model architecture consists of 50 neurons and is configured with return sequences=*True*, allowing it to pass the entire sequence of outputs to the next LSTM layer. This setup is essential for capturing and preserving temporal dependencies across time steps. Mathematically, at each time step t_t, the hidden state h_t is computed using the LSTM function based on the current input x_t. The previous hidden state h_{t-1}, and the previous cell state c_{t-1} is updated as given in Equation (4).

$$h_t = LSTM\ (x_t,\ h_{t-1},\ C_{t-1}) \tag{4}$$

Here, x_t represents the input feature vector at time step t while h_{t-1}, and c_{t-1} retain the memory of prior computations, enabling the model to effectively learn long-term dependencies in the sequential data. Following the first LSTM layer, a dropout layer is incorporated with a dropout rate of 20%. This layer randomly deactivates 20% of the neurons during each training iteration, which helps prevent overfitting by reducing reliance on specific neurons and encouraging the network to learn more robust and generalizable features from the data.

The second LSTM layer also consists of 50 neurons, but differs from the first one, where configured with return sequences=False, meaning it outputs only the final hidden state rather than the full sequence. This final state, denoted effectively captures and summarizes information from the entire 24-time-step input sequence, serving as a compact representation that is passed to the subsequent layers for prediction. Mathematically it can be represented in Equation (5).

$$h_{final} = h_{t=24\mathrm{h}} \tag{5}$$

To further enhance the GridNet generalization capability and reduce the risk of overfitting, a *second dropout layer* was introduced, applying a 20% dropout rate. This additional regularization step helps ensure that the model does not become overly reliant on specific neurons, thereby improving its robustness and performance on unseen data.

The GridNet architecture concludes with a *dense output layer* comprising a single neuron, which is responsible for producing the final prediction of electricity consumption for the next hour. This layer takes the final hidden state from the preceding LSTM layer and maps it to a scalar value, effectively summarizing the learned temporal features into a single output. The output is computed through a linear transformation of the final LSTM state, allowing the model to generate a continuous-valued prediction as given in Equation (6).

$$\hat{y} = \mathrm{W}h_{\mathrm{final}} + \mathrm{b} \tag{6}$$

where W is the weight matrix and b are the bias term.

C. Evaluation Metrics

To assess the performance of the proposed forecasting models GridNet, a set of standard regression evaluation metrics was employed. The most important one is the R^2, which measures how well the predicted electricity consumption values approximate the actual values. A value close to 1 indicates that the GridNet explains most of the variability in the target variable, which is crucial for reliable energy demand forecasting in smart grid [21]. Mathematically it can be represented as given in Equation (7).

$$R^2 = 1 - \frac{\sum_{J=1}^{N}(y_j - \hat{y_j})^2}{\sum_{J=1}^{N}(y_j - \bar{y})^2} \tag{7}$$

Where y_i indicate actual values, $\hat{y_j}$ represent the predicted values and n indicate the number of samples. Furthermore, MAE gives the average magnitude of the errors in a set of predictions, without considering their direction. In electricity prediction, it directly indicates how far off the forecasted values are from the actual values, making it easy to interpret and important for cost estimation or energy allocation planning. Mathematically it can be represented as given in Equation (8).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_j - \hat{y_j}\right| \tag{8}$$

Additionally, MSE penalizes larger errors more than MAE due to the squaring of residuals. In power systems, this is helpful for catching and correcting significant deviations (e.g., sudden spikes or drops in energy consumption), which may impact operational stability or cause resource misallocation. Mathematically it can be represented as given in Equation (9).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_j - \hat{y}_j)^2 \tag{9}$$

Similarly, RMSE [3] is simply the square root of MSE and expresses the error in the same unit as the target variable (e.g., megawatts). It is widely used in power load forecasting as it gives a tangible sense of error magnitude, especially when large errors are critical to avoid. Mathematically it can be represented as given in Equation (10).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_j - \hat{y}_j)^2} \tag{10}$$

Together, these evaluation metrics provide a comprehensive assessment of the GridNet forecasting accuracy, robustness, and reliability in predicting electricity consumption, ensuring its practical applicability in smart grid energy management systems.

III. RESULTS AND DISCUSSION

A. *Dataset Description*

The dataset, retrieved from Kaggle, comprises six years of hourly records of electricity consumption and production in Romania's smart grid system, covering a diverse range of energy sources including nuclear, wind, hydroelectric, solar, biomass, oil and gas, and coal. The dataset contains 54,146 hourly samples, each with 11 features including a timestamp (DateTime) and energy-related attributes. It captures electricity Consumption and Production, with the latter detailed by individual energy sources such as Nuclear, Wind, Hydroelectric, Oil and Gas, Coal, Solar, and Biomass. While rich in temporal and energy sector data the data was split into a training set of 43,316 samples (80%) and a test set of 10,830 samples (20%) for model development and evaluation. In Fig 2 , the stacked area plot shows hourly electricity production by source, with Oil and Gas (blue) and Coal (orange) being the primary contributors. Solar (green) and Biomass (red) have smaller shares, with Solar showing daily patterns.

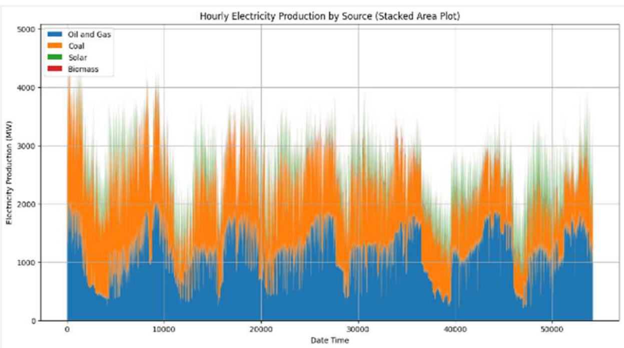

Fig 2: Hourly Electricity Production by Source

The histogram illustrates in Fig 3, shows the distribution of electricity generation across three energy sources: Nuclear, Wind, and Hydroelectric. Nuclear energy shows a narrow distribution centered around lower electricity values (~100 MW), indicating consistent output. Wind energy displays a wider spread, peaking around 150 MW, suggesting greater variability. Hydroelectric generation is distributed at higher values, mainly between 170 MW and 230 MW, reflecting both higher capacity and variability compared to Nuclear and Wind sources.

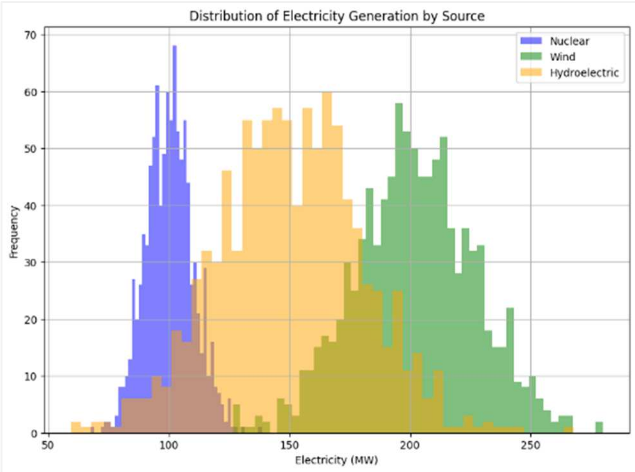

Fig 3: Electricity generation distribution by source

B. *Model Configuration and Training Parameters*

The Table 1 summarizes the key configuration parameters used for model development, including sequence length, input features, network architecture, optimizer, loss function, training settings, and data scaling approach.

TABLE 1: GRIDNET TRAINING PARAMETERS

Parameters	Value
Sequence Length	24 (hours)
Input Features	9
RNN Units (per layer)	50
Output Layer	Dense (1 unit)
Optimizer	Adam
Loss Function	(MSE+L1)
Epochs	30
Batch Size	32
Validation Split	0.2 (20%)
Scaling	MinMaxScaler
Regularization coefficient (λ)	0.001

C. *Evaluation of Power Consumption Predictions*

Table 2, presents the first 10 hourly records of power consumption from the test dataset. The proposed model, GridNet, was applied to forecast the power consumption for each hour. Initially, both the actual and predicted values were in a normalized form (0,1) when model tested, however, they have been converted back to megawatts (MW) for better understanding and interpretation. The "Hours" column represents the specific hour of observation, while the "Actual Power Consumption" column shows the true electricity demand at that time. Correspondingly, the "Predicted Power Consumption" column provides the predicted power consumption values generated by the proposed GridNet algorithm. The table below presents the actual and predicted power consumption values (in megawatts) for the first 10 hours of the test dataset.

TABLE 2: ACTUAL AND PREDICTED POWER CONSUMPTION (MW) FOR THE FIRST 10 HOURS

Actual Power consumption		Predicted Power consumption
Hours	**Megawatt (MW)**	**Megawatt (MW)**
1	8185	8204.391316
2	8506	8355.695371
3	8416	8590.510251
4	8294	8403.957994
5	8032	8172.053336
6	7671	7644.052368
7	6963	7071.17413
8	6513	6535.686769
9	6171	6194.976883
10	5990	5991.220204

The above table presents a comparison between actual and predicted power consumption (in megawatts) over a 10-hour period. The corresponding line plot visually illustrates this comparison, where the blue line represents actual values, and the orange dashed line shows predictions made by the model. Across all hours, the predicted values closely follow the actual consumption trend, with minimal deviation. For example, at hour 1, the actual consumption is 8185 MW, while the model predicts 8204.4 MW. Similarly, at hour 10, the values are nearly identical (5990 MW actual vs. 5991.2 MW predicted), indicating high prediction accuracy. In Fig 4 the close alignment of both lines in the graph confirms the model's ability to effectively forecast short-term power demand, showcasing strong generalization and low prediction error.

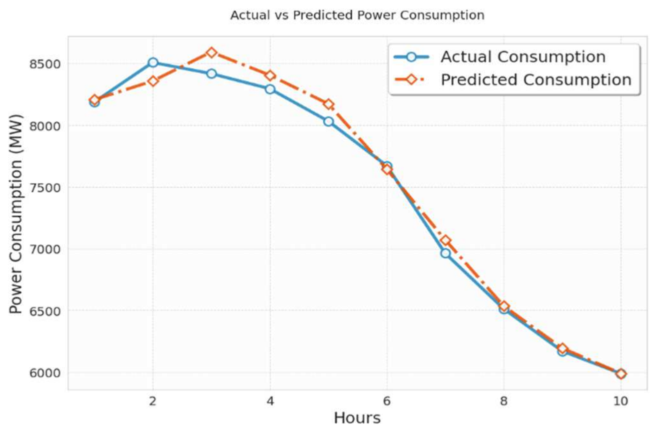

Fig 4: Line graph of actual vs. predicted power.

D. Performance Evaluation on Test Dataset

The Fig 5 illustrates the GridNet performance on the test dataset by comparing actual and predicted electricity consumption over time. The x-axis represents time steps (ranging from 0 to approximately 10,000), while the y-axis denotes the normalized electricity consumption (in megawatts, MW) ranging from 0.0 to 1.0. The blue line indicates the actual consumption, the orange line represents the predicted consumption by the model, and the green shading highlights the error range. The close alignment between the actual and predicted curves suggests that the model accurately captures the underlying consumption patterns and generalizes well to unseen data, demonstrating strong predictive capability and minimal deviation across the entire test set.

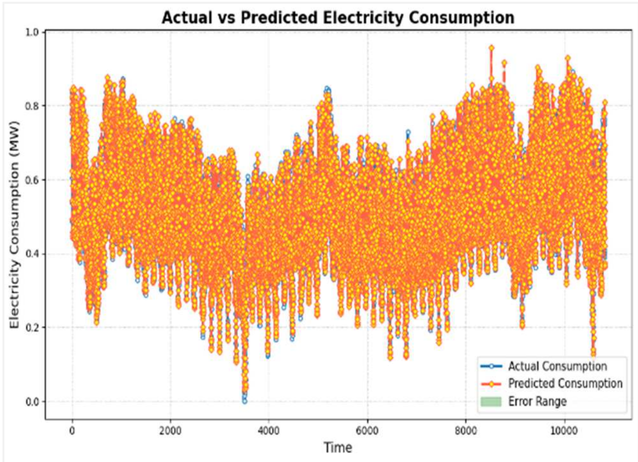

Fig 5: Actual vs Predicted Electricity Consumption

The plot in Fig 6 illustrates the training and validation loss MSE of the proposed GridNet model over 30 epochs. Both losses decrease rapidly during the initial epoch and gradually converge to low values, indicating efficient learning. The validation loss closely follows the training loss throughout, demonstrating good generalization and no signs of overfitting. Overall, GridNet exhibits stable and effective performance across the training process.

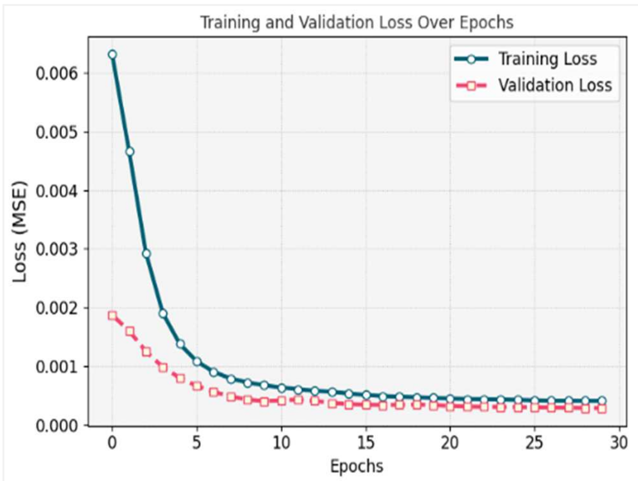

Fig 6: Training and Validation Loss Over Epochs

In Fig 7 the correlation matrix heatmap in displays the relationships between different energy sources, including Nuclear, Wind, Hydroelectric, Oil and Gas, Coal, Solar, and Biomass. All correlation values are close to zero, indicating weak or no linear relationships among the features. This suggests that the energy sources are largely independent, which is beneficial for GridNet. The low inter-feature correlation helps GridNet learn distinct patterns from each input feature, enhancing its ability to model complex relationships and improving its predictive performance.

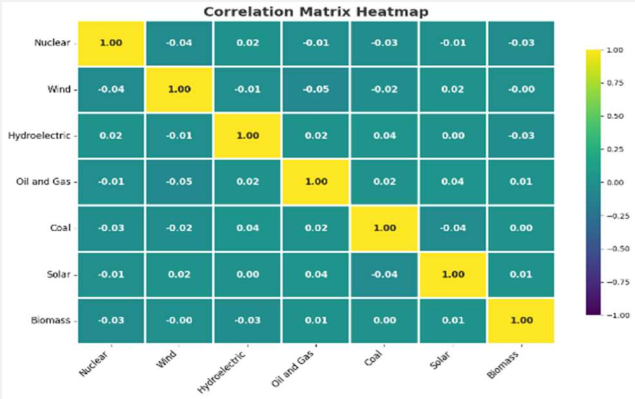

Fig 7: Correlation matrix heatmap of energy sources

Comparative Analysis with Baseline Models

The Table 3 summarizes the performance comparison of various machine learning models XGBoost, K-Nearest Neighbors (KNN), Random Forest, Support Vector Machine (SVM), and the proposed GridNet, based on four key evaluation metrics: R^2 Score, MAE, MSE, and RMSE. Among all models, GridNet achieved the highest R^2 Score of 0.959, indicating a very strong correlation between the predicted and actual values. It also recorded the lowest MAE (0.021), MSE (0.001), and RMSE (0.031), demonstrating superior predictive accuracy and robustness.

In contrast, the conventional models such as XGBoost, KNN, Random Forest, and SVM showed relatively lower performance, with R^2 scores around 0.65 and notably higher error metrics. These results clearly highlight the effectiveness and reliability of the proposed GridNet model in forecasting electricity consumption with significantly better precision compared to traditional methods. It shows a 47% to 60% increase in R^2 score, and reduces MAE, MSE, and RMSE by 65% to 89%. These results highlight GridNet's superior accuracy in predicting electricity consumption, demonstrating its effectiveness for more reliable energy forecasting.

TABLE 3:PERFORMANCE COMPARISON: GRIDNET VS. BASELINE MODELS

Model	R^2 Score↑	MAE↓	MSE↓	RMSE↓
XGBoost	0.653	0.069	0.0099	0.088
KNN	0.599	0.074	0.009	0.095
Random Forest	0.655	0.069	0.008	0.089
SVM	0.649	0.069	0.008	0.089
GridNet (Proposed)	**0.959**	**0.021**	**0.001**	**0.031**

IV. DISCUSSION

The proposed GridNet model, built upon LSTM layers and optimized with a hybrid loss function combining MSE and L1 regularization, demonstrated superior performance compared to traditional models like XGBoost, KNN, Random Forest, and SVM. This performance gain can be attributed to the model's ability to capture complex temporal dependencies in energy data and prevent overfitting through regularization. The consistent improvement across all evaluation metrics especially the significantly higher R^2 score highlights the effectiveness of combining sequence modeling with robust loss penalization. Despite the model's success, it may still be influenced by data variability, and its performance could fluctuate with different feature sets or sequence lengths. Future work may consider integrating external data sources or architectural enhancements for further improvements.

V. CONCLUSION AND FUTURE WORK

In this study, a novel deep learning architecture, GridNet, was proposed for forecasting electricity consumption in smart grid systems. The novelty of GridNet lies in its integration of Long Short-Term Memory (LSTM) layers with a custom-designed hybrid loss function, which combines Mean Squared Error (MSE) with L1 regularization. This dual-objective approach not only enables the model to accurately capture temporal dependencies inherent in electricity consumption data but also enhances generalization by penalizing unnecessary model complexity, thus reducing the risk of overfitting. Furthermore, GridNet introduces a structurally optimized architecture tailored for smart grid forecasting scenarios, offering improved stability and forecasting precision compared to conventional LSTM-based models. This combination of architectural refinement and loss function design positions GridNet as a robust and efficient framework for intelligent energy demand prediction in dynamic environments.

GridNet significantly outperformed traditional machine learning models across key evaluation metrics such as R^2, MAE, MSE, and RMSE, demonstrating its robustness and predictive strength in handling multivariate time-series energy data. For future work, enhancements may include incorporating exogenous variables such as weather data, economic indicators, and calendar effects to further enrich model performance. Additionally, integrating attention mechanisms could help prioritize relevant features or time steps, while expanding the model for multi-step forecasting may improve its utility in real-world grid management. Exploring graph-based neural networks and transfer learning approaches may also offer promising directions to better model spatial-temporal relationships across interconnected energy infrastructures.

DATA AVAILABILITY STATEMENT

Upon reasonable request, the corresponding author will provide access to the data that supports this study.

CONFLIECT OF INTEREST

The authors declare that they have no conflicts of interest related to this study and that the research findings are free from any external influence or personal bias

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the Kuwait Foundation for the Advancement of Sciences (KFAS).

REFERENCES

[1] R. Mathumitha, P. Rathika, and K. Manimala, *Intelligent deep learning techniques for energy consumption forecasting in smart buildings: a review*, vol. 57, no. 2. Springer Netherlands, 2024. doi: 10.1007/s10462-023-10660-8.

[2] S. Ali, S. Bogarra, M. N. Riaz, P. P. Phyo, D. Flynn, and A. Taha, “From Time-Series to Hybrid Models: Advancements in Short-Term Load Forecasting Embracing Smart Grid Paradigm,” *Appl. Sci.*, vol. 14, no. 11, 2024, doi: 10.3390/app14114442.

[3] M. Zubair, H. M. Rais, and T. Alazemi, “A Novel Attention-Guided Enhanced U-Net with Hybrid Edge-Preserving Structural Loss for Low-Dose CT Image Denoising,” *IEEE Access*, no. October 2024, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3526619.

[4] A. Ahamed, M. T. R. Tarafder, S. M. T. H. Rimon, E. Hasan, and M. Al Amin, “Optimizing Load Forecasting in Smart Grids with AI-Driven Solutions,” *Proc. 2024 IEEE Int. Conf. Data Softw. Eng. Data-Driven Innov. Transform. Ind. Soc. ICoDSE 2024*, pp. 182–187, 2024, doi: 10.1109/ICODSE63307.2024.10829903.

[5] M. Zubair, A. Abubakar, S. Arsalaan, and K. Yousafzai, “Prediction of oxide glass refractive index using a novel deep learning architecture , DualNet (U - Net and ANN) with hybrid loss function,” 2025.

[6] P. Arévalo and F. Jurado, “Impact of Artificial Intelligence on the Planning and Operation of Distributed Energy Systems in Smart Grids,” *Energies*, vol. 17, no. 17, pp. 1–22, 2024, doi: 10.3390/en17174501.

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

[8] P. Velayutham, V. S. Balaji, and N. Krissh Shankaran, “Advanced Energy Consumption Forecasting in Smart Grids Using Quantum Boltzmann Machines,” *2024 Glob. Conf. Commun. Inf. Technol. GCCIT 2024*, pp. 1–6, 2024, doi: 10.1109/GCCIT63234.2024.10862398.

[9] E. Aljarrah, “AI-based model for Prediction of Power consumption in smart grid-smart way towards smart city using blockchain technology,” *Intell. Syst. with Appl.*, vol. 24, no. April, 2024, doi: 10.1016/j.iswa.2024.200440.

[10] J. M. Aguiar-Pérez and M. Á. Pérez-Juárez, “An Insight of Deep Learning Based Demand Forecasting in Smart Grids,” *Sensors*, vol. 23, no. 3, 2023, doi: 10.3390/s23031467.

[11] W. Softah, L. Tafakori, and H. Song, “Analyzing and predicting residential electricity consumption using smart meter data: A copula-based approach,” *Energy Build.*, vol. 332, no. January, p. 115432, 2025, doi: 10.1016/j.enbuild.2025.115432.

[12] U. K. Lilhore, S. Dalal, M. Radulescu, and M. Barbulescu, “Smart grid stability prediction model using two-way attention based hybrid deep learning and MPSO,” *Energy Explor. Exploit.*, 2024, doi: 10.1177/01445987241266892.

[13] S. Nowak, F. Widmer, J. Taylor, and V. Science, “Decentralised Privacy-preserving Power Flow Forecast on Smart Meters,” vol. 4, no. 4, pp. 100–108, 2024, doi: 10.1145/3717413.3717422.

[14] W. C. Olumba *et al.*, “Balancing Accuracy and Interpretability : GradientSHAP for Enhanced Energy Demand Predictions,” vol. 199, no. January, pp. 142–155, 2025.

[15] J.-H. Syu, J. C.-W. Lin, and P. S. Yu, “Distributed Multi-Head Learning Systems for Power Consumption Prediction,” pp. 1–11, 2025, [Online]. Available: http://arxiv.org/abs/2501.12133

[16] T. Khan and C. Choi, “Attention enhanced dual stream network with advanced feature selection for power forecasting,” *Appl. Energy*, vol. 377, no. PC, p. 124564, 2025, doi: 10.1016/j.apenergy.2024.124564.

[17] Y. Liu, R. Zheng, M. Liu, J. Zhu, X. Zhao, and M. Zhang, “Short-Term Load Forecasting Model Based on Time Series Clustering and Transformer in Smart Grid,” *Electron.*, vol. 14, no. 2, pp. 1–17, 2025, doi: 10.3390/electronics14020230.

[18] M. Zubair, H. B. Md Rais, and Q. Al-Tashi, “U-Net Autoencoder for Edge-Preserved Denoising of Low Dose Computed Tomography Images: A Novel Technique,” *2023 13th Int. Conf. Inf. Technol. Asia, CITA 2023*, pp. 19–24, 2023, doi: 10.1109/CITA58204.2023.10262803.

[19] M. Zubair, H. Rais, Q. Al-tashi, and F. Ullah, “Enabling Predication of the Deep Learning Algorithms for Low-Dose CT Scan Image Denoising Models : A Systematic Literature Review,” *IEEE Access*, vol. PP, p. 1, 2024, doi: 10.1109/ACCESS.2024.3407774.

[20] M. Zubair, H. Rais, F. Ullah, T. Alazemi, and A. K. Yousafzai, “A Hybrid U-Net Framework for Low-Dose CT Image Denoising : Leveraging EdgeNet + for Structural Integrity and Image Quality A Hybrid U-Net Framework for Low-Dose CT Image Denoising : Leveraging EdgeNet + for Structural Integrity and Image Quality,” no. November, 2024, doi: 10.37934/araset.64.1.121136.

[21] A. A. Suleiman, A. K. Yousafzai, and M. Zubair, “Comparative Analysis of Machine Learning and Deep Learning Models for Groundwater Potability Classification †,” *Eng. Proc.*, vol. 56, no. 1, 2023, doi: 10.3390/ASEC2023-15506.