Optimizing Short-Term Electrical Load Forecasting with Bi-LSTM and Advanced Temporal Encoding

Dmitrii Vasenin[b], Marco Pasetti[b], Savvin Nikita[b], Davide Astolfi[b]

[a]*Universit'a degli Studi di Brescia, Department of Information Engineering, Brescia, , , Italy*
[b]*Universit'a degli Studi di Brescia, Department of Information Engineering, Brescia, , , Italy*

Abstract

This paper introduces a novel approach to short-term electrical load forecasting that emphasizes the integration of calendar features and a unique method for temporal encoding, while critically examining the limited influence of weather data on forecasting accuracy. Through our analysis, we found that weather variables have minimal impact on improving forecast precision. To address this, we proposed an innovative method that enhances the model's understanding of time by encoding minute, hour, day of the week, week of the month, and year using sine and cosine transformations.

Our approach involves the application of advanced machine learning architectures, including LSTM (Long Short-Term Memory), Bi-LSTM (Bidirectional Long Short-Term Memory), and hybrid models such as CNN-LSTM (Convolutional neural network - Long Short-Term Memory) and CNN-Bi-LSTM (Convolutional neural network - Bidirectional Long Short-Term Memory), to predict electrical load. The forecasting process begins with the collection and normalization of load and weather data. Although weather data was initially considered, its negligible effect led us to focus on temporal features and their encoding.

Among the tested models, the Bi-LSTM model demonstrated the highest accuracy, particularly when combined with our temporal encoding technique. This method ensures that the model accurately captures the patterns inherent in different time scales, leading to precise and reliable load forecasts. Our findings highlight the importance of temporal feature engineering and suggest that while weather data may not significantly enhance forecasting accuracy, the careful encoding of time-related features is crucial for improving predictive performance.

Keywords: Electricity consumption forecasting, Time series models, Machine learning, Hybrid approaches

1. Introduction

In recent years, the rapid expansion and growth of electricity networks, coupled with the emergence of new types of electrical loads, have necessitated a thorough analysis of energy consumption patterns and forecasting models Almihat et al. (2022). The complexity of designing electrical system components is increased by the growth of electrical loads, and the reorganization of the energy system has led to the establishment of specialized generation, transmission, and distribution companies Das et al. (2018). These entities face challenges in meeting the increasing requirements for the reliable operation of power system networks. Traditional methods of predicting electrical loads are becoming inadequate due to various factors affecting electrical loads both directly and indirectly, such as population growth, temperatures, climate changes, economic systems, human behavior, and industrial developments Avtar et al. (2019). Accurate electricity load forecasting is essential for ensuring reliable and economic planning, control, and operation of power systems, helping electricity companies make critical decisions regarding power generation, transmission, and distribution infrastructure Nadtoka and Balasim (2015); Adedeji et al. (2019); Apadula et al. (2012).

With the escalating demand for energy and the complexity of developing reliable load prediction models Kabeyi and Olanrewaju (2022); Raza and Khosravi (2015), the importance of enhancing these models is becoming increasingly apparent. Electricity load forecasting is not only critical for the power sector but also beneficial for all economic sectors in preparing future development plans. Load forecasting helps in designing electrical networks and formulating strategic plans that contribute to a stronger economy, a cleaner environment, and energy sustainability Abumohsen et al. (2023).

Electricity consumption forecasting is an essential task for efficient and optimal distributed energy management in the electricity sector. Accurate and timely forecasts help in planning and scheduling power generation, transmission, and distribution, which ultimately leads to better decision-making and resource management. Various forecasting techniques have been proposed and applied to predict electricity consumption, ranging from traditional time series models to advanced machine learning methods.

This paper presents a novel approach to short-term electrical load forecasting by integrating calendar features and employing an advanced method for temporal encoding. Our study critically assesses the impact of weather data on forecasting accuracy, revealing its limited influence on precision. To overcome this limitation, we introduce an innovative encoding technique that enhances the model's temporal understanding by representing minute, hour, day of the week, week of the month, and year through sine and cosine transformations.

The remainder of this paper is organized as follows: Section 2 presents a literature review of electricity load forecasting methods and previous studies. Section 3 provides a discussion of the strengths and weaknesses of the reviewed techniques, as well as their implications for future research. Section 4 describes the methodology employed in building and comparing the selected forecasting models. Section 5 discusses the experimental results and compares them with previous studies. Finally, Section 6 concludes the paper and outlines future work.

2. Related work

The electricity generation process is predominantly reliant on fossil fuels, which consequently leads to an increase in demand for such fuels with a rise in electricity consumption. In this context, the present study (Zolfaghari and Sahabi, 2019) centers on the analysis of Iran's electricity market during the period spanning from January 1, 2013, to March 2, 2018, considering the impact of these factors on global warming.

2.1. Traditional Time Series Models

Traditional time series models, such as Auto regressive Integrated Moving Average (ARIMA) and its variants, have been widely used in electricity consumption forecasting Alberg and Last (2018); Hyndman and Khandakar (2008). These models are based on the assumption that the historical data patterns can be used to predict future values. Some studies have incorporated external factors, such as weather and calendar data, to improve the performance of these models. For example, the study by Zolfaghari and Sahabi (2019) combined wavelet decomposition with ARIMA-GARCH models to predict electricity consumption in Iran's electricity market. The authors considered the calendar (month, day of the week, and special days) as independent variables in their ARIMA model and achieved accurate predictions for a 60-day forecast horizon.

In Van der Meer et al. (2018), two models were employed to forecast electricity consumption in a residential area. The first model was based on the AutoRegressive Integrated Moving Average (ARIMA) approach, utilizing the auto.arima function from the forecast package. The second model, also presented in this study, used Gaussian Processes (GPs) as a nonparametric model to generate probabilistic forecasts for a time horizon of 30 minutes. Both models, including the ARIMA benchmark, were applied for forecasting the half-hour resolution residential load in a single household.

2.2. Machine Learning Techniques

Machine learning techniques have gained popularity in electricity consumption forecasting due to their ability to model complex nonlinear relationships and handle large datasets. The most commonly used machine learning methods include artificial neural networks (ANNs), support vector machines (SVMs), and extreme learning machines (ELMs). ""For instance, Theile et al. (2018) compared the performance of Support Vector Machines (SVMs) and Recurrent Neural Networks (RNNs) for short-term electricity consumption forecasting in France. The study showed that RNNs outperformed SVMs in terms of prediction accuracy, especially during peak demand periods.""

In Dmitri et al. (2016), regression analysis was used to forecast daily power consumption by selecting the functional form of the regression equation and determining the crucial independent variables. The work Theile et al. (2018) employed the Recurrent Neural Network (RNN) technique for accurate energy forecasting. The special property of RNNs lies in their ability to adapt to time-series forecasting problems. The study also considered other machine learning techniques such as Support Vector Machines (SVM) that were computationally more efficient than RNNs for larger datasets. However, for accurate energy forecasting with the given training dataset, RNNs were found to be more suitable than SVMs.

Notably, the study also considered other machine learning techniques such as Support Vector Regression (SVR), which is a variant of SVM specifically designed for regression tasks. SVR was found to be computationally very efficient for larger datasetsAli and Yaman (2013). This efficiency in computational processing makes SVR particularly suitable for handling substantial amounts of data. The notable efficiency of SVR in computational processing renders it well-suited for handling substantial amounts of data. This advantage, coupled with factors such as reduced computation time, ease of implementation, and high accuracy, positions SVR as a compelling choice for short-term load forecasting applications. The study's Alrashidi and Qamar (2023), Gao et al. (2023) findings underscore the practicality and effectiveness of employing SVR in the context of electricity consumption forecasting.

2.3. Hybrid Approaches

Hybrid approaches combine the strengths of multiple forecasting techniques to achieve better accuracy and robustness. These methods typically involve the integration of traditional time series models with machine learning techniques, or the combination of multiple machine learning methods. An example of a hybrid approach is the study by Wang and Meng (2012), which proposed a hybrid model combining ARIMA and ANN for energy consumption forecasting in China. The authors found that the hybrid model outperformed both the ARIMA and ANN models individually, with a lower mean absolute percentage error (MAPE).

In the study Naz et al. (2019), load forecasting techniques were categorized into three main groups: data-driven, classical, and Artificial Intelligence (AI). The results showed that the Enhanced Extreme Learning Machine (ELM) with optimization outperformed Convolutional Neural Network (CNN) and Linear Regression (LR) in terms of prediction accuracy. Additionally, the Enhanced Logistic Regression (ELR) and Enhanced Recurrent Extreme Learning Machine (ERELM) were employed for both datasets, and ERELM was found to outperform ELM and RELM.

In Zhu et al. (2021), a novel EMD-Fbprophet-LSTM model was proposed for short-term electricity consumption forecasting. The model integrates the Long Short-Term Memory (LSTM) method to predict short-term electricity consumption and the Fbprophet time-series model to predict electricity consumption based on the decomposition of electricity consumption by Empirical Mode Decomposition (EMD). The EMD-Fbprophet-LSTM model is designed to account for the unique fluctuating characteristics of electricity demand on different time scales and the impact of weather and other factors on customer demand. As a more advanced hybrid model, it provides lower forecast errors and better results when forecasting short-term electricity consumption of customers. The EMD-Fbprophet-LSTM prediction model is effective in predicting data with strong holiday effects and seasonal fluctuations, and can handle outliers well, leading to improved prediction accuracy and reduced errors.

The proposed integrated electricity demand strategy in Eseye et al. (2019) is based on the hybridization of the Hilbert-Huang Transform (HHT), Regular Particle Swarm Optimization (Reg-PSO), and Adaptive Neuro-Fuzzy Inference System (ANFIS), demonstrates its effectiveness in accurately forecasting 24-hour ahead building electricity demand. The strategy utilizes historical data for model development and achieves high prediction accuracy when tested with a one-year testing dataset. The proposed approach outperforms other evaluated forecasting models, yielding superior results across different building types. Moreover, the model exhibits efficient execution time, making it suitable for practical applications. These findings validate the capability and suitability of the devised integrated approach for short-term electricity demand prediction in building energy systems.

Zhang et al. (2023) introduces a hybrid deep learning framework, CNN-LSTM, based on multi-task learning, convolutional neural networks, and long short-term memory, demonstrating its competitiveness and high performance in short-term and medium-term power forecasting compared to other models.

2.4. *Grey Prediction Models*

Grey prediction models, based on Grey System Theory, have been proposed as an alternative to traditional time series models and machine learning methods. These models are designed to handle small and incomplete datasets, making them suitable for situations where data availability is limited. In the study by Thành (2019), a Grey Model (GM) was employed to forecast electricity consumption in Vietnam. The results indicated that the Fourier Residual Modified GM model achieved satisfactory accuracy, even with a small sample size and limited historical data.

In the paper Bahrami et al. (2014), the proposed model combined Wavelet Transformation (WT) and GM, which was enhanced by Particle Swarm Optimization (PSO). Model inputs included weather data such as mean temperature, average relative humidity, mean wind speed, and load based on past days. Hybrid models that incorporate WT with other techniques are widely used for short-term load forecasting.

2.5. *Incorporating Weather, Calendar, and Occupancy Data*

The accuracy of electricity consumption forecasting models can be significantly influenced by the incorporation of external factors such as weather, calendar, and occupancy data. For example, the study by Nadtoka and Balasim (2015) explored the impact of weather data, including temperature and natural illumination. The paper contributes to the literature on short-term electricity consumption forecasting by demonstrating the potential benefits of incorporating meteorological factors and using SVM with PSO for model development. The paper provides a comparison between two popular machine learning techniques and algorithms, for day-ahead electricity consumption forecasting. This comparison can help researchers and practitioners understand the trade-offs between accuracy and computational efficiency when choosing a model.Time,day and week are considered, Type of day (working day or holiday) is designed as a feature of the models Theile et al. (2018).The paper analysis of the relevance of various types of data (weather, indoor ambient, calendar, and building occupancy) for building load forecasting Massana et al. (2015). Alrashidi and Qamar (2023) asserts that utilizing the Support Vector Regression (SVR) algorithm with specific features as: air temperature, cloud capacity, global horizontal irradiance, relative humidity, surface pressure, wind direction and wind speed yields the most precise predictions for factory load forecasting. Development of a simple and economical method for building load forecasting with high accuracy and low computational cost. The authors found that incorporating weather data led to a substantial improvement in the prediction accuracy of models.

Similarly, Zolfaghari and Sahabi (2019) demonstrated the importance of calendar data in forecasting electricity consumption. The authors used a combination of daily and hourly calendar data, such as holidays, weekends, and time of day, to enhance the performance of an NeuroWavelet and ARIMAX-GARCHmodels model for predicting electricity consumption in Iran. The results showed that the inclusion of calendar data improved the model's accuracy and generalization ability.

Occupancy data, which provides information about the number of people present in a building or area, can also play a critical role in electricity consumption forecasting. In the study by Massana et al. (2015); Eseye et al. (2019); Abdulrahman et al. (2022), occupancy data was used as an input to forecasting models for different type of buildings. The authors observed that incorporating occupancy data led to a significant improvement in the model's prediction accuracy, particularly during peak demand periods.

2.6. *Feature selection and preprocessing*

In the field of short-term load forecasting, data gaps caused by missing or incomplete data present substantial challenges. These gaps can arise due to a range of factors, including communication failures, measurement errors, and equipment malfunctions. However, filling these data gaps is crucial for achieving accurate and reliable load forecasts, as it enables a more comprehensive understanding of the underlying system dynamics.

To tackle this issue effectively, this paper explores various techniques and methodologies from existing literature in the area of feature selection and preprocessing. These techniques aim to address the challenge of data gaps by incorporating approaches such as imputation, where missing or incomplete data is estimated or completed using available information. By investigating the existing literature, we aim to identify the most effective methods for handling data gaps and improving the reliability of load forecasts.

Data Preprocessing: To address the data gaps in short-term load forecasting, we employ a combination of interpolation techniquesHyman (1983); Huang (2021), data imputation algorithms Borges et al. (2020), statistical methods Makridakis et al. (2018), and introduce the use of a generative adversarial network (GAN) Hammad Alharbi and Kimura (2020).

Interpolation techniques, such as linear interpolation and spline interpolation, are utilized to estimate missing values by considering the surrounding observed data points. These techniques interpolate values based on the trends and patterns observed in the available data..

Data imputation algorithms, including k-nearest neighbors (KNN)Dong et al. (2021) and expectation-maximization (EM)Dempster et al. (1977); Jeong et al. (2021), are employed to infer missing values by leveraging the patterns and similarities present in the available data. KNN imputation involves estimating missing values by considering the values of the nearest neighboring data points, while EM imputation utilizes an iterative approach to estimate missing values based on the expectation-maximization algorithm.

In addition, statistical methods such as seasonal decompositionAhmad (2017) and time series decompositionPrema and Rao (2015); Mbuli et al. (2020) are applied to extract relevant features and trends from the dataset. These methods aid in identifying patterns and structures in the data, enabling more effective handling of data gaps.

Furthermore, to enhance the data preprocessing process, we introduce the use of a generative adversarial network (GAN)Hammad Alharbi and Kimura (2020). GANs consist of two components: the generator and the discriminator. The generator learns to generate plausible data, including filling in missing values, while the discriminator learns to distinguish between the generator's fake data and real data. Through an adversarial training process, the generator continuously improves its ability to produce realistic and plausible data, which can help in generating accurate estimations for the missing values.

By employing a combination of interpolation techniques, data imputation algorithms, statistical methods, and introducing the use of a GAN, we aim to effectively handle data gaps and improve the reliability and accuracy of short-term load forecasting models. These preprocessing techniques allow us to generate complete and consistent datasets for further analysis and modeling.

Handling Outliers: Outliers, often present in load data, can significantly affect the accuracy of short-term load forecasts. Robust statistical techniques, such as median absolute deviation (MAD)Leys et al. (2013) and Hampel filterLiu et al. (2004), are employed to identify and mitigate the influence of outliers. These methods effectively reduce the impact of extreme values, resulting in a more accurate and robust dataset for load forecasting.

Data Fusion: To further enhance the completeness and reliability of the dataset, we explore data fusion techniques that integrate data from multiple sources. This fusion process allows for cross-validation and data redundancy elimination, minimizing the effects of missing or erroneous data pointsXie et al. (2022).

In summary, this section delves into the importance of addressing data gaps in short-term load forecasting. It highlights the various reasons behind these gaps and emphasizes the need for comprehensive techniques and methodologies to fill them. By exploring the existing literature, we aim to identify effective approaches for feature selection and preprocessing that can enhance the accuracy and reliability of load forecasts.

3. Data Analysis

3.1. Data Description

3.1.1. Load Consumption Data

The dataset utilized in this study was sourced from the Energy Management System (EMS) records of an electric power station located on the engineering campus of the University of Brescia, Italy. This facility supplies power to a public building that includes a variety of amenities, such as departmental offices, classrooms, student dormitories, study halls, a cafeteria, a gym, and a baseball field. The data was collected at a resolution of 5-minute intervals, following a systematic recording method.

3.1.2. Weather Data

To enhance the accuracy of our electrical load forecasting model, we incorporated external weather parameters as additional data inputs. Weather data was collected every minute by the Weather Station at the University of Brescia. The following weather parameters were selected for their potential impact on electrical load:

- **Solar Irradiance:** Measures the power per unit area received from the Sun in the form of electromagnetic radiation.
- **Ambient Temperature:** The temperature of the air in the surrounding environment, which can influence heating and cooling demands.
- **Atmospheric Pressure:** The force exerted by the weight of the atmosphere, affecting weather patterns and potentially energy consumption.
- **Dew Point:** The temperature at which air becomes saturated with moisture and dew forms, indicating humidity levels.
- **Heat Index:** A measure of how hot it feels when relative humidity is factored in with the actual air temperature.
- **Rain Rate:** The amount of rainfall over a specific period, impacting outdoor activities and energy usage.
- **Relative Humidity:** The percentage of moisture in the air relative to the maximum moisture the air can hold at that temperature.

- **UV Index:** A measure of the strength of ultraviolet radiation from the Sun, which can influence energy use for cooling.
- **Wind Chill:** The perceived decrease in air temperature felt by the body on exposed skin due to wind.
- **Wind Speed Average:** The average speed of wind over a set period, affecting wind power generation and heating requirements.
- **Wind Direction Average:** The predominant direction from which the wind is blowing over a set period.

By integrating these diverse weather parameters, we aimed to capture the various environmental factors that could influence electrical load, thereby improving the forecasting accuracy of our model.

3.2. Checking for correlations between electrical consumption and weather/calendar data:

3.2.1. Graphical Analysis

Scatter Plots: Construct scatter plots between electrical consumption and each of the weather and calendar variables. Pair Plots: Utilize pair plots to visualize relationships between multiple variables concurrently.

3.2.2. Pearson Correlation Coefficient

This is one of the most popular methods for examining linear relationships between two continuous variables.

Value: Ranges from -1 (perfect inverse correlation) to +1 (perfect direct correlation), with 0 indicating no linear relation. Hypothesis: H0 (null) assumes no correlation, whereas H1 (alternative) indicates correlation.

3.2.3. Spearman Correlation Coefficient

This method is employed to measure the degree of monotonic relation between variables and is especially useful when the data is ordinal or not normally distributed.

3.2.4. Kendall's Tau Correlation Coefficient

This also measures the degree of monotonic relation between two variables and is often utilized for categorical data.

3.2.5. Cross-Correlation

Time series are often analyzed using cross-correlation to explore the dependency of signals considering time lags.

3.2.6. Correlation Heatmap

Constructing a correlation heatmap among all variables allows visualization and quick assessment of the degree of relations among them.

3.2.7. Partial Correlation

This analysis allows exploring the correlation between two variables while controlling for the effect of a third variable.

3.2.8. Multiple Regression Analysis

Employing multiple regression allows assessing the impact of several independent variables on the dependent variable.

Significance Assessment: Besides calculating the correlation coefficient, it is also vital to assess its statistical significance (p-value). A p-value ¡ 0.05 is often used as a threshold to determine statistical significance. Important Notes: Correlation does not imply causation. Ensure the data is free from outliers or anomalies before conducting correlation analysis, as it is sensitive to such values. Consider utilizing the aforementioned methods to analyze your data and understand which variables have the most impact on electrical consumption and may be employed in forecasting models.

4. Methodology

4.1. Difference Between In-Sample and Out-of-Sample Forecasting

Forecasting in time series analysis involves predicting future values based on past observations. Two key approaches to evaluating a forecasting model's performance are in-sample forecasting and out-of-sample forecasting. Understanding these approaches is essential for selecting the appropriate method for specific applications.

4.1.1. In-Sample Forecasting

Definition. In-sample forecasting involves simulating the model's output on a portion of the dataset that has been used for training. The model is trained and tested on different portions of the same dataset.

Example. If you have data from January to December 2023, you might train the model on data from January to August 2023 and simulate the output on September to December 2023. Here, the test data is part of the same dataset used for training.

Characteristics.

- **Training Data:** The model is trained on part of the dataset (e.g., January to August 2023).
- **Testing Data:** The model is tested on another part of the same dataset (e.g., September to December 2023).
- **Performance Measurement:** The accuracy is measured within the same dataset, which can sometimes lead to overfitting, meaning the model may perform well on the test set but not generalize to new data.

4.1.2. Out-of-Sample Forecasting

Definition. Out-of-sample forecasting involves training the model on one dataset and simulating the output on a separate, unseen dataset. This approach simulates a real-world scenario where the model predicts future data not used during training.

Example. Using the same dataset from January to December 2023, you might train the model on data from January to September 2023 and forecast the load for October to December 2023. The validation dataset (October to December 2023) has not been seen by the model during training.

Characteristics.

- **Training Data:** The model is trained on a historical dataset (e.g., January to September 2023).
- **Validation Data:** The model is validated on future data that was not part of the training process (e.g., October to December 2023).
- **Performance Measurement:** This method provides a realistic assessment of the model's ability to generalize to new, unseen data, reducing the risk of overfitting.

4.1.3. Summary of Differences

- **Dataset Usage:**
  - **In-Sample Forecasting:** Uses different portions of the same dataset for training and testing.
  - **Out-of-Sample Forecasting:** Uses one dataset for training and an unseen dataset for validation.
- **Objective:**
  - **In-Sample Forecasting:** Evaluates how well the model fits the data it was trained on.
  - **Out-of-Sample Forecasting:** Evaluates how well the model predicts new, unseen data.
- **Evaluation:**
  - **In-Sample Forecasting:** Can lead to overfitting if the model is too closely tailored to the training data.
  - **Out-of-Sample Forecasting:** Provides a better indication of the model's performance in real-world scenarios by testing on truly unseen data.

4.2. Description of the Model

The model is designed to predict a single time step ahead based on a lag of historical data, specifically 12 points in our case 1. To achieve this, we denote the historical data sequence as:

$$X = \{x_1, x_2, \ldots, x_{12}\}. \tag{1}$$

where each x_i represents a vector of features including primary parameters and weather-related variables.

The trained model, which has been exclusively trained on data excluding the date we intend to forecast (out-of-sample), generates a single prediction $\hat{y}_{13}$, representing the next point in the time series.

Subsequently, this predicted point $\hat{y}_{13}$ is appended to the sequence of historical values, and define as follows:

$$X' = \{x_2, x_3, \ldots, x_{12}, \hat{y}_{13}\} \tag{2}$$

The oldest point x_1 in the sequence X' is then removed, maintaining a window of 12 points—11 actual values and 1 generated value. This updated sequence X' is fed back into the model to obtain the next prediction $\hat{y}_{14}$.

This iterative process continues, where each forecast $\hat{y}_t$ is generated based on the updated sequence X', until we accumulate enough predictions to cover a full day of forecasting.

Given the dataset's resolution of 5 minutes per data point, executing this cycle 288 times provides us with a forecast for the next 24 hours. Throughout this process, our framework leverages the logic of a chosen model architecture, such as LSTM or Bi-LSTM, which incorporates a set of parameters θ.

These parameters θ include:

- 12 primary features: $\{P_l$, year, week, day_of_week, hour, minute, second, hour_sin, hour_cos, day_of_week_sin, day_of_week_cos, week_sin, week_cos, is_holiday$\}$
- Additional weather-related variables: {temperature, pressure, dew_point, heat_index, humidity, solar_irradiance, uv_index, wind_chill}.

This approach ensures that each forecast point $\hat{y}_t$ in the cycle adheres to the underlying model's framework, providing reliable predictions for operational planning or analysis purposes.

4.3. Defining correlation parameters

To enhance the accuracy of our electrical load forecasting model, it is crucial to identify and incorporate relevant external data. As an initial step in our methodology, we focus on correlating the collected weather data with electrical load patterns. Based on an extensive literature review, we identified several methods for determining these correlations:

- **Heat Map:** A simple yet effective visualization tool that shows the strength of correlations between different variables using color gradients. This method helps in quickly identifying potential relationships between weather parameters and electrical load.
- **Pearson Correlation Coefficient:** A statistical measure that quantifies the linear relationship between two variables. It provides a numerical value between -1 and 1, indicating the direction and strength of the correlation. This method is useful for identifying linear dependencies between weather data and electrical load.
- **Random Forest:** A machine learning approach that can handle non-linear relationships and interactions between variables. By using feature importance metrics from the Random Forest algorithm, we can assess the impact of each weather parameter on the electrical load. This method is particularly powerful for capturing complex dependencies that may not be evident through simpler statistical methods.

By employing these diverse correlation methods, we aim to comprehensively understand the relationships between weather parameters and electrical load. This understanding is essential for selecting the most relevant features to improve the forecasting accuracy of our model.

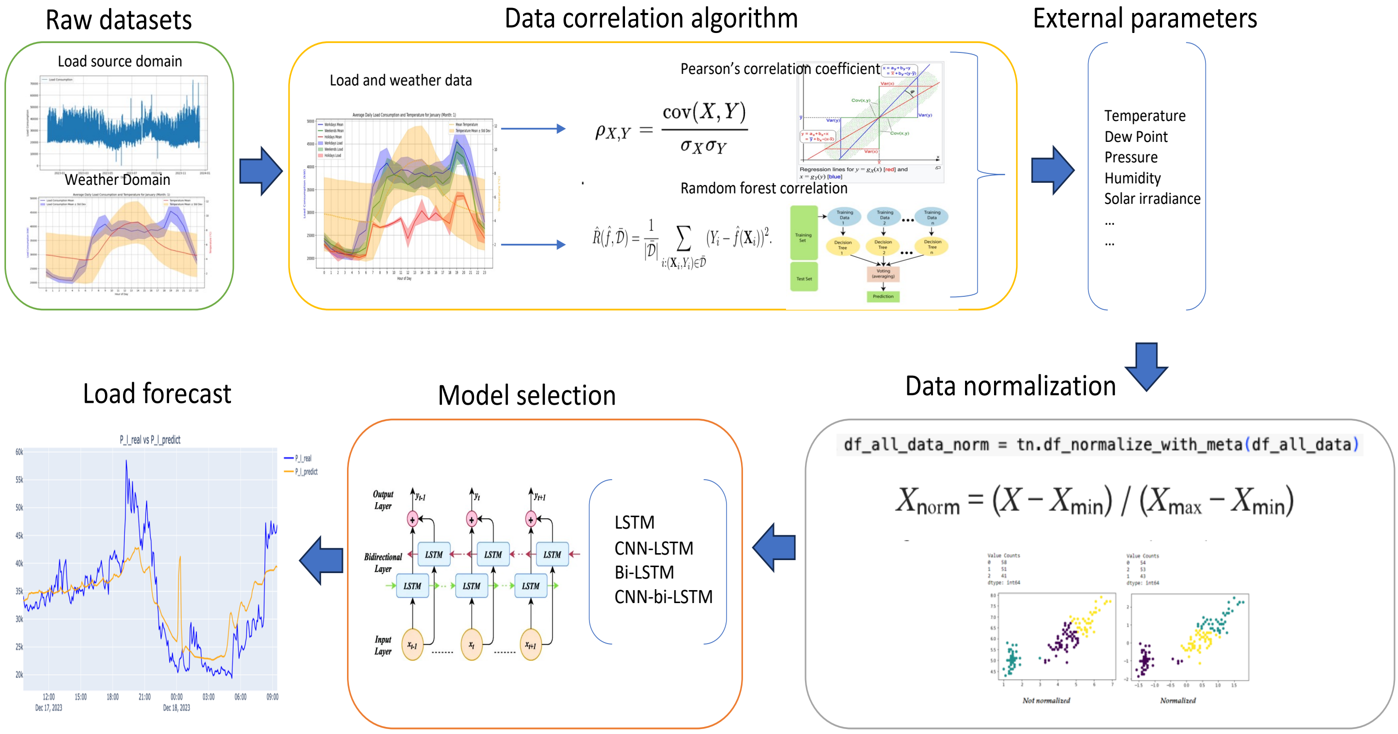

Figure 1: Forecasting framework

4.4. *Heat Map and Pearson Correlation Coefficient*

As a first step of research for determining the correlation between load consumption and weather data was chosen heat map method. Results of implementing the heat map of the correlation matrix and correlation of parameters with load consumption based on Pearson Correlation Coefficient are presented on 2 and 3. Temperature indicators (Dew Point, Heat Index, Wind Chill, Ambient Temperature) exhibit a negative correlation with load consumption. This denotes that with the escalation of temperature, the energy consumption values tend to diminish. This may be attributable to a reduced necessity for heating as the ambient air temperature rises.

The UV Index and Solar Irradiance also display a negative correlation, albeit more tenuous, which could indicate a decrement in the use of artificial lighting or air conditioning systems on bright sunny days.

Wind Speed Average and Wind Direction Average demonstrate negligible correlation with load consumption, suggesting their insubstantial impact on energy usage under the given conditions.

Rain Rate possesses a very low positive correlation, virtually not influencing energy consumption.

Relative Humidity manifests a slight positive correlation, which could imply a minor increase in energy consumption at high humidity levels, potentially due to the operation of dehumidifiers or air conditioning units.

Atmospheric Pressure shows the least positive correlation among all factors, hinting at a marginal increase in energy consumption with elevated atmospheric pressure.

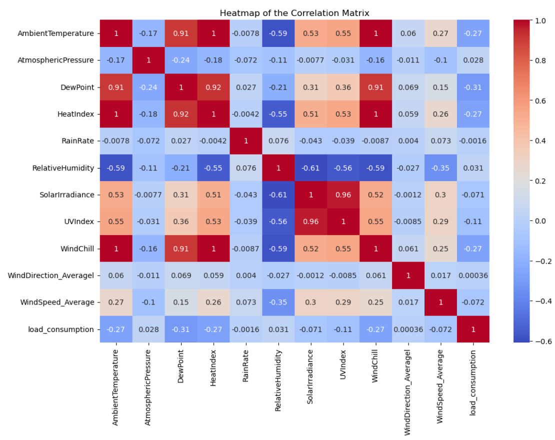

Figure 2: Heatmap of the Correlation Matrix

The Dew Point indicates the most substantial negative correlation, which may testify to a significant reduction in energy consumption with rising dew points. This can be associated with the fact that a higher dew point generally corresponds to warmer weather conditions, where there may be less demand for energy for heating purposes.

4.5. *Random Forest for Correlating Parameters*

To capture the complex dependencies between weather parameters and electrical load, we employed the Random Forest algorithm. This method is adept at handling non-linear relationships and interactions among variables, providing a comprehensive analysis of feature importance.

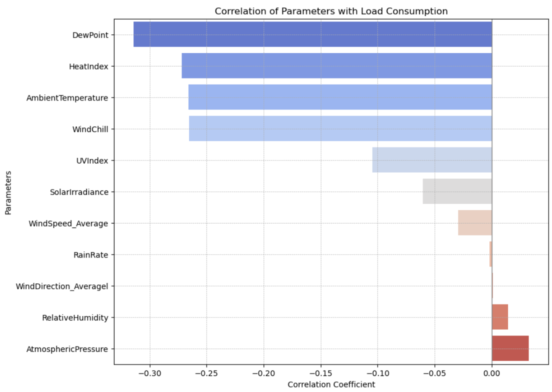

Figure 3: Correlation of Parameters with Load Consumption

Based on the feature importance metrics derived from the Random Forest model (as shown in Table 1), we initially considered a range of weather parameters. However, to refine our model and focus on the most impactful predictors, we decided to use only 8 parameters. This decision was driven by the observation that three parameters—Wind Speed Average, Wind Direction Average, and Rain Rate—had minimal impact on the correlation with load consumption. These parameters showed negligible importance in the Random Forest analysis, indicating that they do not significantly contribute to the model's predictive accuracy.

Feature	Importance
pressure	0.329318
dewpoint	0.199652
humidity	0.157317
heatindex	0.083769
solarirradiance	0.081718
windchill	0.072896
temperature	0.056828
uvindex	0.018502

Table 1: Feature Importance from Random Forest

The feature importance metrics derived from the Random Forest model indicate that atmospheric pressure is the most significant predictor of electrical load, followed by dew point and humidity. This highlights the strong influence of these weather parameters on energy consumption patterns.

Pressure, with an importance score of 0.329318, significantly impacts load consumption, possibly due to its influence on temperature and humidity, which in turn affect heating and cooling demands. Dew point (0.199652) and humidity (0.157317) also show strong correlations, underscoring their relevance in understanding energy usage patterns.

Heat index, solar irradiance, and wind chill have moderate importance scores, suggesting that while they do impact load consumption, their effects are not as pronounced as pressure, dew point, and humidity. Temperature and UV index exhibit the least importance, indicating a relatively minor impact on load consumption within the context of this study.

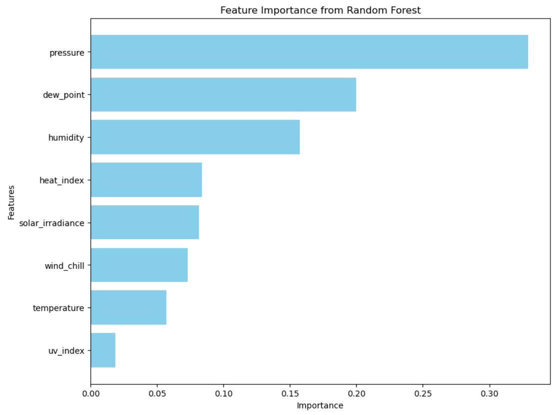

Figure 4: Feature Importance from Random Forest

To visually represent these findings, we created a bar chart of the feature importances from the Random Forest model, as shown in Figure 4.

This comprehensive analysis using Random Forest enables us to prioritize the most relevant weather parameters for improving the accuracy of our electrical load forecasting model. By incorporating these key features, we can enhance the model's predictive capability and better manage energy resources.

4.5.1. Analysis of Weather Parameters Influence on Forecast Accuracy

To further assess the impact of weather parameters on forecast accuracy, I conducted additional tests by sequentially adding each parameter to the input and evaluating the model's performance. This testing provided a clearer understanding of how individual weather features contribute to the prediction outcomes.

Among the tested parameters, **UV Index** had the most significant positive impact on forecast accuracy, substantially reducing error metrics 2. Specifically, its inclusion resulted in the lowest RMSE and WMAPE, along with a strong improvement in R^2, indicating that the UV Index effectively captures key environmental variations that influence energy consumption.

Conversely, **Dew Point**, **Heat Index**, and **Temperature** also contributed to improved model performance, showing moderate reductions in RMSE and an increase in R^2 values, although their impact was not as pronounced as UV Index.

Other parameters, such as **Humidity**, **Pressure**, and **Wind Chill**, demonstrated either a negligible or even negative effect on model accuracy, with high error metrics and negative R^2 values. These results suggest that these variables do not provide significant predictive value in the context of this load forecasting task, and their inclusion could lead to overfitting or unnecessary complexity.

Solar Irradiance, while having a weaker correlation with energy consumption, showed a slight positive effect on accuracy but was less impactful compared to other temperature-related parameters.

This additional testing confirmed that focusing on the

Weather Parameter	Influence Coefficient	Rank
UV Index	1.00	1
Heat Index	0.82	2
Temperature	0.78	3
Dew Point	0.71	4
Solar Irradiance	0.66	5
Pressure	0.45	6
Humidity	0.34	7
Wind Chill	0.10	8

Table 2: Ranking of Weather Parameters by Influence on Forecast Accuracy

most influential weather parameters, such as UV Index and temperature-related metrics, enhances the predictive performance of the model. The results of this sequential parameter analysis align with the feature importance rankings obtained from Random Forest, where UV Index and temperature parameters demonstrated substantial relevance.

In summary, the data suggests that energy consumption escalates in colder weather (potentially due to heating requirements) and diminishes in warmer conditions, but not during hot periods when the air conditioning system is active. Moderate climates witness an intermediate level of energy consumption. These observations should be considered in the context of local climate peculiarities and customary energy usage patterns.

4.6. Feature Engineering and Temporal Encoding for Forecasting

In our forecasting methodology, we leverage a sophisticated feature engineering process that transforms raw temporal data into a rich set of cyclic and contextual attributes. This process enhances the model's ability to recognize and predict patterns across different time scales, which is critical for accurate short-term load forecasting.

4.6.1. Extraction of Temporal Features

For each timestamp k, the following temporal attributes are extracted, forming the foundation of our forecasting model:

- **Year** ($y[k]$): The year corresponding to each timestamp.
- **Week** ($w[k]$): The ISO week number of the year.
- **Day of the Week** ($\mathrm{dow}[k]$): A numerical representation of the day of the week (0 for Monday, 6 for Sunday).
- **Hour** ($h[k]$): The hour of the day.
- **Minute** ($\mathrm{min}[k]$): The minute within the hour.
- **Second** ($\mathrm{sec}[k]$): The second within the minute.

4.6.2. Trigonometric Transformations of Temporal Features

To capture cyclic patterns inherent in time series data, we apply trigonometric transformations to the temporal features. These transformations are crucial for enabling the model to understand the periodic nature of time-related data.

- **Hour Sine and Cosine**: Capture daily cycles.

$$\mathrm{h\_sin}[k] = \sin\left(\frac{2\pi \cdot \mathrm{h}[k]}{24}\right) \tag{3}$$

$$\mathrm{h\_cos}[k] = \cos\left(\frac{2\pi \cdot \mathrm{h}[k]}{24}\right) \tag{4}$$

- **Day of the Week Sine and Cosine**: Capture weekly cycles.

$$\mathrm{dow\_sin}[k] = \sin\left(\frac{2\pi \cdot \mathrm{dow}[k]}{7}\right) \tag{5}$$

$$\mathrm{dow\_cos}[k] = \cos\left(\frac{2\pi \cdot \mathrm{dow}[k]}{7}\right) \tag{6}$$

- **Week Sine and Cosine**: Capture yearly cycles.

$$\mathrm{w\_sin}[k] = \sin\left(\frac{2\pi \cdot \mathrm{w}[k]}{52}\right) \tag{7}$$

$$\mathrm{w\_cos}[k] = \cos\left(\frac{2\pi \cdot \mathrm{w}[k]}{52}\right) \tag{8}$$

4.6.3. Holiday Information Extraction

The model also considers whether each timestamp k corresponds to a working day or a holiday, based on Italy's holiday calendar. This is captured by a binary flag:

$$\mathrm{is\_holiday}[k] = \begin{cases} 1 & \text{if } k \text{ is a holiday} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

4.6.4. Data Set Arrangement for Forecasting

After processing, the dataset for each time step k is augmented from a simple load measurement $P[k]$ to a season-informed vector $\mathbf{P}_{SI}[k]$. The final dataset includes the following features:

- Load Measurement: $P[k]$
- Year: $y[k]$
- Week Sine and Cosine: $\mathrm{w\_sin}[k]$, $\mathrm{w\_cos}[k]$
- Hour Sine and Cosine: $\mathrm{h\_sin}[k]$, $\mathrm{h\_cos}[k]$
- Day of the Week Sine and Cosine: $\mathrm{dow\_sin}[k]$, $\mathrm{dow\_cos}[k]$
- Holiday Flag: $\mathrm{is\_holiday}[k]$

This augmented dataset forms the input for our forecasting models, enhancing their ability to detect and predict cyclical and contextual patterns in the load data. This approach is particularly effective in capturing the nuances of temporal dynamics, leading to more accurate and reliable load forecasts.

4.7. Forecast Model Architecture

4.7.1. Bidirectional LSTM (BiLSTM) with Seasonal and Weather Components

The BiLSTM (Bidirectional Long Short-Term Memory) model processes time-series data in both forward and backward directions, allowing it to capture past and future context simultaneously. This architecture is highly effective for tasks like load forecasting, where temporal dependencies and external factors, such as weather, significantly influence the outcome.

Key Components.

- **LSTM Units:** Two stacked layers with 100 and 150 units, respectively.
- **Bidirectional Processing:** The model processes sequences forward and backward, combining hidden states to capture dependencies from both directions.
- **Input Features:** Includes temporal metadata (e.g., year, week, hour) and cyclical representations (e.g., `hour_sin`, `hour_cos`) as well as weather parameters.
- **Seasonal and Weather Components:** Critical for modeling periodic fluctuations and external factors (e.g., temperature, humidity) that influence energy consumption.

Parameter	**Value**
LSTM Units	100, 150 (two stacked layers)
Lag	12 (time steps)
Dropout	0.01
Dense Units	5
Activation	ReLU
Optimizer	Adam
Epochs	25

Table 3: Key BiLSTM Model Parameters

Model Parameters.

Feature Selection and Tuning. Features like pressure, dew point, and humidity were included based on their importance as determined by Random Forest analysis, but based on practice experiments UV index was chosen as a weather parameter, while less relevant features (e.g., wind speed) were excluded. The model parameters were tuned across 10 test experiments, and the best configuration was selected based on validation metrics, such as mean squared error (MSE) and mean absolute error (MAE).

Comparison with Other Architectures. The BiLSTM model was compared with LSTM, CNN-LSTM, and CNN-BiLSTM models. Each method was tuned using a similar hyperparameter grid, and the best-performing configurations were selected through validation testing.

4.8. Overall Metrics

The evaluation of the accuracy of various load forecasting models is based on the analysis of key metrics, providing insights into the efficacy of each forecasting technique. The selected metrics include RMSE, MAE, MAPE, and WMAPE.

4.8.1. Coefficient of Determination (R^2)

The coefficient of determination measures the proportion of the variance in the forecasted load values that is predictable from the actual load values. It ranges from 0 to 1, where 1 indicates a perfect fit. The formula for R^2 is given by:

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(P[k] - \hat{P}[k])^2}{\sum_{k=1}^{n}(P[k] - \bar{P})^2} \tag{10}$$

Where:

- $\bar{P}$ is the mean of the actual load values.

Root Mean Square Error (RMSE). Root Mean Square Error is a widely used metric to measure the average deviation of forecasted load values from actual load values. It is calculated as the square root of the average of squared differences between forecasted and actual values. The formula for RMSE is given in (11):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(P[k] - \hat{P}[k])^2} \tag{11}$$

Mean Absolute Error (MAE). Mean Absolute Error is a measure of the average absolute difference between actual and forecasted load values. It is calculated as the average of the absolute differences between the forecasted and actual values. The formula for MAE is given in (12):

$$\text{MAE} = \frac{1}{n}\sum_{k=1}^{n}|P[k] - \hat{P}[k]| \tag{12}$$

Mean Absolute Percentage Error (MAPE). Mean Absolute Percentage Error measures the average absolute percentage difference between actual and forecasted load values. It is often used to assess the accuracy of forecasts. The formula for MAPE is given in (13):

$$\text{MAPE} = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{P[k] - \hat{P}[k]}{P[k]}\right| \tag{13}$$

Weighted Mean Absolute Percentage Error (WMAPE). Weighted Mean Absolute Percentage Error is an extension of MAPE that accounts for the magnitude of the actual load values, providing a weighted average of the absolute percentage errors. This metric is particularly useful when dealing with varying load scales. The formula for WMAPE is given in (14):

$$\text{WMAPE} = \frac{\sum_{k=1}^{n}|P[k] - \hat{P}[k]|}{\sum_{k=1}^{n}P[k]} \tag{14}$$

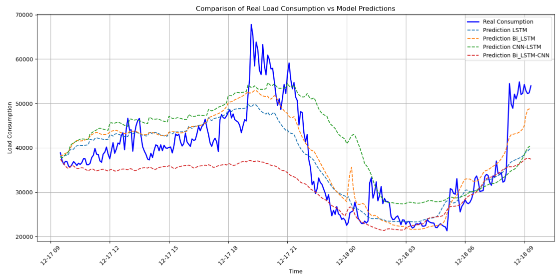

Figure 5: Comparison of Real Load Consumption vs Model Predictions

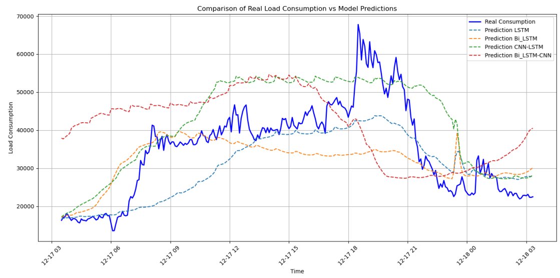

Figure 6: Comparison of Real Load Consumption vs Model Predictions with weather data

5. Experiments and Results

To evaluate the effectiveness of the proposed Bi-LSTM model, we conducted a series of experiments across various forecasting scenarios and time horizons. The primary objective was to assess the model's performance in comparison to other advanced methods, including CNN-LSTM, LSTM, and CNN-BiLSTM. Each experiment was designed to test the robustness and accuracy of the Bi-LSTM model under different conditions, ensuring comprehensive evaluation across multiple contexts.

The experiments were structured as follows:

5.1. Case A

Case A focuses on short-term load forecasting for a 24-hour time horizon using only the raw time series data. This scenario tests the model's ability to accurately predict immediate future load without incorporating any external data, which serves as a baseline for evaluating the effectiveness of additional features.

In this case, various models were tested, including BiLSTM, BiLSTM-CNN, CNN-LSTM, and LSTM. The results of these models, evaluated based on Root Mean Square Error (RMSE), R-squared, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Weighted Mean Absolute Percentage Error (WMAPE), are summarized in Table 4.

Method	RMSE	R-squared	MAE	MAPE	WMAPE
BiLSTM	3990.96	0.8038	3288.68	10.05	9.94
BiLSTM-CNN	7781.04	0.2543	5432.86	13.97	16.42
CNN-LSTM	7029.28	0.3915	5418.97	18.26	16.37
LSTM	4789.52	0.7175	3491.44	9.96	10.55

Table 4: Results of short-term load forecasting using raw time series data

On figure 5 the BiLSTM model demonstrated the best performance among all tested methods in Case A. It achieved the lowest RMSE of 3990.96, indicating the smallest prediction error. The R-squared value of 0.8038 suggests that the BiLSTM model explains 80.38% of the variance in the load data, which is significantly higher than the other models.

Additionally, the BiLSTM model recorded the lowest MAE (3288.68) and MAPE (10.05), showing that its predictions are both closer to the actual values and relatively more accurate in percentage terms. The WMAPE of 9.94 further highlights the model's robustness in handling variability in load data.

In comparison, the BiLSTM-CNN and CNN-LSTM models showed higher RMSE and lower R-squared values, indicating less accuracy in their predictions. The LSTM model performed better than the CNN-based models but was still outperformed by the BiLSTM model in all metrics.

These results establish the BiLSTM model as the baseline for short-term load forecasting using raw time series data. The performance of the BiLSTM model in this scenario serves as a benchmark for evaluating the improvements achieved by incorporating additional features in subsequent cases.

5.2. Case B

Case B extends the short-term forecasting by incorporating weather data along with the raw time series data. The model predicts load for the next 24 hours, evaluating the added value of weather parameters in improving prediction accuracy for real-time energy management and operational efficiency.

In this case, various models were tested, including BiLSTM, BiLSTM-CNN, CNN-LSTM, and LSTM. The results of these models, evaluated based on Root Mean Square Error (RMSE), R-squared, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Weighted Mean Absolute Percentage Error (WMAPE), are summarized in Table 5.

Method	RMSE	R-squared	MAE	MAPE	WMAPE
BiLSTM	8328.57	0.3376	6119.75	19.59	19.83
BiLSTM-CNN	4982.30	0.5920	3936.21	16.93	15.45
CNN-LSTM	7811.50	0.4173	5776.97	16.20	18.72
LSTM	7073.41	0.5222	5457.06	16.60	17.68

Table 5: Results of short-term load forecasting with weather data

The BiLSTM-CNN model demonstrated the best performance among all tested methods in Case B fig 6. It achieved the lowest RMSE of 4982.30, indicating the smallest prediction error. The R-squared value of 0.5920 suggests that the BiLSTM-CNN model explains 59.20% of the variance in the load data, which is higher than the other models in this scenario.

Additionally, the BiLSTM-CNN model recorded the lowest MAE (3936.21) and WMAPE (15.45), indicating that its predictions are closer to the actual values and handle variability in load data more effectively. However, the MAPE of 16.93 indicates that the percentage error in the predictions is relatively higher compared to the best performing model in Case A.

In comparison, the BiLSTM model showed the highest RMSE (8328.57) and the lowest R-squared (0.3376), indicating less accuracy in its predictions when incorporating weather

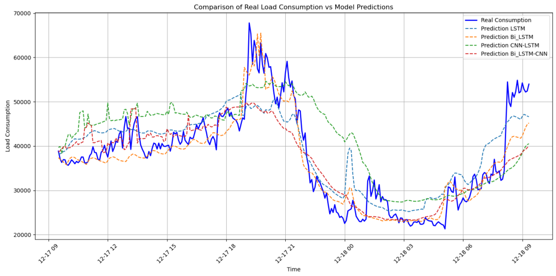

Figure 7: Comparison of Real Load Consumption vs Model Predictions with temporal metadata

data. The CNN-LSTM and LSTM models performed better than the BiLSTM model but were still outperformed by the BiLSTM-CNN model in most metrics.

These results illustrate the added value of incorporating weather data into the forecasting model. The BiLSTM-CNN model, in particular, benefits from the enriched information provided by the weather parameters, leading to improved accuracy in short-term load forecasting.

5.3. *Case C*

Case C further enhances the forecasting model by adding temporal metadata to the time series data, including day of the week and time features. This scenario tests the model's ability to leverage these temporal features for predicting load over the next 24 hours, which is crucial for understanding daily and weekly consumption patterns.

The incorporation of temporal metadata significantly improves the model's performance. Temporal features help the model to recognize patterns such as higher energy consumption on weekdays compared to weekends, or different load patterns during various times of the day. By including this information, the model can better anticipate changes in load, leading to more accurate forecasts.

In this case, various models were tested, including CNN-LSTM, LSTM, CNN-BiLSTM, and BiLSTM. The results of these models, evaluated based on Root Mean Square Error (RMSE), R-squared, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Weighted Mean Absolute Percentage Error (WMAPE), are summarized in Table 6.

Method	RMSE	R-squared	MAE	MAPE	WMAPE
CNN-LSTM	5864.28	0.5765	5007.01	17.36	15.13
LSTM	5184.20	0.6690	3945.43	12.57	11.92
CNN-BiLSTM	6055.11	0.5484	4224.82	12.77	12.77
BiLSTM	3699.31	0.8314	2640.43	7.68	7.97

Table 6: Results of forecasting

The BiLSTM model demonstrated the best performance among all tested methods fig7. It achieved the lowest RMSE of 3699.31, indicating the smallest prediction error. The R-squared value of 0.8314 suggests that the BiLSTM model explains 83.14% of the variance in the load data, which is significantly higher than the other models.

Furthermore, the BiLSTM model recorded the lowest MAE (2640.43) and MAPE (7.68), showing that its predictions are both closer to the actual values and relatively more accurate in percentage terms. The WMAPE of 7.97 reinforces the model's robustness, indicating that the BiLSTM model handles variability in load data more effectively.

In comparison, the CNN-LSTM and CNN-BiLSTM models showed higher RMSE and lower R-squared values, suggesting less accuracy in their predictions. The LSTM model performed better than the CNN-based models but was still outperformed by the BiLSTM model in all metrics.

These results illustrate that incorporating temporal metadata significantly enhances the forecasting model's ability to capture load patterns, and the BiLSTM model particularly benefits from this enriched information, making it the most accurate and reliable model for short-term load forecasting in this scenario.

5.4. *Case D*

Case D combines all previous elements by integrating weather data and temporal metadata with the raw time series data. This scenario aims to predict the load for the next 24 hours, demonstrating the comprehensive model's ability to utilize both weather parameters and temporal features for the most accurate short-term load forecasting.

In this case, various models were tested, including BiLSTM, BiLSTM-CNN, CNN-LSTM, and LSTM. The results of these models, evaluated based on Root Mean Square Error (RMSE), R-squared, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Weighted Mean Absolute Percentage Error (WMAPE), are summarized in Table 7.

Method	RMSE	R-squared	MAE	MAPE	WMAPE
BiLSTM	5035.85	0.5832	4193.82	18.85	16.46
BiLSTM-CNN	5341.83	0.5310	4329.28	21.21	16.99
CNN-LSTM	4401.07	0.7614	2936.88	8.54	8.87
LSTM	5323.46	0.5342	4314.55	20.81	16.93

Table 7: Results of short-term load forecasting with integrated weather and temporal data

The CNN-LSTM model demonstrated the best performance among all tested methods in Case D fig 8. It achieved the lowest RMSE of 4401.07, indicating the smallest prediction error. The R-squared value of 0.7614 suggests that the CNN-LSTM model explains 76.14% of the variance in the load data, which is significantly higher than the other models in this scenario.

Additionally, the CNN-LSTM model recorded the lowest MAE of 2936.88 and MAPE of 8.54, indicating that its predictions are both closer to the actual values and more accurate in percentage terms compared to other models. The WMAPE of 8.87 further reinforces the model's effectiveness in handling variability in load data.

In comparison, the BiLSTM model also performed well, achieving an RMSE of 5035.85 and an R-squared value of 0.5832, indicating a moderate level of accuracy. However, it was outperformed by the CNN-LSTM model in all metrics. The BiLSTM-CNN and LSTM models had similar performance, with RMSE values of 5341.83 and 5323.46, respectively, but

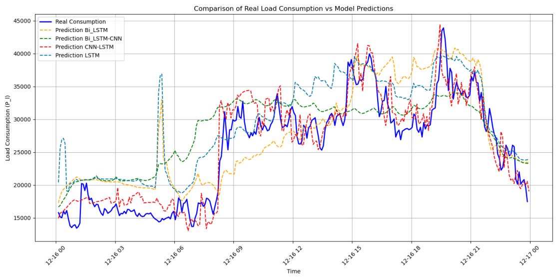

Figure 8: Comparison of Real Load Consumption vs Model Predictions with integrated weather and temporal data

both recorded higher MAE and MAPE compared to the CNN-LSTM model.

These results illustrate the advantage of integrating both weather data and temporal metadata into the forecasting model. The CNN-LSTM model particularly benefits from this comprehensive approach, leading to the most accurate short-term load forecasting in this scenario. This demonstrates the potential of leveraging diverse data sources to enhance forecasting performance.

5.5. Case E

5.5.1. The Sliding Window technique with 5 minutes resolution forecast

Based on the results from the previously implemented models, the Bi-LSTM method with seasonal metadata was selected for short-term load forecasting over a one-week time horizon. To ensure the accuracy of this approach, the Sliding Window technique was implemented for model training and prediction.

The Sliding Window approach involves iteratively training the model on a slightly updated dataset for each subsequent forecast. For the period from December 2 to December 8, 2023, one-day-ahead load forecasts were generated, as shown on Figure 9.

The training process for the December 2 forecast began with data from September 7, 2022. For each subsequent day, the training dataset was shifted by one day. For example, to predict the load for December 3, 2023, the model was trained starting from September 8, 2022. This process was repeated daily, creating a seven-day forecast using the Sliding Window approach.

Mathematically, the Sliding Window technique can be represented as follows:

$$\text{Forecast}(t+n) = f(\text{Train}(t-k, t)) \tag{15}$$

Here: - t represents the current day. - n is the forecast horizon (in this case, one day). - k is the size of the training window. - $\text{Train}(t-k, t)$ represents the training data used for forecasting, starting from $t-k$ and ending at t.

This approach allows the model to incorporate recent trends and patterns into each day's forecast, enhancing the overall accuracy over the one-week period. The results of the forecasts for each day from December 2 to December 8, 2023, are provided in Table 8, showcasing the effectiveness of the Sliding Window approach in short-term load forecasting.

Date	RMSE	R-squared (R^2)	MAE	MAPE	WMAPE
December 2, 2023	5001.74	0.7085	3396.52	10.09%	10.83%
December 3, 2023	4993.97	0.7153	2949.32	7.81%	9.18%
December 4, 2023	6866.79	0.6111	4626.91	10.74%	12.51%
December 5, 2023	5123.94	0.7295	3508.03	9.23%	9.50%
December 6, 2023	5249.56	0.7147	4036.07	11.10%	11.10%
December 7, 2023	5099.58	0.6764	3111.45	8.09%	9.09%
December 8, 2023	5386.42	0.2940	4569.67	16.52%	16.29%
Weekly Summary	**5388.84**	**0.6357**	**3742.85**	**10.66%**	**11.36%**

Table 8: 7-Day Forecast Results with 5 Minutes Data Resolution Using the Sliding Window Approach with Bi-LSTM and Temporal Metadata, Including Weekly Summary

5.5.2. The Sliding Window Technique with 15 Minutes Resolution Forecast

To compare forecasting accuracy, the Sliding Window technique was also applied to data resampled to 15-minute intervals. This interval more closely aligns with typical data collection practices and often better captures the load dynamics.

The results for the same period, December 2 to December 8, 2023, using the 15-minute resolution data, are presented in Table 9.

Date	RMSE	R-squared (R^2)	MAE	MAPE	WMAPE
December 2, 2023	3289.51	0.8480	2553.46	7.93%	8.14%
December 3, 2023	3159.54	0.8610	2279.64	6.71%	7.10%
December 4, 2023	5567.30	0.6958	4485.27	11.06%	12.12%
December 5, 2023	3500.79	0.8392	2806.24	8.42%	7.59%
December 6, 2023	3632.24	0.8369	2698.11	8.08%	7.41%
December 7, 2023	3275.27	0.8204	2627.69	7.61%	7.66%
December 8, 2023	3986.93	0.4221	3355.91	12.73%	11.95%
Weekly Summary	**4151.92**	**0.7087**	**3191.11**	**8.68%**	**8.94%**

Table 9: 7-Day Forecast Results with 15 Minutes Data Resolution Using the Sliding Window Approach with Bi-LSTM and Temporal Metadata, Including Weekly Summary

The results indicate that forecasts based on 15-minute resolution data tend to be more accurate compared to those based on 5-minute resolution data. Specifically, the 15-minute data provides a more stable and realistic representation of the load variations, aligning better with the conditions of sensor data collection. As demonstrated in Figure 10, the forecasts with 15-minute resolution exhibit higher accuracy, as reflected in the lower RMSE and MAPE values.

In summary, while both resolutions offer valuable insights, the 15-minute resolution data proved more effective in capturing the nuances of load patterns, leading to improved forecast accuracy.

5.5.3. Comparison of 7-Day Forecast Results and Actual Data with 15-Minute Data Resolution and Average Load Values

To evaluate the effectiveness of the proposed method, we compared it with the average load values μ calculated for each hour of each month. This analysis helps to understand the load consumption patterns over time. Let $L_{h,m}$ denote the load consumption for hour h in month m. The average load $\mu_{h,m}$ is computed using the following formula:

$$\mu_{h,m} = \frac{1}{N_{h,m}} \sum_{i=1}^{N_{h,m}} L_{h,m,i} \tag{16}$$

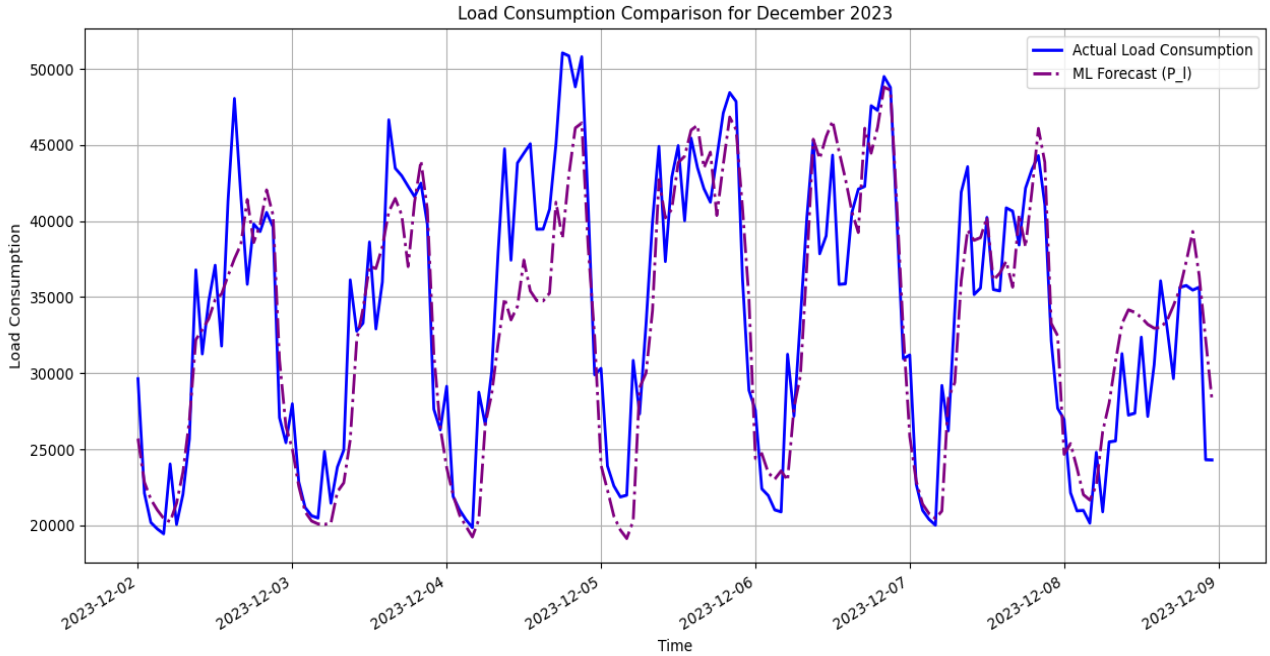

Figure 9: Comparison of 7-Day Forecast Results and Actual Data with 5 Minutes Data Resolution Using the Sliding Window Approach with Bi-LSTM and Temporal Metadata, Including Weekly Summary

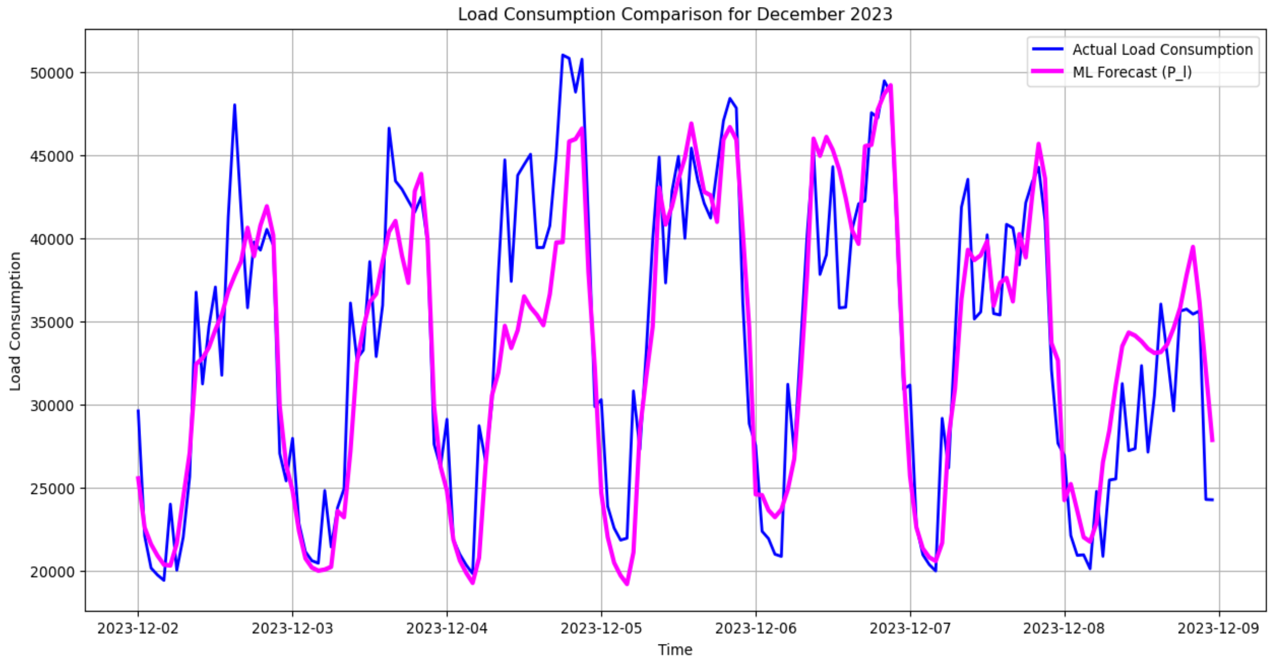

Figure 10: Comparison of 7-Day Forecast Results and Actual Data with 15 Minutes Data Resolution Using the Sliding Window Approach with Bi-LSTM and Temporal Metadata, Including Weekly Summary

where $N_{h,m}$ represents the number of load measurements for hour h in month m, and $L_{h,m,i}$ is the load consumption value at the i-th measurement. This average load $\mu_{h,m}$ provides insights into the consumption patterns based on seasonal and hourly variations throughout the year.

The results presented in Table 10 show the performance of the proposed method compared to the average load values. The metrics indicate that the suggested method consistently outperforms the average load approach. Specifically, the model demonstrates superior performance, often nearly doubling the accuracy of the average load values.

Figure 11 visually compares the 7-day forecast results with actual data at a 15-minute resolution against the average load values. The figure illustrates that the proposed model offers significantly improved forecast accuracy, particularly on atypical days.

Date	RMSE μ	MAPE μ	MAE μ	R-squared (R^2) μ	WMAPE μ
December 2, 2023	4334.37	10.85%	3235.73	0.7361	10.31%
December 3, 2023	4036.15	9.61%	3003.21	0.7732	9.35%
December 4, 2023	5681.29	10.36%	4337.64	0.6832	11.72%
December 5, 2023	4649.53	10.89%	4129.05	0.7163	11.17%
December 6, 2023	4736.19	9.70%	3784.69	0.7227	10.40%
December 7, 2023	2819.14	6.01%	2131.70	0.8669	6.22%
December 8, 2023	6115.03	17.60%	5072.56	-0.3595	18.07%
Period Average	**4337.18**	**10.36%**	**3366.68**	**0.5912**	**10.19%**

Table 10: Daily Metrics for Mean Load (μ) with 15-Minute Data Resolution, Including Period Average

5.6. Discussion

The experiments conducted across various forecasting scenarios and time horizons emphasize the robustness and adaptability of deep learning models, particularly the Bi-LSTM and CNN-LSTM architectures, in short-term load forecasting. Each case study provides insights into how different input data types—ranging from raw time series to enriched datasets with weather and temporal metadata—affect model performance.

Case A focused on forecasting load using raw time series data. The Bi-LSTM model established itself as a strong baseline, outperforming other models, including BiLSTM-CNN, CNN-LSTM, and LSTM, by achieving the lowest RMSE and the highest R-squared values. This result highlights the Bi-LSTM's capability to effectively capture inherent patterns in raw load data.

In **Case B**, the integration of weather data showed the advantages and limitations of different models when considering external factors. The BiLSTM-CNN model excelled in this scenario, leveraging convolutional layers to extract relevant features, achieving the lowest RMSE and highest R-squared values. However, the Bi-LSTM model's performance decline when additional non-temporal data was introduced suggests that its strength lies primarily in processing time series data.

Case C examined the impact of temporal metadata (e.g., day of the week, time of day) on forecasting accuracy. Here, the Bi-LSTM model again demonstrated superior performance, indicating that temporal metadata significantly enhances the model's ability to capture daily and weekly consumption patterns, thereby improving forecast accuracy.

In **Case D**, combining raw time series data, weather data, and temporal metadata led the CNN-LSTM model to outperform all others, achieving the lowest RMSE and highest R-squared values. This result underscores the CNN-LSTM's ability to integrate diverse data types effectively, capturing both temporal dependencies and spatial features. While the Bi-LSTM model performed well, it was surpassed by the CNN-LSTM in this more complex scenario, suggesting that model architecture should be tailored to the specific characteristics of the dataset.

These findings suggest that while the Bi-LSTM model is highly effective for time series data with temporal metadata, the CNN-LSTM model offers superior performance when integrating contextual information such as weather data and temporal metadata. This indicates that hybrid models like CNN-LSTM provide a better balance of accuracy and robustness in real-world applications involving multiple data sources.

Overview of the Sliding Window Technique: To refine forecast accuracy over a one-week horizon, the Sliding Window technique was applied with the Bi-LSTM model, enhanced by seasonal metadata. This approach involves iterative model training with a slightly updated dataset for each forecast, allowing the model to adapt to recent data trends. Forecasts were generated daily from December 2 to December 8, 2023, each using training data starting from September 7, 2022, onwards. This technique improved forecast accuracy by maintaining the model's relevance with the latest data.

Impact of Data Resolution on Forecast Accuracy: The study compared forecasts using 5-minute and 15-minute resolution data, finding that the latter consistently yielded more accurate results. The 15-minute data provided a more stable representation of load variations, leading to lower RMSE and MAPE values, as shown in Figure 10. This suggests that 15-minute resolution data effectively captures load patterns by reducing noise and smoothing trends over slightly longer intervals.

Comparison with Average Load Values: The Bi-LSTM model's performance, particularly when combined with the Sliding Window technique and 15-minute data resolution, was also compared with average load values ($\mu_{h,m}$) calculated for each hour of the month. The results, presented in Table 10, demonstrate that the Bi-LSTM model consistently outperformed the average load approach, with lower error metrics (RMSE, MAPE, WMAPE), indicating its superior ability to capture dynamic and seasonal load fluctuations.

Overall, the results establish benchmarks for short-term load forecasting and emphasize the importance of selecting the appropriate model architecture based on the input data's characteristics. Future research could explore combining these models or developing new architectures to further enhance prediction accuracy by leveraging diverse datasets.

6. Conclusion

This study introduced a Bi-LSTM model enhanced with temporal metadata for short-term electrical load forecasting, demonstrating its superior accuracy across various forecasting scenarios. The model's ability to capture both past and future dependencies, coupled with the integration of temporal metadata such as time of day and day of the week, proved critical in improving forecast precision, particularly in scenarios where immediate load fluctuations are crucial for real-time energy management.

Experimental results consistently showed that the Bi-LSTM model outperformed state-of-the-art methods, including LSTM and CNN-LSTM, especially when focusing on time series data enriched with temporal metadata. While weather data provided some additional context, its influence on forecast accuracy was relatively minor compared to the strong impact of temporal features. This highlights the effectiveness of the Bi-LSTM model in scenarios where load patterns are predominantly driven by time-dependent factors.

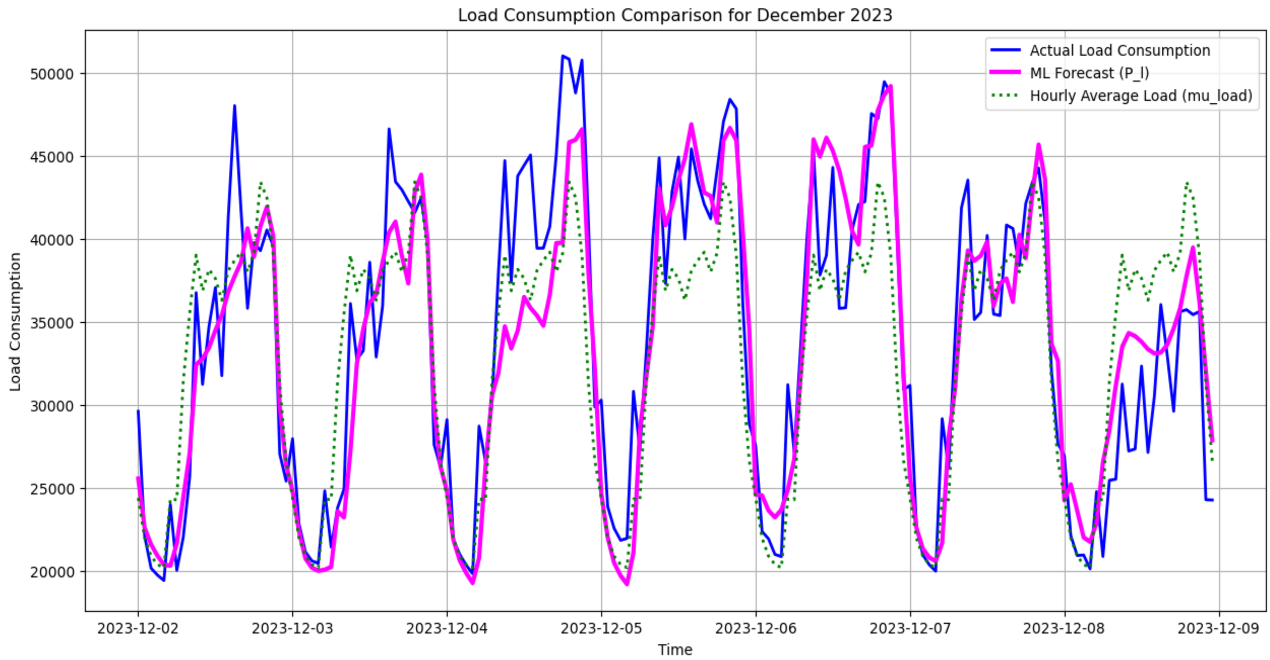

Figure 11: Comparison of 7-Day Forecast Results and Actual Data with 15-Minute Data Resolution and Average Load Value Using the Sliding Window Approach with Bi-LSTM and Temporal Metadata, Including Weekly Summary

The success of the Sliding Window technique, applied in a one-week forecasting horizon, further underscores the model's adaptability and relevance in dynamic environments, ensuring that the most recent data trends are incorporated into the predictions.

In summary, the Bi-LSTM model with temporal metadata offers a robust, scalable, and efficient solution for short-term load forecasting, outperforming existing methods and proving its value in energy management and planning. Future work could explore the integration of the developed Bi-LSTM model into different datasets, particularly those from diverse geographical regions or sectors with varying load consumption patterns.

References

Abdulrahman, M.L., Gital, A.Y., Ibrahim, K.M., Zambuk, F.U., Umar, I.M., Yakubu, Z.I., 2022. Predicting electricity consumption in residential building's using deep transfer learning with attention-lstm, in: 2022 International Mobile and Embedded Technology Conference (MECON), IEEE. pp. 540–546.

Abumohsen, M., Owda, A.Y., Owda, M., 2023. Electrical load forecasting using lstm, gru, and rnn algorithms. Energies 16, 2283.

Adedeji, P.A., Akinlabi, S., Madushele, N., Olatunji, O., 2019. Neuro-fuzzy mid-term forecasting of electricity consumption using meteorological data, in: IOP Conference Series: Earth and Environmental Science, IOP Publishing. p. 012017.

Ahmad, M., 2017. Seasonal decomposition of electricity consumption data. Review of Integrative Business and Economics Research 6, 271.

Alberg, D., Last, M., 2018. Short-term load forecasting in smart meters with sliding window-based arima algorithms. Vietnam Journal of Computer Science 5, 241–249.

Ali, Ö.G., Yaman, K., 2013. Selecting rows and columns for training support vector regression models with large retail datasets. European Journal of Operational Research 226, 471–480.

Almihat, M.G.M., Kahn, M., Aboalez, K., Almaktoof, A.M., 2022. Energy and sustainable development in smart cities: An overview. Smart Cities 5, 1389–1408.

Alrashidi, A., Qamar, A.M., 2023. Data-driven load forecasting using machine learning and meteorological data. Computer Systems Science & Engineering 44.

Apadula, F., Bassini, A., Elli, A., Scapin, S., 2012. Relationships between meteorological variables and monthly electricity demand. Applied Energy 98, 346–356.

Avtar, R., Tripathi, S., Aggarwal, A.K., Kumar, P., 2019. Population–urbanization–energy nexus: a review. Resources 8, 136.

Bahrami, S., Hooshmand, R.A., Parastegari, M., 2014. Short term electric load forecasting by wavelet transform and grey model improved by pso (particle swarm optimization) algorithm. Energy 72, 434–442.

Borges, C.E., Kamara-Esteban, O., Castillo-Calzadilla, T., Andonegui, C.M., Alonso-Vicario, A., 2020. Enhancing the missing data imputation of primary substation load demand records. Sustainable Energy, Grids and Networks 23, 100369.

Das, C.K., Bass, O., Kothapalli, G., Mahmoud, T.S., Habibi, D., 2018. Overview of energy storage systems in distribution networks: Placement, sizing, operation, and power quality. Renewable and Sustainable Energy Reviews 91, 1205–1230.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the em algorithm. Journal of the royal statistical society: series B (methodological) 39, 1–22.

Dmitri, K., Maria, A., Anna, A., 2016. Comparison of regression and neural network approaches to forecast daily power consumption, in: 2016 11th International Forum on Strategic Technology (IFOST), IEEE. pp. 247–250.

Dong, Y., Ma, X., Fu, T., 2021. Electrical load forecasting: A deep learning approach based on k-nearest neighbors. Applied Soft Computing 99, 106900.

Eseye, A.T., Lehtonen, M., Tukia, T., Uimonen, S., Millar, R.J., 2019. Short-term forecasting of electricity consumption in buildings for efficient and optimal distributed energy management, in: 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), IEEE. pp. 1103–1110.

Gao, J., Chen, Y., Hu, W., Zhang, D., 2023. An adaptive deep-learning load forecasting framework by integrating transformer and domain knowledge. Advances in Applied Energy 10, 100142.

Hammad Alharbi, H., Kimura, M., 2020. Missing data imputation using data generated by gan, in: 2020 the 3rd International Conference on Computing

and Big Data, pp. 73–77.
Huang, G., 2021. Missing data filling method based on linear interpolation and lightgbm, in: Journal of Physics: Conference Series, IOP Publishing. p. 012187.
Hyman, J.M., 1983. Accurate monotonicity preserving cubic interpolation. SIAM Journal on Scientific and Statistical Computing 4, 645–654.
Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for r. Journal of statistical software 27, 1–22.
Jeong, D., Park, C., Ko, Y.M., 2021. Short-term electric load forecasting for buildings using logistic mixture vector autoregressive model with curve registration. Applied Energy 282, 116249.
Kabeyi, M.J.B., Olanrewaju, O.A., 2022. Sustainable energy transition for renewable and low carbon grid electricity generation and supply. Frontiers in Energy Research 9, 1032.
Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L., 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology 49, 764–766.
Liu, H., Shah, S., Jiang, W., 2004. On-line outlier detection and data cleaning. Computers & chemical engineering 28, 1635–1647.
Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: Concerns and ways forward. PloS one 13, e0194889.
Massana, J., Pous, C., Burgas, L., Melendez, J., Colomer, J., 2015. Short-term load forecasting in a non-residential building contrasting models and attributes. Energy and Buildings 92, 322–330.
Mbuli, N., Mathonsi, M., Seitshiro, M., Pretorius, J.H.C., 2020. Decomposition forecasting methods: A review of applications in power systems. Energy Reports 6, 298–306.
Van der Meer, D.W., Shepero, M., Svensson, A., Widén, J., Munkhammar, J., 2018. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using gaussian processes. Applied energy 213, 195–207.
Nadtoka, I., Balasim, M.A.Z., 2015. Mathematical modelling and short-term forecasting of electricity consumption of the power system, with due account of air temperature and natural illumination, based on support vector machine and particle swarm. Procedia Engineering 129, 657–663.
Naz, A., Javed, M.U., Javaid, N., Saba, T., Alhussein, M., Aurangzeb, K., 2019. Short-term electric load and price forecasting using enhanced extreme learning machine optimization in smart grids. Energies 12, 866.
Prema, V., Rao, K.U., 2015. Time series decomposition model for accurate wind speed forecast. Renewables: Wind, Water, and Solar 2, 1–11.
Raza, M.Q., Khosravi, A., 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renewable and Sustainable Energy Reviews 50, 1352–1372.
Thành, P.V., 2019. Using improved grey forecasting model to estimate the electricity consumption demand in vietnam. Hue University Journal of Science: Economics and Development 128, 13–21.
Theile, P., Towle, A.L., Karnataki, K., Crosara, A., Paridari, K., Turk, G., Nordstrom, L., 2018. Day-ahead electricity consumption prediction of a population of households: analyzing different machine learning techniques based on real data from rte in france, in: 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), IEEE. pp. 1–6.
Wang, X., Meng, M., 2012. A hybrid neural network and arima model for energy consumption forcasting. J. Comput. 7, 1184–1190.
Xie, J., Zhong, Y., Xiao, T., Wang, Z., Zhang, J., Wang, T., Schuller, B.W., 2022. A multi-information fusion model for short term load forecasting of an architectural complex considering spatio-temporal characteristics. Energy and Buildings 277, 112566.
Zhang, S., Chen, R., Cao, J., Tan, J., 2023. A cnn and lstm-based multi-task learning architecture for short and medium-term electricity load forecasting. Electric Power Systems Research 222, 109507.
Zhu, G., Peng, S., Lao, Y., Su, Q., Sun, Q., 2021. Short-term electricity consumption forecasting based on the emd-fbprophet-lstm method. Mathematical Problems in Engineering 2021, 1–9.
Zolfaghari, M., Sahabi, B., 2019. A hybrid approach to model and forecast the electricity consumption by neurowavelet and arimax-garch models. Energy Efficiency 12, 2099–2122.