CHAPTER 1

INTRODUCTION

1.1 Rationale

Electrical load is a fundamental component in managing electric power systems and is crucial for ensuring a reliable and adequate electricity supply [1]. The growing electricity demand, driven by economic growth and population expansion, requires precise forecasting to optimize power system operations and prevent issues such as power shortages or overcapacity, which can have significant economic and environmental consequences [2, 3]. Industrial activity, household consumption, and weather conditions significantly influence electrical loads [4]. Given the instantaneous nature of electricity, maintaining a real-time balance between power generation and consumption is essential for system stability [5]. Short-term power load forecasting is therefore critical to achieving this balance and supporting the structured development of society [4, 6]. Common short-term power load forecasting approaches encompass traditional techniques, machine learning/deep learning methods, and integrated model forecasting strategies [5].

Numerous traditional methods for electrical load forecasting exist, including time series analysis, regression analysis, and grey forecasting methods. These statistical techniques are known for their straightforward structures, easy understanding, and quick execution. However, they have limitations, such as limited ability to utilize information, poor performance predicting nonlinear and complex data, and restricted generalization capabilities [2, 7]. The time series method, for example, forecasts future data based on past and current data but often overlooks the impact of complex factors [3]. These traditional methods generally rely on simple formulas, which can lead to poor fitting and significant errors [8]. Therefore, using machine learning technology to address existing problems by handling nonlinear and complex data offers faster learning speed, strong convergence, and the ability to extract deep data features [4].

Previous studies have used machine learning methods for electricity load forecasting with various approaches. Huang et al. use regression and neural network techniques to predict electricity consumption patterns [9]. These methods use historical data to identify future electricity load trends and anomalies. However, these methods are often limited in capturing sudden changes or new trends not reflected in historical data [9]. Therefore, a more adaptive and sophisticated approach is needed to improve prediction accuracy, especially when dealing with increasingly complex electricity consumption dynamics, such as the LSTM model. The LSTM model can overcome problems such as vanishing and exploding gradients in previous models [10], so this method is very effective for predicting time series, as shown by Wang et al. in estimating nonlinear load data [11]. The LSTM model, while effective, has notable limitations, including the vanishing gradient problem, which persists in deep networks or long training sessions [12]. The numerous hyperparameters make tuning difficult [13]. To overcome these challenges, the BiLSTM model improves prediction accuracy by utilizing bidirectional information flow, proving effective in power load forecasting [14, 15].

Regarding integrating BiLSTM with other methods, Ying et al.'s research demonstrated that the use of BiLSTM combined with phase space reconstruction can enhance the accuracy of short-term wind power forecasting compared to models like ANN, KNN, and RF [16]. Despite its effectiveness, this approach has drawbacks, particularly in detecting subtle frequency components and unexpected short-term variations. Consequently, applying signal decomposition can more effectively separate different frequency components, enabling the model to manage data variability and boost prediction accuracy.

Numerous studies have extensively explored the enhancement of prediction accuracy in time series analysis through Signal Decomposition. For instance, Wang et al. applied the Empirical Mode Decomposition (EMD) technique in forecasting electrical load, which effectively divides the original signal into multiple intrinsic components. These components can be examined individually to uncover more intricate patterns [17]. Gao et al. illustrate that Variational Mode Decomposition (VMD) effectively addresses the issue of mode mixing, which is a common challenge in Empirical Mode Decomposition (EMD), thus enhancing the quality of decomposition outcomes [18]. Furthermore, research presented in [17] highlights that integrating signal decomposition techniques with machine learning models can greatly enhance the accuracy of short-term electrical load forecasting. Nevertheless, signal decomposition methods still need help managing noise and non-stationary signals, which could potentially affect the resulting decomposition's accuracy. To address these issues, methods that are more robust to noise and capable of providing a more stable decomposition by combining the results of several different EMD iterations have been explored [19]. This approach allows the model to handle more complex signal variations, improving overall prediction performance. Therefore, this study uses Ensemble Empirical Mode Decomposition (EEMD).

In this study, we propose a machine learning approach for electricity load estimation using Bidirectional Long Short-Term Memory (BiLSTM). A spatial correlation method is introduced to identify the most relevant geographical regions, enhancing the model's ability to capture localised consumer behaviour and weather patterns. Additionally, we implement an advanced feature extraction technique employing Ensemble Empirical Mode Decomposition (EEMD), where their correlation with the target variable informs the selection of Intrinsic Mode Functions (IMFs). This approach is applied in a case study of the Jakarta-Banten region in Indonesia, demonstrating its effectiveness in improving prediction accuracy and underscoring the innovative aspects of our methodology.

1.2 Statement of the Problem

The complexity of electricity consumption patterns, driven by economic growth, urbanisation, and varying weather conditions, presents significant challenges for accurate load forecasting. Traditional forecasting methods, including time series analysis, regression, and grey forecasting, often fall short when handling non-linear and complex data—critical components for effective load prediction in modern electric power systems [2, 7]. These methods typically rely on historical data, which may not adequately capture sudden changes or emerging trends, leading to unreliable forecasts [3, 8]. Consequently, there is a pressing need for more advanced models that can better address the dynamic nature of electricity demand. While machine learning and deep learning models, particularly Long Short-Term Memory (LSTM) networks, have shown promise in modelling non-linear data patterns and extracting deep features from extensive datasets, they still face significant issues such as the vanishing gradient problem, difficulty in hyperparameter tuning, and sensitivity to long-term dependencies [4, 10, 12].

To address these limitations, the Bidirectional Long Short-Term Memory (BiLSTM) model has been proposed, which utilises bidirectional information flow to improve prediction accuracy in power load forecasting [14, 15]. However, BiLSTM alone may still be insufficient in detecting subtle frequency components and sudden short-term variations in complex electricity consumption data [16]. Therefore, this study proposes integrating BiLSTM with advanced signal decomposition techniques, such as Ensemble Empirical Mode Decomposition (EEMD), to enhance the model's ability to handle data variability and extract relevant features for more accurate short-term load forecasting [17, 18]. Additionally, incorporating spatial correlation methods to identify the most pertinent geographical regions and their specific consumption patterns can further refine model predictions. This integrated approach is designed to improve the robustness of forecasting models by better adapting to the increasingly complex dynamics of electricity consumption, as demonstrated in the case study of the Jakarta-Banten region in Indonesia.

1.3 Objective and Hypotheses

This section outlines the study's primary objectives and hypotheses. The study aims to address the identified problems through specific research goals and testable hypotheses.

1.3.1 Objective

The primary objective of this research is to enhance the accuracy of short-term electricity load forecasting by integrating Bidirectional Long Short-Term Memory (BiLSTM) models with advanced signal decomposition techniques, specifically Ensemble Empirical Mode Decomposition (EEMD) and by incorporating spatial correlation methods. This study aims to develop a robust forecasting model to effectively capture complex consumption patterns influenced by temporal dynamics and geographical variations. Specifically, the objectives of this research are:

1. To evaluate the accuracy of combining BiLSTM models with EEMD in decomposing load signals and extracting relevant features to improve forecasting accuracy.
2. To assess the impact of incorporating spatial correlation methods in refining the model's predictions by identifying relevant geographical regions and their unique electricity consumption patterns.

1.3.2 Hypotheses

The hypotheses of this research are based on the premise that integrating advanced machine learning models with sophisticated signal decomposition and spatial analysis techniques will significantly enhance the accuracy and reliability of electricity load forecasting. Specifically, the following hypotheses are proposed:

1. Integrating BiLSTM models with EEMD for signal decomposition and feature extraction will significantly improve the accuracy of short-term electricity load forecasting by providing a more detailed signal representation and enhancing the model's ability to manage data variability.
2. Applying spatial correlation methods to identify relevant geographical regions will refine the model's predictions by effectively capturing localised electricity consumption patterns, thereby improving forecasting performance.

1.4 Assumption

This research assumes that temporal factors, such as daily and seasonal variations, and spatial factors, including geographical location and localised weather, influence electricity consumption patterns. These factors are considered predictable and measurable using appropriate data collection and analysis techniques. Additionally, it is assumed that integrating advanced machine learning models with signal decomposition methods will enhance the modelling of electricity load dynamics by effectively capturing complex non-linear relationships and breaking them down into interpretable components. It is further assumed that spatial correlation methods can effectively model geographical variations in electricity consumption by considering unique local factors. Ultimately, combining Bidirectional Long Short-Term Memory (BiLSTM) models, signal decomposition techniques, and spatial correlation methods is expected to significantly improve the accuracy and reliability of short-term electricity load forecasting.

1.5 Scope and Delimitation

This study focuses on developing a short-term electricity load forecasting model by integrating Bidirectional Long Short-Term Memory (BiLSTM) networks with signal decomposition techniques, specifically Ensemble Empirical Mode Decomposition (EEMD) and spatial correlation methods. The critical variables analysed in this research include electricity load data, weather variables, and geographical information relevant to the Jakarta-Banten region in Indonesia. The study is centred on this locale to account for localised consumption patterns influenced by regional weather and geographical characteristics. The research timeframe covers one year of historical data to capture seasonal variations and short-term fluctuations in electricity consumption. This scope is justified as it allows for a comprehensive analysis of electricity demand dynamics and validates the proposed model's effectiveness in handling the complexity and variability inherent in short-term electricity load forecasting.

1.6 Significance of the Study

This study contributes significantly to electricity load forecasting by proposing a novel approach integrating Bidirectional Long Short-Term Memory (BiLSTM) networks with signal decomposition techniques, such as Ensemble Empirical Mode Decomposition (EEMD) and spatial correlation methods. The findings provide more precise insights into the benefits of combining advanced machine learning models with signal processing and spatial analysis to improve forecasting accuracy. The developed model offers a robust framework that can be utilised by utility companies, policymakers, and researchers to enhance the reliability of electricity supply management, optimise power system operations, and devise targeted strategies for energy consumption planning. Furthermore, the outcomes of this study can serve as a guide for future research to apply similar integrated approaches in other regions and contexts, thereby expanding the understanding of dynamic electricity consumption patterns under diverse environmental and geographical conditions.