An LSTM-Based Approach to Day-Ahead Electricity Load Forecasting

Mohammad Alhashem
dept. of Electrical Engineering
University of Calgary
Calgary, Canada
mohammad.alhashem@ucalgary.ca

Yuhao Huang
dept. of Electrical Engineering
University of Calgary
Calgary, Canada
yuhao.huang@ucalgary.ca

Noureldin Amer
dept. of Software Engineering
University of Calgary
Calgary, Canada
noureldin.amer@ucalgary.ca

Ali Karimi
dept. of Physics and Astronomy
University of Calgary
Calgary, Canada
ali.karimi2@ucalgary.ca

***Abstract*—Accurate forecasting of short-term electrical load plays a vital role in ensuring the reliability and efficiency of modern power systems. This project presents a day-to-day load forecasting framework that uses both classical statistical methods and advanced deep learning techniques. A comparison was made between a traditional ARIMA model with a long-short-term memory (LSTM) neural network trained on a decade of hourly load and weather data from the Calgary region. The LSTM model incorporates multiple contextual features, including temperature, wind speed, weekends, and holidays, to capture the complex dynamics of energy consumption. Evaluation on both specific days and the entire test set highlights the superior ability of the LSTM model to learn temporal patterns and adapt to variations in demand. These findings reinforce the significance of deep learning in advancing power system forecasting capabilities.**

***Index Terms*—Load forecasting, time series forecasting, long short-term memory (LSTM), ARIMA, deep learning.**

I. Introduction

Day-ahead electricity load forecasting is a fundamental task in power system operations, enabling utilities and grid operators to anticipate next-day energy demand on an hourly basis. Accurate forecasting supports energy scheduling, bidding in electricity markets, optimal resource allocation, and grid reliability. As the integration of renewable energy sources and the volatility of consumption patterns continue to rise, precise day-ahead forecasting becomes increasingly critical for minimizing operational costs and enhancing energy efficiency [1].

Conventional methods for load forecasting, such as the AutoRegressive Integrated Moving Average (ARIMA) model, have been widely used due to their simplicity and interpretability. However, these statistical approaches often struggle to capture the complex, nonlinear, and multi-dimensional patterns found in electricity demand data, particularly when influenced by weather and temporal variables [2]. To address these challenges, recent research has increasingly turned to deep learning techniques, with Long Short-Term Memory (LSTM) networks emerging as a popular choice due to their ability to model long-term dependencies and sequential relationships in time series data [3].

In this project, we focus on developing a forecasting model for day-ahead hourly electricity load prediction in Calgary, Alberta. We use two main approaches: ARIMA as a classical time series benchmark and LSTM to explore the benefits of deep learning for this task. The model utilizes historical load data from the Alberta Electric System Operator (AESO) [4], weather variables obtained from NOAA's integrated surface database [5] [6], and time-based features such as hour of day, day of week, seasonality, and holidays.

While our implementation centers on ARIMA and LSTM, we acknowledge the growing interest in more advanced models—such as CNN-LSTM hybrids and Transformer-based architectures—that have been investigated in the literature and show promising results for multivariate, high-frequency load forecasting tasks [7] [8]. These models represent potential future directions for enhancing accuracy and interpretability in operational forecasting.

By applying and comparing both classical and neural network-based forecasting methods to Calgary's electricity load, this study aims to evaluate performance trade-offs and provide insights into the impact of weather and temporal features on energy consumption. The findings can inform more resilient grid planning and real-time energy management in urban environments.

II. Related Work

The forecasting of electricity-related data, particularly electricity prices and loads, has emerged as a critical area of research due to its significant impact on market operations, resource allocation, and decision-making in the energy sector. This literature review examines the evolution of electricity forecasting methods from traditional statistical approaches to sophisticated deep learning models, highlighting the key advancements and challenges in this rapidly developing field.

A. Evolution of Forecasting Methodologies

The early approaches to electricity forecasting primarily employed traditional statistical methods, including autoregressive integrated moving average (ARIMA), autoregressive moving average (ARIMA), and generalized autoregressive conditional heteroskedasticity (GARCH) models [9], [10]. While these methods provided a foundation for time series forecasting, they were inherently limited by their requirement for stationary data and inability to capture complex nonlinear relationships present in electricity time series. To address these limitations, machine learning approaches gained popularity, with support vector machines (SVM), random forests (RF), and various neural network architectures being increasingly adopted [9], [10]. These methods demonstrated improved forecasting accuracy due to their enhanced robustness and nonlinear mapping capabilities. However, they still lacked the ability to effectively model the temporal dynamics inherent in electricity time series data [10].

B. Deep Learning Approaches

The recognition of these limitations led to the adoption of deep learning methods specifically designed for sequence modeling, most notably recurrent neural networks (RNNs) and their variants such as long short-term memory (LSTM) networks and gated recurrent units (GRUs) [9], [11]. These architectures were better suited to handle the sequential nature of time series data and could effectively capture both short-term and long-term dependencies within the data. LSTM networks, in particular, have gained significant attention for electricity price forecasting due to their ability to address the vanishing gradient problem that plagues standard RNNs, allowing them to learn long-term dependencies more effectively [10], [11]. Several studies have demonstrated the superior performance of LSTM-based models compared to traditional statistical and machine learning methods for electricity price forecasting tasks [11]. Recent advancements have seen the emergence of hybrid approaches that combine different models and techniques to further enhance forecasting accuracy. For instance, Wang and Luo (2017, cited in [10]) proposed a hybrid model based on two-layer decomposition and neural networks optimized by firefly algorithms for electricity price forecasting.

C. Transformer-Based Models

A significant development in time series forecasting has been the application of Transformer-based architectures, which have demonstrated remarkable success in natural language processing and computer vision tasks [12]. The Transformer architecture, originally proposed by Vaswani et al. (2017), replaces the recurrent components of traditional sequence models with self-attention mechanisms that can efficiently model dependencies across time steps. Several Transformer-based models have been proposed for electricity forecasting, including Informer [13], Autoformer [11], FEDformer [13], and Pyraformer [12], [13]. These models have introduced various modifications to the original Transformer architecture to better suit the characteristics of time series data and address challenges specific to electricity forecasting. However, the efficacy of Transformer-based models for time series forecasting has been questioned by some researchers. In [13], the authors argue that while Transformers excel at extracting semantic correlations in sequence data, their permutation-invariant self-attention mechanism makes them less suitable for modeling the temporal relations in time series data, where order is crucial. Their study suggests that simpler linear models can outperform sophisticated Transformer-based models on various time series forecasting benchmarks, including electricity data.

III. Dataset

The dataset used for model training in this project comprises three main components: electricity load data, weather-related features, and time-based variables. This section provides a detailed overview of the data sources and the specific features included.

A. Load Data

Historical electricity load data is obtained from the Alberta Electric System Operator's (AESO) website [4], which provides hourly load measurements (in megawatts, MW) for areas and regions across Alberta from 00:00 on January 1, 2011, to 23:00 on December 31, 2024. In this project, only the load data corresponding to the Calgary region is utilized.

B. Weather Data

Weather data is sourced from the US National Centers for Environmental Information (NCEI) affiliated with the National Oceanic and Atmospheric Administration (NOAA) [5], which maintains a global climate monitoring database. Records from five climate observation stations located in Calgary are collected, covering the same time period as the load data. The dataset includes a wide range of meteorological variables, such as temperature, dew point, wind speed, wind type, wind direction, and more. For this project, only hourly temperature and wind speed are extracted, as they are identified as the primary weather factors influencing electricity load [6]. The raw weather data underwent preprocessing to ensure consistency and completeness. Only hourly entries valid under the Integrated Surface Data (ISD) format [14] were retained. Temperature and wind speed values were parsed and converted to degrees Celsius and meters per second, respectively. For each hour, the average across the five Calgary stations was computed. Missing timestamps were addressed by linear interpolation for gaps under 24 hours, while larger gaps were imputed using historical records from Weather Spark [15], a secondary low-resolution climate source.

C. Time Data

In addition to load and weather data, the final dataset includes time-based features to capture temporal load patterns. Each data point has an hourly timestamp from 2011 to 2024, along with binary indicators for weekends and statutory holidays in Alberta. These features help account for variations in electricity usage between weekdays, weekends, and holidays. The resulting dataset is a time series with hourly entries, where each row includes the load (MW), average temperature ($^{\circ}C$), average wind speed (m/s), and the aforementioned temporal indicators.

IV. Methodology[1]

A. Baseline model: ARIMA

As a classical time-series forecasting approach, the AutoRegressive Integrated Moving Average (ARIMA) model was employed as the baseline. ARIMA models predict future values using a linear combination of past values (autoregression), the differencing of observations (integration), and past forecast errors (moving average). Univariate ARIMA model was used and trained only on historical hourly electricity load data. To capture short-term temporal dependencies and daily patterns, we used a 24-hour input window for each forecast. The ARIMA model parameters (p, d, q) were empirically chosen as (2, 1, 2) after testing a few common configurations. The model was re-fitted for each target forecast date to avoid information leakage. Despite its simplicity, ARIMA provides a useful benchmark against more complex models like LSTM, particularly regarding interpretability and computational efficiency [2].

B. LSTM model design

To capture nonlinear patterns and dependencies in the data, we designed a deep learning model based on the Long Short-Term Memory (LSTM) architecture. LSTMs are a type of recurrent neural network (RNN) effective for sequence modeling tasks due to their ability to retain long-term dependencies [3]. The model employs a sequence-to-one structure where the input is a 24-hour window of multivariate features, and the output is the subsequent 24-hour forecast of electricity load. Input features, derived from a 10-year hourly dataset, include electricity load in the Calgary region, temperature, wind speed, and three temporal indicators extracted from the date: weekend status, weekday information (as a numerical value), and holiday flag. The LSTM network comprises two stacked LSTM layers, each with 64 hidden units, followed by a fully connected layer. The final output layer predicts 24 hourly load values simultaneously. Training utilized the Adam optimizer with a learning rate of 0.001 and mean squared error (MSE) loss function. The model was trained for 30 epochs with early stopping based on validation set performance to monitor overfitting. All input features were normalized using MinMax scaling to ensure stable training. The dataset was split into training (70%), validation (15%), and testing (15%) sets. Although no explicit dropout layers were used in this version, future iterations could incorporate dropout to mitigate overfitting [7]. Hyperparameters such as learning rate, number of hidden units, and LSTM layers were selected based on prior experience and initial experimentation. Future work may involve systematic tuning using methods like grid search or Bayesian optimization [16].

[1]The source code is available at https://github.com/NoureldinAmer/An-LSTM-Based-Approach-to-Day-Ahead-Electricity-Load-Forecasting

V. Results and Discussion

A. Overall model performance

The LSTM model was implemented and trained on Google Colab using GPUs. Both training and validation loss decreased significantly during the initial epochs, with the validation loss stabilizing around 0.0011 after approximately 25 epochs, indicating that the model had effectively converged without overfitting as observed from the training and validation loss curves over the epochs Figure 1. To evaluate performance, the

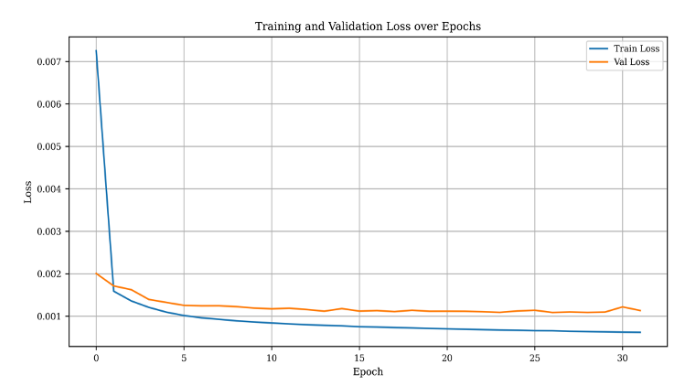

Fig. 1. Training and validation loss across approximately 30 epochs.

LSTM model was benchmarked against a traditional ARIMA model over the test dataset spanning from November 25, 2022, to December 31, 2024. As detailed in Table I, the LSTM model achieved a Mean Absolute Percentage Error (MAPE) of 2.57%, significantly outperforming the ARIMA model, which recorded a MAPE of 16.45%. This represents an 84.38% reduction in prediction error, underscoring the LSTM's superior ability to capture complex temporal dependencies in the data. Further, the LSTM model yielded much lower values in both Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)—31.62 and 38.74, respectively—compared to ARIMA's 188.90 (MAE) and 224.99 (RMSE). These metrics confirm the LSTM model's robustness and accuracy in forecasting tasks, making it a compelling alternative to traditional statistical approaches in time series prediction.

TABLE I
Comparison of overall model performance metrics for LSTM and ARIMA forecasting models.

Model / Metrics	MAE (MW)	RMSE (MW)	MAPE (%)
LSTM	31.62	38.74	2.57
ARIMA	188.90	224.99	16.45

B. Detailed LSTM performance analysis

To better investigate the LSTM model's predictive behavior, we evaluated its performance across different days of the week using the forecasting results in 2024. Figure 2 displays the MAPE for each weekday, alongside a red reference line indicating the overall average MAPE of 2.57%. As summarized in Table II, the LSTM model achieves the lowest MAPE values on Friday (2.31%), Saturday (2.35%), and Thursday (2.35%), indicating more consistent performance on these days. In contrast, Sunday (3.08%) and Monday (2.99%) exhibit the highest MAPE, suggesting slightly reduced model accuracy at the beginning and end of the week. The asymmetric performance over the week suggests that the model struggles to capture certain transitional points at the boundaries of the weekly cycle, particularly on Sundays and Mondays. This indicates a limitation in the model's ability to adapt to the variability observed at the start and end of the week. Despite the inclusion of day-of-week as a feature, the model may be implicitly treating temporal patterns as more uniform than they actually are. To address this, future improvements could involve representing the day-of-week as a categorical input using techniques such as one-hot encoding or embedding layers, which may enable the model to better differentiate day-specific behaviors and improve overall forecast accuracy. Beyond the weekly pattern, the model performs poorly on holidays, with an average MAPE of 7.76% on Alberta's 2024 statutory holidays, compared to 2.57% overall. This suggests the model struggles with the irregular behavior of holidays—such as altered schedules or demand drops—likely due to their limited representation in the dataset and high variability. Future improvements could explore these assumptions and adjust the model accordingly

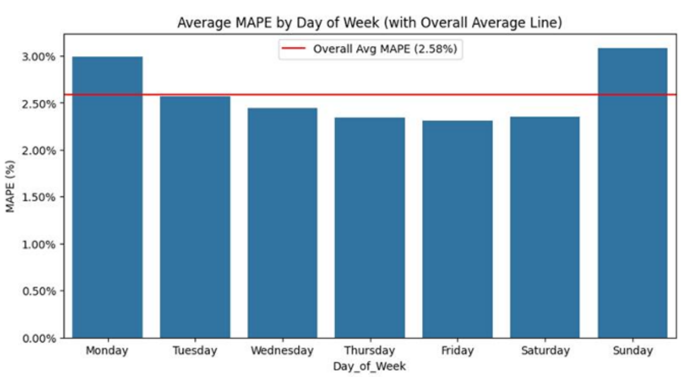

Fig. 2. Average MAPE by day of week in 2024.

C. LSTM performance on special occasions

To further evaluate the model's robustness under varying conditions, we analyzed its forecasting performance on a set of representative special days, as visualized in Figure 2. These include weather-related anomalies (e.g., major temperature drop, high wind, extreme cold), calendar effects (e.g., public holidays, long weekends), and typical day types (weekdays and weekends). Across all cases, the LSTM model consistently

TABLE II
MAPE BY DAY OF WEEK AND AGGREGATED BY WEEKDAY AND WEEKEND.

Day of Week	MAPE by Day of Week (%)	Day Type	MAPE by Day Type (%)
Monday	2.99	Weekday	2.49
Tuesday	2.57		
Wednesday	2.44		
Thursday	2.35		
Friday	2.31		
Saturday	2.35	Weekend	2.78
Sunday	3.08		

outperforms the ARIMA baseline, demonstrating lower MAE and MAPE values and a closer fit to the actual load profiles. Notably, the LSTM model handles extreme weather conditions

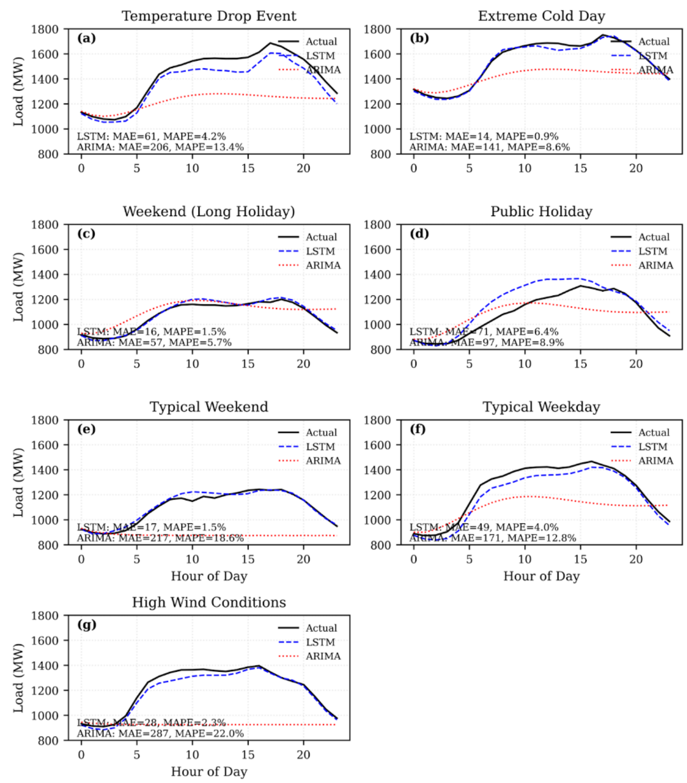

Fig. 3. Forecasting performance of LSTM and ARIMA models on special days, including weather events, holidays, and typical day types.

relatively well—for instance, during the extreme cold day, it achieves a MAPE of only 0.9%, compared to ARIMA's 8.6%. Similarly, on a high wind day, the LSTM maintains reasonable accuracy (2.3% MAPE), while ARIMA underperforms significantly (22.0% MAPE). However, performance tends to degrade under calendar-related disruptions, particularly on public holidays, where LSTM records a MAPE of 6.4%, noticeably higher than typical weekday or weekend values. This supports earlier findings that holiday patterns are less predictable and harder for the model to learn, likely due to low representation and high behavioral variability. Interestingly, on long holiday weekends, the LSTM performs exceptionally well (1.5% MAPE), suggesting that when holiday behavior mimics regular weekend trends, the model adapts more effectively. Typical weekends and weekdays also show low errors (1.5% and 4.0% MAPE, respectively), reinforcing the model's strength in learning regular temporal cycles. Overall, these case studies confirm that the LSTM model is capable of capturing non-linear patterns under both typical and atypical conditions. Yet, its performance remains sensitive to rare or irregular events, indicating room for improvement through the inclusion of richer temporal and contextual features.

VI. Conclusion

This study presents a comparative analysis of short-term electricity load forecasting using both classical and deep learning approaches, demonstrating the superior performance of the LSTM-based model in capturing complex temporal patterns. By integrating multivariate features—including historical load, weather data (temperature and wind speed), and calendar-based indicators (weekends and holidays)—the LSTM model effectively learns nonlinear dependencies and adapts to real-world consumption behavior. Across the full test set, the LSTM model achieved a mean absolute percentage error (MAPE) of 2.57%, significantly outperforming the ARIMA baseline, which recorded a MAPE of 16.45%. The LSTM model also demonstrated strong robustness under both typical and atypical conditions, such as extreme weather events and long holiday weekends. However, its performance declined during public holidays and at weekly transition points (e.g., Sundays and Mondays), highlighting areas where further model refinement is needed. While classical models such as ARIMA remain valuable for their interpretability and simplicity, they fall short in handling the complex and dynamic nature of electricity load data. In contrast, the deep learning-based approach provides a more flexible and scalable solution for high-resolution forecasting tasks.

References

[1] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207015001508

[2] C. Tarmanini, N. Sarma, C. Gezegin, and O. Ozgonenel, "Short term load forecasting based on arima and ann approaches," *Energy Reports*, vol. 9, pp. 550–557, 05 2023.

[3] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches †," *Energies*, vol. 11, p. 1636, 06 2018.

[4] Alberta Electric System Operator (AESO), "Hourly load by area and region," https://www.aeso.ca/market/market-and-system-reporting/data-requests/hourly-load-by-area-and-region, 2024, accessed: 2025-04-04.

[5] NOAA National Centers for Environmental Information, "Hourly weather observations map," https://www.ncei.noaa.gov/maps/hourly/, 2024, accessed: 2025-04-04.

[6] NOAA, "Integrated surface data (isd) format documentation," https://www.ncei.noaa.gov/data/global-hourly/doc/isd-format-document.pdf, 2024, accessed: 2025-04-04.

[7] K. Wu, J. Wu, L. Feng, B. Yang, R. Liang, S. Yang, and R. Zhao, "An attention-based cnn-lstm-bilstm model for short-term electric load forecasting in integrated energy system," *International Transactions on Electrical Energy Systems*, vol. 31, no. 1, p. e12637, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/2050-7038.12637

[8] B. Lim, S. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207021000637

[9] J. Lago, G. Marcjasz, B. D. Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Applied Energy*, vol. 293, p. 116983, Jul. 2021, arXiv:2008.08004 [stat]. [Online]. Available: http://arxiv.org/abs/2008.08004

[10] X. Xiong and G. Qing, "A hybrid day-ahead electricity price forecasting framework based on time series," *Energy*, vol. 264, p. 126099, Feb. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0360544222029851

[11] Y. Li, Y. Ding, Y. Liu, T. Yang, P. Wang, J. Wang, and W. Yao, "Dense Skip Attention Based Deep Learning for Day-Ahead Electricity Price Forecasting," *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4308–4327, Sep. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9931505/

[12] O. Llorente and J. Portela, "A Transformer approach for Electricity Price Forecasting," Mar. 2024, arXiv:2403.16108 [cs]. [Online]. Available: http://arxiv.org/abs/2403.16108

[13] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are Transformers Effective for Time Series Forecasting?" *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 9, pp. 11 121–11 128, Jun. 2023. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/26317

[14] M. U. Fahad and N. Arbab, "Factor affecting short term load forecasting," *Journal of Clean Energy Technologies*, vol. 2, no. 4, pp. 305–309, 1 2014. [Online]. Available: https://doi.org/10.7763/jocet.2014.v2.145

[15] WeatherSpark, "Weatherspark: Beautiful weather graphs and charts," https://weatherspark.com, 2024, accessed: 2025-04-04.

[16] M. A. Alotaibi, "Machine learning approach for short-term load forecasting using deep neural network," *Energies*, vol. 15, no. 17, 2022.