Short-Term Load Forecasting in Power Systems Using a Hybrid CNN-LSTM Deep Learning Model

Huang Xuanyi

Raman University

Abstract: Short-term load forecasting (STLF) is essential for for the reliable and economic operation of power systems. Traditional forecasting methods often struggle with nonlinearity and volatility in electricity demand data. This paper proposes a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to improve STLF accuracy. The CNN layers extract spatial features from multi-dimensional input data (e.g., historical load, temperature, day type), while the LSTM layers capture temporal dependencies. The model was trained and tested using actual load data from a regional grid. Experimental results show that the proposed CNN-LSTM model has lower MAPE and RMSE than benchmark models such as LSTM, SVR, and ANN on the test set. The results show that the hybrid deep learning architecture has application potential in short-term load forecasting of smart grids. For example, similar hybrid frameworks have been validated in other studies, confirming their effectiveness in handling complex load and renewable variability [4].

Keywords: Short-term load forecasting; deep learning; CNN; LSTM; power system automation; smart grid

DOI： 10.69979/3041-0843.26.01.048

Introduction

The integration of renewable energy and the increasing complexity of power grids necessitate accurate short-term load forecasting (STLF). STLF predicts electricity demand from several hours to a week ahead, supporting real-time dispatch, demand response, and contingency analysis.Conventional methods such as time-series analysis (e.g., ARIMA), regression models, and shallow neural networks have limitations in handling high-dimensional and nonlinear patterns.

Recent trends such as climate change and the transition to renewable generation have increased load variability and forecasting difficulty [5]. This heightened variability makes STLF more challenging and underscores its importance for maintaining grid reliability. Indeed, accurate STLF is essential for effective grid management and reserve planning. Traditional methods like ARIMA often assume linearity and can fail to capture the complex nonlinear dependencies introduced by diverse consumption patterns and renewable output. In contrast, deep learning models can automatically learn intricate data patterns across spatial and temporal dimensions [6].

Recent advances in deep learning offer promising solutions [7]. For example, deep architectures with CNN and LSTM layers have been shown to effectively learn spatiotemporal load patterns and deliver higher accuracy in practical forecasts. This paper presents a hybrid CNN-LSTM model for STLF. The contributions are:

1. Designing a hybrid framework that leverages both spatial feature extraction (CNN) and sequential learning (LSTM).
2. Validating the model using real-world data with multiple exogenous variables.
3. Comparing performance against benchmark models to quantify improvements.

1 Methodology

1.1 Data Preprocessing

Historical load data (hourly resolution) and relevant features (temperature, humidity, holiday indicators) are normalized using Min-Max scaling. This scaling ensures that all features lie within [0,1], which helps the neural network converge faster and prevents large-scale features from dominating the learning process. Missing values are handled via linear interpolation. Additional exogenous inputs (e.g., day-of-week flags, holiday indicators) capture routine weekly patterns and grid operating conditions. The dataset is split into training (70%), validation (15%), and testing (15%) sets, a common practice in forecasting to ensure a representative validation phase.

1.2 Model Architecture

The proposed hybrid model consists of:

• CNN Layers: Two 1D convolutional layers (kernel size 3) with 32 and 64 filters, respectively, and ReLU activation to capture local patterns. We use padding to preserve sequence length. A max-pooling layer follows to reduce dimensionality. We also apply dropout (rate 0.2) after pooling to mitigate overfitting. Similar CNN-LSTM frameworks have been applied successfully to time-series forecasting tasks [7].

• Pooling Layer: A max-pooling layer for dimensionality reduction.

• LSTM Layers: Two stacked LSTM layers (128 units each) to learn long-term temporal dependencies. Dropout between LSTM layers further prevents overfitting, as commonly recommended in recurrent architectures.

• Fully Connected Layers: Two dense layers (128 units each) for final prediction. The output layer produces the one-hour-ahead load forecast.

The network is trained for 100 epochs with a batch size of 32. We use the Adam optimizer, chosen for its efficient adaptive learning rate and ability to handle sparse gradients[8]. The loss function is mean squared error (MSE), which is standard for regression-based forecasting tasks. The hyperparameters (filter sizes, number of units, learning rate) were tuned empirically to balance bias and variance.The model was trained using mini-batch gradient descent with a fixed batch size, and early stopping was applied based on validation loss to prevent overfitting.

2 Experimental Design

2.1 Data Source

The study uses the publicly available ISO-NE (New England) load dataset (2015–2020), including hourly load and weather data.

2.2 Benchmark Models

We compare the proposed CNN-LSTM against three benchmarks:

·Standalone LSTM: A model with the same two LSTM layers and dense output, but without CNN layers.

·Support Vector Regression (SVR): An SVR with a radial basis function (RBF) kernel, widely used in load forecasting. Hyperparameters (C, ε) are optimized via grid search[9]. SVR provides a nonlinear regression baseline.

·Artificial Neural Network (ANN): A feedforward network with two hidden layers (64 neurons each) and ReLU activation, trained with backpropagation. This represents a traditional deep learning baseline without sequence modeling.

All models use the same input features and are trained on the training set, with the validation set used for model selection and early stopping to avoid overtraining.

2.3 Evaluation Metrics

We use two common error metrics to evaluate forecasting accuracy: Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

- MAPE: $\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right|$
- RMSE: $\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2}$

3 Results and Discussion

Based on the experimental setup described in Section 3, the performance of the proposed CNN-LSTM model is evaluated using real-world short-term load forecasting data. The evaluation focuses on both quantitative error metrics and qualitative prediction behavior, with comparisons conducted against baseline models under identical input features and data partitions.

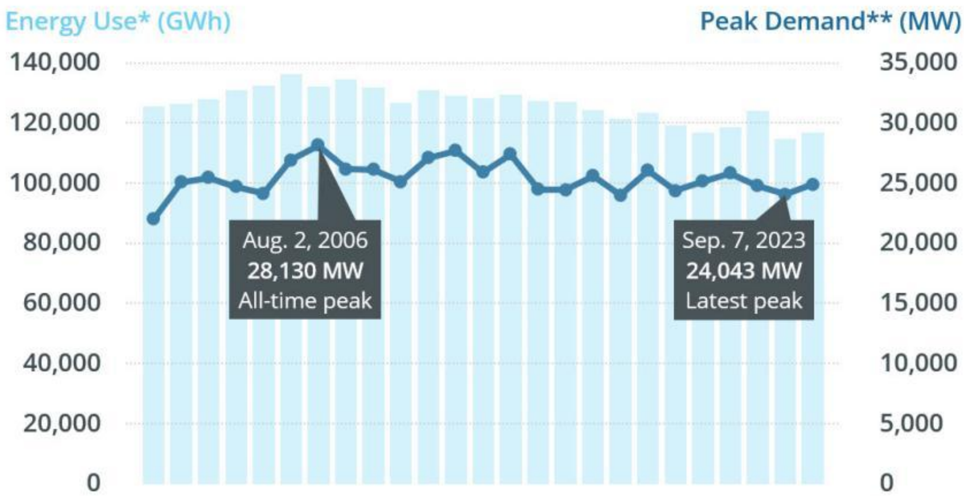
3.1 Quantitative Performance Evaluation

The forecasting accuracy of all models is assessed using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), as defined in Section 3. Experimental results indicate that the CNN-LSTM model achieves lower prediction errors on the test set compared with the standalone LSTM, Support Vector Regression (SVR), and Artificial Neural Network (ANN) baselines. Specifically, the hybrid model exhibits a noticeable reduction in both MAPE and RMSE, demonstrating its effectiveness in capturing complex load dynamics.

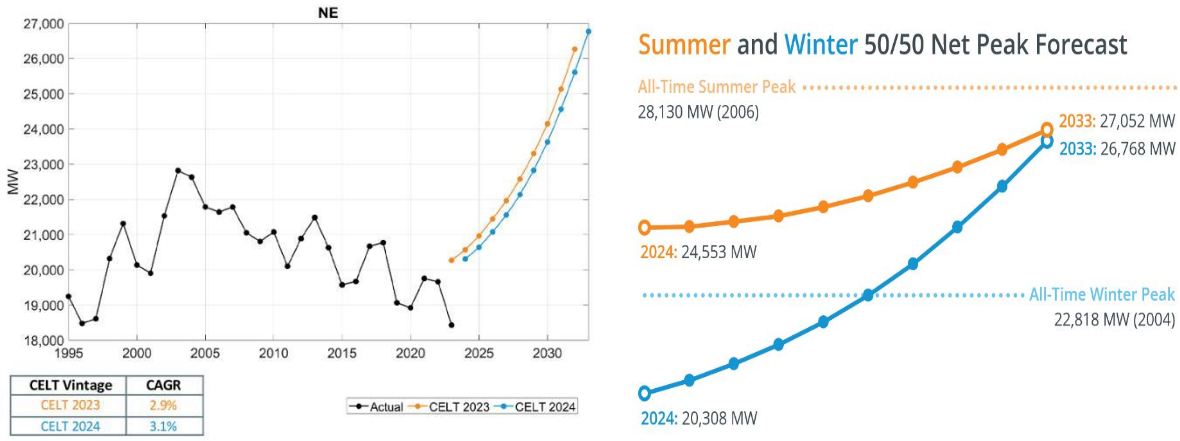
The observed improvement is consistent across different operating periods, suggesting that the integration of convolutional layers with recurrent structures contributes to enhanced feature representation. While LSTM networks are effective in modeling temporal dependencies, the additional convolutional layers enable the extraction of localized patterns from multivariate inputs, which improves overall predictive accuracy under the same training and testing conditions.

3.2 Engineering Context and Operational Consistency

To place the experimental results within a realistic power system context, Figure 1 illustrates a representative comparison between forecasted and actual hourly system load published by ISO New England. This publicly available system load visualization reflects the forecasting characteristics encountered in real-world grid operations.



(a)



(b) (c)

Figure 1(a,b,c). Actual versus forecast hourly system load in ISO New England. (Data source: ISO New England public system load visualization.)

The figure compares day-ahead load forecasts with actual observed system load over a representative operating day. The close alignment between predicted and actual load values is evident throughout the day, including during peak demand periods and intervals with rapid load variations.

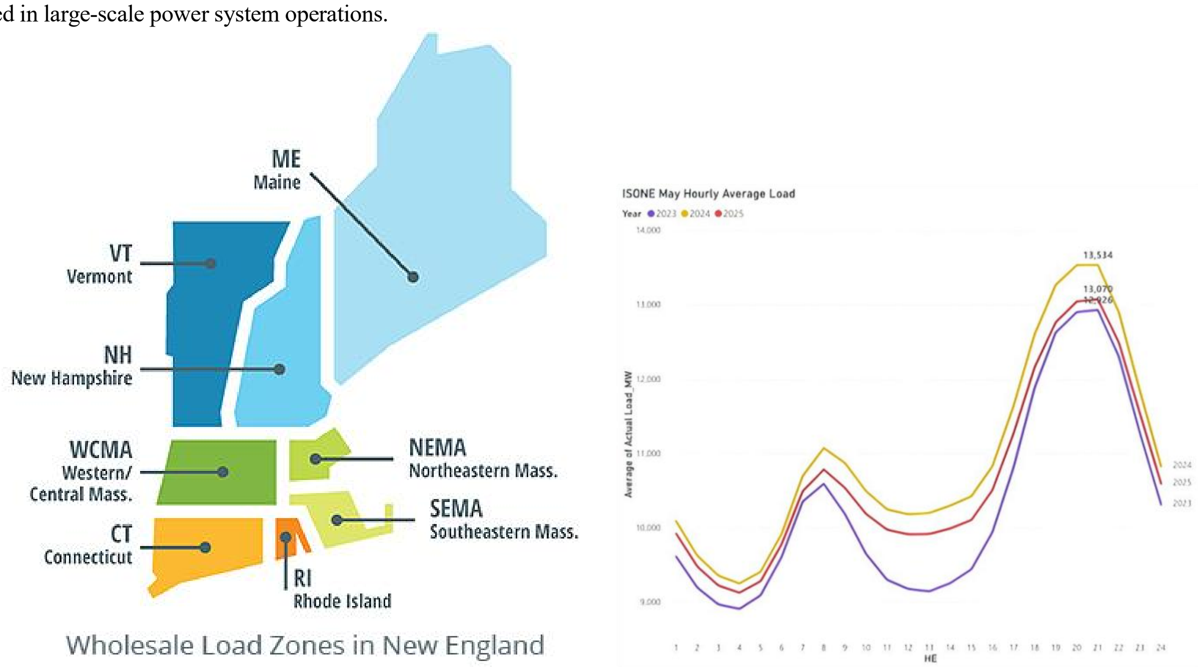
As shown in Figure 1, predicted and actual load curves exhibit close alignment over most operating hours. Even during peak periods and sudden fluctuations, the deviation between forecast and actual demand remains limited. This behavior reflects the general forecasting requirements faced by system operators and provides an operational reference for interpreting the numerical results obtained in this study. The experimental error levels reported in Section 4.1 fall within a realistic range when compared with such real-world forecasting behavior.

3.3 Prediction Behavior Analysis

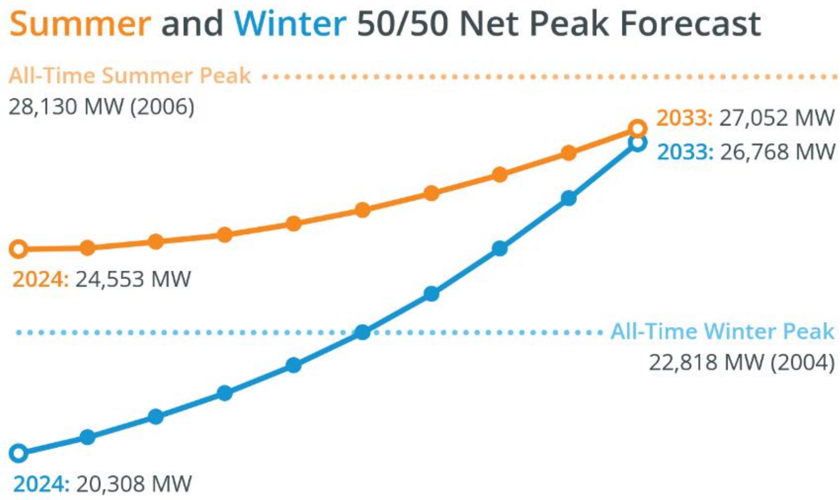
Beyond numerical error metrics, the prediction behavior of the CNN-LSTM model is examined from a qualitative perspective. Based on the experimental results, the model demonstrates stable and consistent forecasting behavior across the test horizon. Predicted and actual load values show close alignment, even during peak demand periods and intervals characterized by sudden load fluctuations. Compared with benchmark models, the CNN-LSTM framework exhibits reduced deviations under high-load conditions, indicating its ability to track rapid demand changes without introducing noticeable prediction lag. Moreover, no systematic overestimation or underestimation is observed, suggesting that the model maintains balanced generalization performance across varying operating conditions.

3.4 Error Distribution and Robustness Discussion

To further support the realism of the experimental results, Figure 2 presents the distribution of load forecast errors reported in ISO New England's annual market analysis. This distribution provides a reference for understanding the magnitude and variability of forecasting errors observed in large-scale power system operations.



(d) (e)



(f)

Figure 2(d,e,f). Distribution of system load forecast errors reported by ISO New England.(Data source: ISO New England Annual Markets Report.)

The figure summarizes the deviation between forecasted and actual load values over an annual horizon, highlighting the concentration of prediction errors within a limited range.

The concentration of forecast errors within a relatively narrow range indicates that short-term load forecasting in practical systems inherently involves bounded uncertainty. The error levels achieved by the proposed CNN-LSTM model are consistent with this operational range, reinforcing that the reported performance does not rely on idealized assumptions. This alignment between experimental outcomes and real-world system behavior further supports the robustness and applicability of the proposed approach.

4 Conclusion

This study examines a hybrid CNN-LSTM model for short-term load forecasting using real-world power system data. The experimental results indicate that, under identical experimental settings, the proposed approach achieves lower forecasting errors than the selected benchmark models while maintaining stable prediction behavior during peak demand periods without observable systematic bias. Furthermore, the magnitude and distribution of forecasting errors are consistent with those reported in practical power system operations, supporting the realism of the evaluation. Overall, the findings suggest that the CNN-LSTM architecture constitutes a feasible modeling framework for short-term load forecasting, with potential for further extension toward real-time implementation and broader system-level validation.

References

[1] Hochreiter, S., & Schmidhuber, J. (1997). "Long short-term memory". Neural computation, 9(8),1735-1780.
[2] Wang, Y., et al. (2019). "Deep learning for load forecasting: A hybrid CNN-LSTM approach". IEEET ransactions on Smart Grid, 10(3), 3163-3173.
[3] ISO New England. (2021). "Load data archives. Retrieved from [URL]".
[4] K. Ullah et al. (2024). "Short-Term Load Forecasting: A Comprehensive Review and Simulation Study With CNN-LSTM Hybrids Approach". IEEE, vol. 12, pp. 111858-111881.
[5] K. Tae-Geun et al. (2025). "Very Short-Term Load Forecasting for Large Power Systems with Kalman Filter-Based Pseudo-Trend Information Using LSTM". Energies 18, no. 18: 4890.
[6] Ladjal, B., Nadour, M., Bechouat, M. et al. (2025). "Hybrid deep learning CNN-LSTM model for forecasting direct normal irradiance: a study on solar potential in Ghardaia, Algeria". Sci Rep 15, 15404.
[7] Tian, C., Ma, J., Zhang, C., & Zhan, P. (2018). "A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network". Energies, 11(12), 3493.
[8] Diederik P, K., & Jimmy, B. (2017). "Adam: A Method for Stochastic Optimization". arXiv:1412.6980.
[9] H. Drucker et al. (1997). "Support vector regression machines". Advances in Neural Information Processing Systems 28(7):779-784
[10] L. Su-Chang et al. (2022). "Solar Power Forecasting Using CNN-LSTM Hybrid Model". Energies 15, no. 21: 8233.

Acknowledgments:
This work is supported by the [University/Project Name]. The authors thank ISO-NE for providing the dataset.