CHAPTER 1

INTRODUCTION

1.1 Rationale

Renewable energy (RE) is one of the essential solutions to the Sustainable Development Goals (SDGs) [1, 2]. Renewable energy has been actively used worldwide in recent years and is an unlimited ecological resource without emitting a greenhouse effect [1]. RE can be produced from sunlight, water, geothermal, and wind and used as a substitute for fossil fuels to generate electricity [1]. Many countries are now abandoning fossil fuels because they can cause serious problems related to environmental pollution and global warming [1–3], which can result in the melting of the North and South Pole ice and mass drought in agricultural areas [4]. With these problems, it is currently necessary to produce energy sources that are environmentally friendly and low in carbon [3].

Solar energy is an alternative solution because it is environmentally friendly, clean, and included in sustainable energy with unlimited potential and economic benefits [2–4]. Solar energy comes from sunlight, a renewable resource [5]. Solar Photovoltaic (PV) is used to convert sunlight into electrical energy [4], which can be used to meet daily needs, such as communicating to obtain information, transportation, education, and other aspects of life [6]. PV technology has now been developed and researched by several countries, such as Portugal, building solar power plants with smart grid, microgrid or nano grid systems [5]; besides that, Tanzania is conducting research related to comparing PV Technologies between a-Si and CISG to review the power output effectiveness of PV power [7]. Australia researched the effectiveness of PV technologies with the practical order of PV modules m-Si, p-Si, a-Si, h-Si, CdTe, and CISG and p-Si technology produces the most stable power output [2]. Technological breakthroughs are increasingly increasing efficiency and reducing costs in using active power in solar panels [2].

Prediction of active power in solar panels can benefit PV system planning and operation. However, some challenges must be faced, such as sunlight is only sometimes available throughout the day, so active power optimization is needed so that there is no overestimation of active power costs, which can result in material losses[3]. In addition, the PV power generated depends on various weather parameters. As [4] stated, temperature is an important factor that impacts PV panels and solar systems, which are also affected by dust, ambient temperature, humidity, and wind [4].

Several predictions of PV power output have been carried out to improve prediction accuracy. Using a dihybrid Symbolic Regression (SR) model with Multi-Layer Perceptron (MLP) has shown good results in forecasting PV power output. In testing, this model achieved an RSME score of 5.58, an MAE of 3.3, and an R2 of 0.993. This shows that using the SR method can reduce the number of iterations and produce higher forecasting accuracy. In addition, the feature selection used uses the Extreme Boosting and Elastic Method with features including Wind Speed, Temperature, Relative humidity, Previous PV power, Horizontal Radiation, and Diffuse Horizontal Radiation, based on the method used to select horizontal radiation and previous PV features power is the most important feature compared to other features [1]. In [2], the MLP-Controlled ARIMA (CARIMA)-Gumbel Probabilistic-based Model (GPM) algorithm has also been tested in predicting PV power output with various PV module technologies. The results show good performance, with R 2 ranging from 0.927 to 0.9999, RSME from 0.082 to 1.913, MAPE from 0.0511 to 7.321, and BIC from -0.811 to 1.404. This algorithm is suitable for forecasting PV power output with various PV module technologies. The features used include Ambient Temperature, Relative humidity, Global horizontal radiation, Solar radiation, Wind speed, and Clearness index. Based on the research that has been done, the Global solar radiation and clarity index features are the most influential features. From the results of two studies that have been conducted in Australia with Solar farm data [1] and Australia Solar Center desert knowledge data, it can be concluded that several algorithms have been tested for prediction of PV power output providing important information about model performance and methods that can be used in PV power output forecasting.

Short-term active power forecasting is needed to minimize decomposition residuals as much as possible. The Variational Mode Decomposition (VMD) algorithm effectively overcomes denoising effects and modal aliasing and increases fault diagnosis accuracy by 91.67%, with a running time of 18 seconds. The VMD- Mutual Dimensionaless Indicator(MDI) data preprocessing method is more effective and accurate than the empirical mode decomposition (EMD) [8]. EMD is an empirical recursive technique that can be affected by noise and lacks a strong theoretical foundation, leading to potential mode aliasing and inconsistent results and VMD, in contrast, employs a non-recursive optimization-based approach with a solid mathematical basic, providing more accurate, stable, and reliable signal decomposition without mode aliasing issues [8]. Specifically [9], for increased Signal Noise Ratio (SNR) values and the best effect of using VMD or Local Mean Decomposition (LMD)-MDI. The absolute value of the SNR increases by 146.5% with a comparison of the filtering signal from real data. The RSME and MAR of the noise cancellation model are relatively reduced compared to the noise cancellation model. The VMD group had the best effect, and the metrics of all models were the lowest among all groups, the mean values of RSME and MAE decreased by 22.5% and 31.3% respectively compared to the original. For individual models, VMD-SVR has the greatest RSME and MAE reduction rate among all experimental models, reaching 31.0% and 38.5%, respectively. Thus, VMD can provide higher accuracy than others.

Variational Mode Decomposition (VMD) in the feature selection process can greatly enhance the accuracy of models predicting solar panel power output [10]. VMD is a signal processing technique that decomposes a complex signal, like the time series of power output, into intrinsic mode functions (IMFs), each representing different frequency components of the original signal [11]. This decomposition allows for analysing power data across various temporal scales, capturing high-frequency fluctuations and low-frequency trends [12]. The VMD algorithm effectively reduces noise, avoids modal aliasing, and improves fault diagnosis accuracy by 91.67% with a processing time of 18 seconds [8]. In addition to selecting appropriate features, choosing the right model algorithm is crucial.

Ensemble learning techniques, which combine the predictions of multiple models, can significantly improve accuracy and robustness. The key idea behind ensemble learning is that combining different model outputs helps reduce overfitting, enhances generalizability, and improves predictive performance [13]. Prominent ensemble methods include Bagging, Boosting, and Stacking [14]. Bagging trains several instances of the same model on different data subsets and averages their predictions to lower variance, with Random Forest being a well-known example. Boosting builds models sequentially, with each new model correcting the errors of its predecessor, thereby minimizing bias and variance. Stacking involves training a meta-model to integrate the predictions of base models [14]. In this approach, various models like Support Vector Machines, Decision Trees, and Deep Neural Networks are independently trained, and their predictions are used as inputs for a higher-level meta-model, which learns the best combination of these base predictions to enhance overall performance [15].

In this study, we will predict output power production to minimize active power using the Variational Mode Decomposition (VMD) algorithm, the Ensemble method, and the most influential features in the Likupang case study.

1.2 Statement of the Problem

The research addresses the challenge of forecasting the electricity power output of the Solar Power Plant (SPP) in Likupang. Since SPP's electricity generation is highly dependent on solar radiation and weather conditions, accurate forecasting is crucial for maintaining electricity stability on the grid. The specific problem the research aims to solve is how to effectively model features using machine learning and signal decomposition algorithms to improve the accuracy of these forecasts, thereby reducing operational costs and the reliance on gensets.

1.3 Objective and Hypothesis

1.3.1 Objective

The primary objective of this research is to develop a highly accurate solar power prediction model, a task of utmost importance in the renewable energy industry. This will be achieved by leveraging Variational Mode Decomposition (VMD) for feature extraction and optimized ensemble learning techniques, specifically on bagging and stacking methods. The model aims are:

1. How to choose the weather parameters that impact electricity generation in Solar Power Plants.
2. How to find the best model combination for ensemble learning; stacking bagging.
3. How to compare the performance of machine learning for predicting electricity production in Solar Power Plants by using Variational Mode Decomposition and without using Variational Mode Decomposition.

1.3.2 Hypothesis

In consonance with the research objectives, the study has formulated the following hypotheses to guide the investigation and assess the efficacy of the proposed methodologies:

1. Electricity production in SPP is heavily influenced by certain weather parameters.
2. An effective machine learning model that accurately forecasts electricity production in Solar Power Plants (SPP) can be developed by selecting specific weather parameters as input features, resulting in higher prediction accuracy.
3. Enhancing models for forecasting electricity production in SPP can be achieved by applying feature engineering and the Variational Mode Decomposition (VMD) method to incorporate additional features.

1.4 Assumption

Certain foundational assumptions must be established in conducting this research on solar power prediction using machine learning and feature extraction techniques such as Variational Mode Decomposition (VMD). These assumptions provide a basis for the methodologies adopted in this study and ensure the validity and reliability of the results obtained. They help define the scope and boundaries of the research, ensuring that the objectives and hypotheses are tested under controlled conditions that are realistic and applicable to real-world scenarios. The following assumptions are considered:

1. Relevance of Weather Parameters: The chosen weather parameters are assumed to be the most relevant for predicting solar power output in the studied region.
2. Model Generalizability: The developed machine learning models, including those using VMD for feature extraction and ensemble learning methods, are assumed to be accurate and highly versatile. They are expected to be generalizable to different

locations and types of solar power plants, provided similar weather and environmental conditions.

3. Impact of VMD on Prediction Accuracy: VMD is assumed to play a significant role in improving the accuracy and reliability of solar power predictions. By better handling non-linearities and noise in the data, VMD is expected to enhance the precision of our predictions.

1.5 Scope and Delimitation

This study focuses on developing an accurate model for short-term solar power prediction by incorporating key independent variables such as weather parameters (Surface Temperature, Dew Point Temperature, Surface Radiation Downward Clear Sky, Surface Radiation Downward, Wind Speed, Wind Direction, Relative Humidity) and temporal features. The study is specifically conducted in Likupang, Minahasa, Indonesia, a region chosen for its high potential for solar energy production due to its tropical climate and geographical characteristics. The timeframe of the study encompasses short-term prediction periods, ranging from minutes to several hours ahead, which is essential for effective grid management and energy optimization.

1.6 Significance of the Study

This study introduces a novel short-term solar power prediction approach by integrating Variational Mode Decomposition (VMD) for feature extraction with ensemble learning methods, specifically bagging and stacking techniques. By decomposing the solar power signal into intrinsic mode functions (IMFs) using VMD and selecting relevant weather parameters through spatial correlation methods, the proposed model aims to significantly enhance solar power output's prediction accuracy. This improvement in forecasting precision is vital for effective energy management, grid stability, and optimizing solar power plant operations. Additionally, incorporating weather parameters and temporal features into the prediction model addresses the inherent variability of solar energy, contributing to more reliable and efficient renewable energy planning and utilization. By integrating these theories and concepts, this research aims to develop a comprehensive short-term model for solar power prediction. It utilizes Variational Mode Decomposition (VMD) for feature extraction to improve model inputs and compares ensemble learning techniques, such as stacking and bagging, for prediction accuracy. The development of this short-term solar power prediction model ultimately helps PLN efficiently manage electricity supply to cover deficits not met by solar power production, thereby reducing the cost of using gensets.