成都理工大學中英合作办学
CDUT Sino-British Collaborative Education

OXFORD BROOKES UNIVERSITY

UNDERGRADUATE PROJECT REPORT

Project Title:	**Wind Power Prediction Using Attention-PowerWiNet**
Surname:	**HE**
First Name:	**SHIJIA**
Student Number:	**202118010109**
Supervisor Name:	**Grace Ugochi Nneji**
Module Code:	**CHC 6096**
Module Name:	**Project**
Date Submitted:	**May 6, 2025**

Chengdu University of Technology Oxford Brookes College

Chengdu University of Technology

BSc (Single Honours) Degree Project

Programme Name: Projects

Module No.: CHC 6096

Surname: He

First Name: Shi Jia

Project Title: Wind Power Prediction Using Attention-PowerWiNet

Student No.: 202118010109

Supervisor: Grace Ugochi Nneji

Date submitted: May 6 2025

A report submitted as part of the requirements for the degree of BSc (Hons) in Computer Science

At

Chengdu University of Technology Oxford Brookes College

Declaration

Student Conduct Regulations:

Please ensure you are familiar with the regulations in relation to Academic Integrity. The University takes this issue very seriously and students have been expelled or had their degrees withheld for cheating in assessment. It is important that students having difficulties with their work should seek help from their tutors rather than be tempted to use unfair means to gain marks. Students should not risk losing their degree and undermining all the work they have done towards it. You are expected to have familiarized yourself with these regulations. https://www.brookes.ac.uk/regulations/current/appeals-complaints-and-conduct/c1-1/

Guidance on the correct use of references can be found on www.brookes.ac.uk/services/library, and also in a handout in the library.

The full regulations may be accessed on-line at https://www.brookes.ac.uk/students/sirt/student-conduct/. If you do not understand what any of these terms mean, you should ask your Project Supervisor to clarify them for you.

I declare that I have read and understood Regulations C1.1.4 of the Regulations governing Academic Misconduct, and that the work I submit is fully in accordance with them

Signature: Nicole (He ShiJia) Date: 6/05/2025

REGULATIONS GOVERNING THE DEPOSIT AND USE OF OXFORD BROOKES UNIVERSITY MODULAR PROGRAMME PROJECTS AND DISSERTATIONS.

Copies of projects/dissertations, submitted in fulfilment of Modular Programme requirements and achieving marks of 60% or above, shall normally be kept by the library.

I agree that this dissertation may be available for reading and photocopying in accordance with the Regulations governing use of the library.

Signature: Nicole (He ShiJia) Date: 6/05/2025

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Dr. Grace Ugochi Nneji, for her steadfast guidance and unwavering support throughout the completion of my undergraduate project. Her expertise, patience, and encouragement have been truly pivotal in shaping the direction and success of this work. My heartfelt thanks also go to Joojo Walker, the module leader, as well as to all the dedicated educators who have shared their knowledge and offered invaluable insights during my undergraduate studies.

Furthermore, I am sincerely grateful for the academic resources and supportive environment provided through the collaborative efforts between Oxford Brookes University and Chengdu University of Technology. These facilities have greatly contributed to my academic development.

Lastly, I would like to extend my heartfelt appreciation to my beloved family and friends. Your unwavering love, support, and encouragement have been a constant source of strength and inspiration. I am deeply thankful for your presence in my life.

Table of Contents

List of Figures

List of Table

Abstract

Wind power is an essential renewable energy source, but its reliance on unpredictable environmental conditions makes forecasting a considerable task. Variables like variable wind speed, direction, and temperature frequently constrain predictive accuracy and influence the operational efficiency of wind farms. To address this, this project proposes a hybrid deep learning model called Attention-PowerWiNet, which integrates PowerNet, WiNet, and an attention mechanism to capture both spatial and temporal features while highlighting the most relevant inputs. Two real-world wind farm datasets were used: one seasonally segmented and one continuous. Comparative experiments were performed on four models: PowerNet, WiNet, PowerWiNet, and Attention-PowerWiNet. Results show that Attention-PowerWiNet achieves the best performance across all evaluation metrics with MAE of 0.035, RMSE of 2.314, and R^2 of 0.906, especially in summer, autumn and winter subsets. Furthermore, also introduces the SHAP method to enhance the interpretability of the model and discovers that wind speed and cloud cover are the most important factors in the prediction. For the convenience of practical application, a graphical user interface (GUI) has also been developed, which supports data upload, obtaining prediction results and visual interpretation. This project provides an effective solution for improving the accuracy and transparency of wind power prediction.

Keywords: wind power forecasting, deep learning, attention mechanism, SHAP

Abbreviations

WiNet: Wind Forecasting Memory Network

PowerNet: Power Network

PowerWiNet: Power Wind Forecasting Memory Network

Attention-PowerWiNet: Attention-Power Wind Forecasting Memory Network

CNN: Convolutional Neural Network

LSTM: Long Short-Term Memory

MAE: Mean Absolute Error

MSE: Mean Squared Error

RMSE: Root Mean Squared Error

R^2: Coefficient of Determination

SHAP: SHapley Additive exPlanations

GUI: Graphical User Interface

CSV: Comma-Separated Values

ReLU: Rectified Linear Unit

IPCC: Intergovernmental Panel on Climate Change

DGF: Double Gaussian Function

KDE: Kernel Density Estimation

NWP: Numerical Weather Prediction

Glossary

Wind Power Forecasting: The process of predicting the amount of electrical power that will be generated by wind turbines based on historical and environmental data.

Deep Learning: A subset of machine learning that uses neural networks with many layers (deep architectures) to model complex patterns in large datasets.

CNN (Convolutional Neural Network): A type of deep learning model particularly effective in extracting local features from data, such as spatial patterns in images or time-series data.

LSTM (Long Short-Term Memory): A type of recurrent neural network (RNN) that is designed to learn and retain long-term dependencies in sequential data, often used for time-series forecasting.

Attention Mechanism: A technique in neural networks that allows the model to focus on the most relevant parts of the input sequence when making predictions, improving performance and interpretability.

SHAP (SHapley Additive exPlanations): A method based on cooperative game theory used to explain individual predictions of machine learning models by assigning each feature an importance value for a particular prediction.

PowerNet (Power Network): A baseline deep learning model composed of convolutional and pooling layers, followed by a dense output layer.

WiNet (WindFormer Network): A recurrent neural network model consisting of stacked LSTM layers with dropout.

PowerWiNet (Power WindFormer Network): A hybrid deep learning model that combines convolutional layers and LSTM units to learn both spatial and temporal features for wind power prediction.

Attention-PowerWiNet (Attention-Power WindFormer Network): An enhanced version of PowerWiNet that incorporates an attention mechanism to further improve prediction accuracy and interpretability.

Chapter 1 Introduction

1.1 Background

Fossil fuels, such as coal, oil, and natural gas, continue to dominate global energy consumption, accounting for over 80% of the total primary energy supply. The impact of rising primary energy consumption on climate change and the environment has become a significant challenge [1]. To keep the average increase in global temperature below 2°C, a worldwide energy transition is essential. The energy industry is significantly impacted by the Paris Agreement, which calls for a shift from fossil fuels to low-carbon alternatives because CO_2 emissions account for approximately two-thirds of total greenhouse gases emissions. This transition is driven by technological innovations, particularly in renewable energy [2]. Renewable energy is gaining increasing attention as its eco-friendly nature and minimal air pollutant emissions, driven by growing public awareness of environmental protection. In addition to promoting sustainability, renewable energy also holds significant economic value. By utilizing natural and renewable resources, it lowers electricity generation costs and supports long-term economic benefits [3]. Shifting to renewable energy for power generation is crucial because it lowers carbon emissions, combats climate change, promotes environmental sustainability, and supports long-term socio-economic advantages like sustainable economic development [4].

1.1.1 Risk and Factor

The fuels of fossil fuels emit a lot of toxic gases and carbon dioxide. The emissions are seriously harmful to human health, and excessive inhalation by humans may cause diseases and cause developmental disorders in children. In addition, it also causes serious environmental pollution and climate change [5]. Human activities have caused around 1.0°C of warming over pre-industrial levels, which, according to the Intergovernmental Panel on Climate Change (IPCC), is expected to increase to 1.5°C by 2030-2052 [6]. Carbon dioxide from fossil fuel combustion is the primary greenhouse gas responsible for climate change.

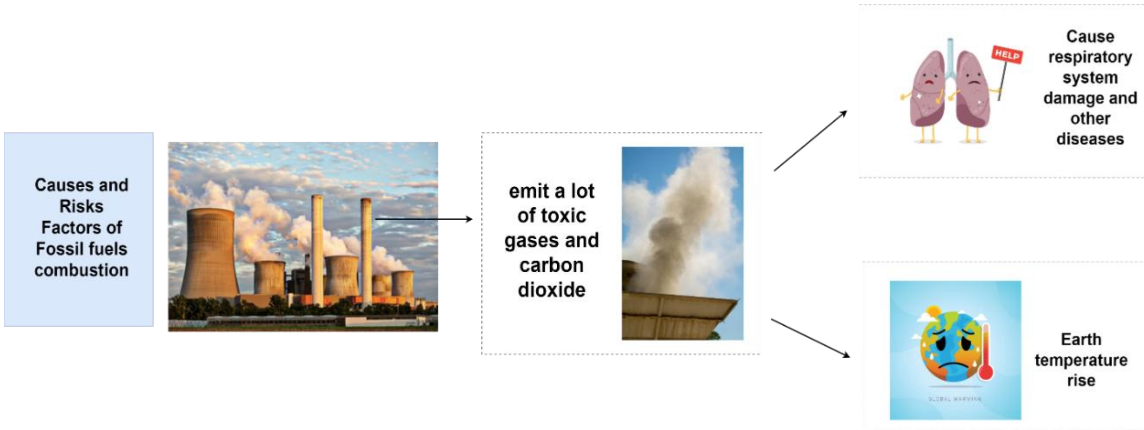


Figure 1. Risks of burning fossil fuels

The importance of renewable energy sources to the world's energy supply is growing as the world faces environmental challenges from global warming [6]. Wind power energy is rapidly becoming one of the renewable energy industry's fastest-growing and most competitive sectors, a trend exemplified by countries like Denmark and Germany. Denmark generates over 50% of its electricity from wind on average, while Germany's wind power share reached more than 30% during the first half of 2020 with its larger grid demands [7]. In addition, wind energy has proven itself to be an effective and sustainable alternative to conventional energy sources, with the capacity to satisfy the world's energy needs while causing the least amount of environmental damage. Unlike fossil fuels, which release greenhouse gases into the atmosphere and contribute to climate change, wind energy is clean and environmentally friendly [8].

In the aspect of the economy, renewable energy has become a crucial driver of global economic development and is experiencing a continuous upward trajectory. Renewable energy has now secured a significant portion of the energy market and is steadily strengthening its position as a key player in the global energy supply system. The wind energy industry continues to show strong and steady growth although other industries recessions as facing a financial crisis. According to some research, the wind power industry is growing at a higher annual rate than other industries in the world as high as 20% to 30%. In addition, the wind power industry also eased the pressure of employment which provides 440,000 employment positions globally [9]. However, despite the impressive growth of the wind energy sector, the efficiency of its energy output still faces problems. Improving wind energy usage and quickening the energy transition process require increasing the wind farm’s wind energy capture efficiency [10].

1.1.2 Challenge

Predicting wind energy production comes with several significant challenges. The wind speed is the main factor determining power. Unfortunately, wind speed can fluctuate significantly over time which is one of the main challenges, making it difficult to predict wind energy production. These fluctuations not only affect short-term power generation but also impact the long-term planning and operation of wind farms. In addition to wind speed fluctuations, various environmental factors significantly impact wind power production. Temperature, air pressure, and humidity influence wind patterns, also affecting both short-term and long-term wind energy output. Additionally, the quality and availability of historical wind farm data often pose issues—missing values, noise, and inconsistencies can degrade model performance. Finally, develop an efficiency predictive models is crucial for practical deployment [11]. Advanced deep learning models, such as CNN-LSTM architectures, have shown promise in capturing spatial and temporal patterns. Addressing these challenges requires continued innovation in forecasting methodologies, improved data processing technique to enhance the precision and reliability of wind energy predictions.

1.2 Aim

In this project, a hybrid deep learning model called Attention-PowerWiNet is used, combining the strengths of PowerNet, WiNet, and attention mechanism. WindFormer Network is employed to capture long-term dependencies in time-series data, while Power Network can be used to extract patterns or relationships between multiple variables in the dataset. The project aims to predict the efficiency of wind farms by analysing historical data in CSV format. By considering time-dependent environmental factors, such as wind speed, direction, and temperature, this approach enhances the accuracy of wind energy production predictions, helping optimize wind farm operations and overall energy output efficiency.

1.3 Objectives

The project will collect two datasets from different wind farms. The dataset1 contains operational data from a wind farm for the year of 2017, while the dataset2 provides time-dependent environmental factors for the wind farm from 2018 to 2020. The dataset1 was divided into four seasonal groups (spring, summer, autumn, and winter) based on the timestamp, and predictions are made for each season separately. The dataset2 is kept intact and split into training, testing, and validation sets with the same proportions (70%, 15%, 15%).

The project aims to build an Attention-PowerWiNet, with hyperparameters such as batch size, dropout rate, and others being tuned for optimal performance. The PowerNet and WiNet models are trained separately to make predictions on the two datasets. Finally, the performance of the PowerNet model, WiNet model, the PowerWiNet and Attention-PowerWiNet model is compared to prove that the Attention-PowerWiNet model will perform better than the individual model. In addition, the evaluation of the model will include metrics such as "MeanAbsoluteError" and "Loss." Moreover, performance will be assessed using RMSE and R2. After the model is trained, five SHAP plots will be used for the interpretability of the model.

Lastly, the project will be deployed through a website that allows uploading of wind farm historical data, and then gives the wind power prediction values result and with SHAP plots.

1.4 Project Overview

This section explores the potential of combining PowerNetwork, WindFormer Network, and attention mechanism to improve wind power production forecasts, emphasizing the benefits for various stakeholders, including wind farms, grid managers, governments, and power purchasers.

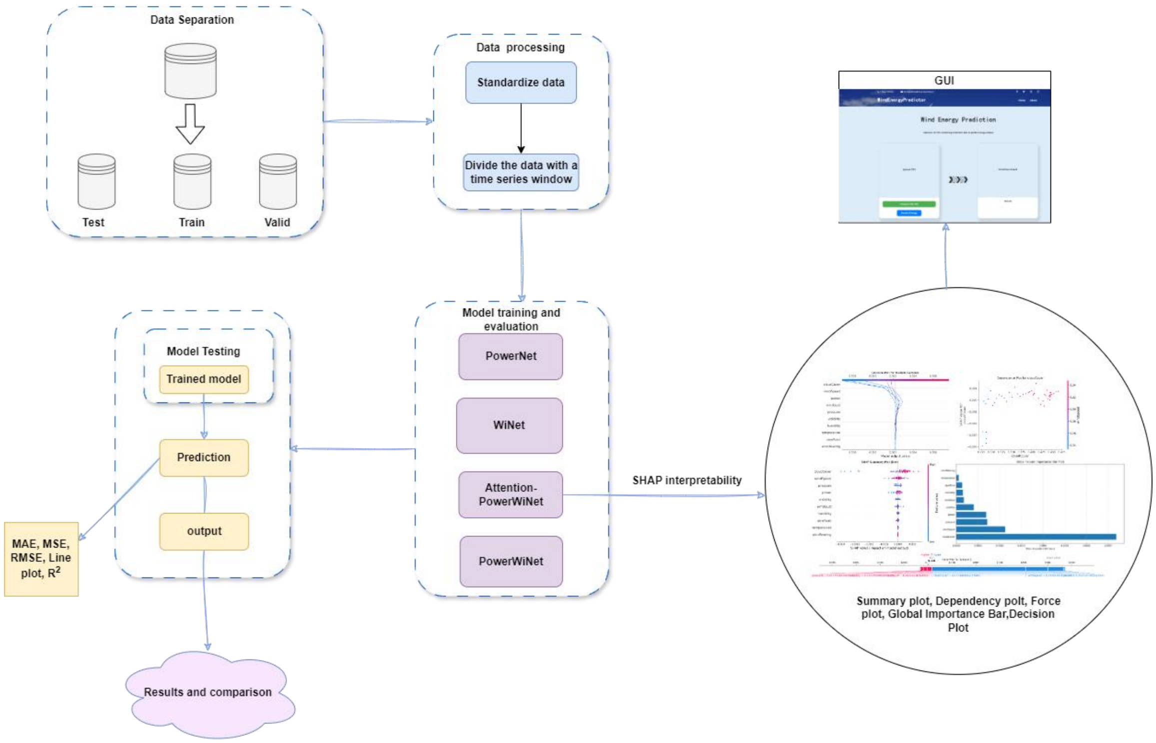


Figure 2. Project Overview

1.4.1 Scope

In this project, the PowerNetwork model is used to extract features from the historical wind farm data (in CSV format), while the WindFormer Network model processes these features to predict wind power production. The model combines convolutional layers for feature extraction and LSTM layers for capturing temporal dependencies.

The significance of this study contains:

- Combining multiple data improves the accuracy of wind energy production of wind farm forecasts.
- Helps maximize the use of renewable energy sources.
- Improve the stability of the power system.
- Reduce the operating cost of wind farms.
- Environment friendly.

1.4.2 Audience

- For the wind farm: this project not only increases the wind energy output by improving the conversion efficiency of wind energy. It also provides more accurate data to support wind farm managers to optimize management strategies and improve overall operational efficiency.

- For grid managers: this project helps them better allocate wind power and ensure the stability of the power supply. In addition, it also could reduce wind power supply fluctuations and enhance the reliability of the grid.
- For the government: this project can help better achieve its energy strategy and sustainable development goals.
- For the purchaser of power: this project can reduce power procurement costs because of the increase in wind energy output.

In summary, using the hybrid model to predict wind power will bring benefits for the wind farm, grid manager, government, and the purchase of power by optimizing wind farm efficiency based on environmental factors.

Chapter 2 Background Review

Many researches are focusing on predicting the wind energy output of wind farms by using the collected data and various parameters. This section will present the works that had been conducted to predicting the wind energy output.

2.1 Wind Power Prediction using Tradition learning Technique

Wang et al. [12] introduced a hybrid system WPFSVer1.0, which combines statistical method and physical method. This approach relies on the understanding of the numerical weather prediction (NWP) techniques and data from wind farms, historical weather, and generation parameters. There was communication between the five core parts in the system. Figure 3 show the frame of WPFSVer1.0.

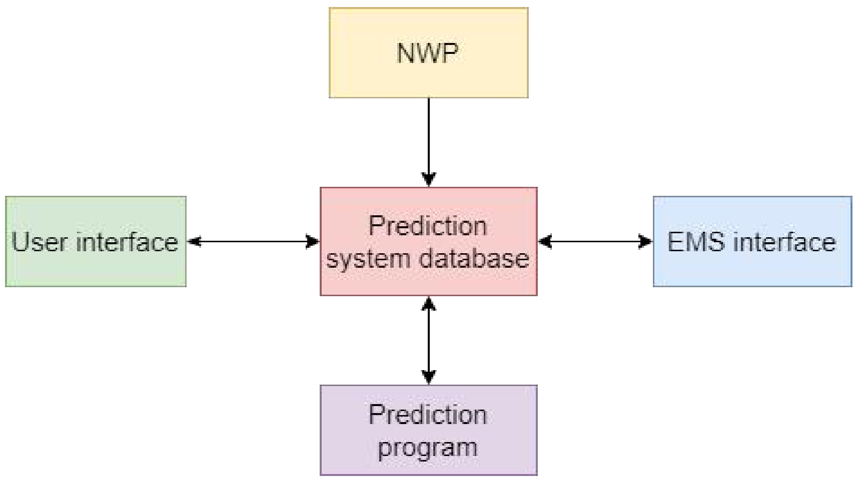


Figure 3. WPFSVer1.0 Frame by Wang et al. [12]

2.2 Wind Power Prediction using Machine learning Techniques

Kusiak et al. [11] proposed a 10-min Time Series Prediction Model which use the KNN algorithm. The time series model is initially used to predict future wind speed. This predicted wind speed is then fed into the KNN model to calculate the power output of the wind farm. Figure 4 shows the structure of the model. It uses 3476 data points from 100 turbines in a wind farm to develop this prediction model and uses 871 data points to test the model's performance. The KNN model (k = 250) achieved a MAE of 1.689 MW and a mean relative error of 4.231%.

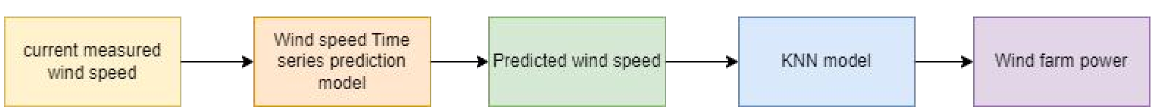


Figure 4. Integrated prediction model Frame by Kusiak et al. [11]

2.3 Wind Power Prediction using Deep learning Techniques

2.3.1 Convolutional Neural Network Method

Hong and Rioflorido [13] developed a hybrid deep learning neural network method which is based on Convolutional Neural Network (CNN) and cascaded with a new Radial Basis Function Neural Network (RBFNN) using a Double Gaussian function (DGF) for activation. Firstly, using CNN to capture the features from wind power time-series. Secondly, the features will as the input to the new RBFNN in the hidden layer. It used historical wind power generation dataset from the Changgong Wind Farm. The value of coefficient of determination, R^2 is achieving 0.9 with this method. Figure 5 shows the basic structure of this approach.

Figure 5. hybrid deep learning neural network method by Hong and Rioflorido [13]

2.3.2 Long Short-Term Memory Method

Zhou et al. [14] introduced a method that combines K-Means clustering with Long Short-Term Memory (LSTM) networks for wind power prediction. and a nonparametric Kernel Density Estimation (KDE) model with bandwidth optimization for wind power probabilistic interval prediction.

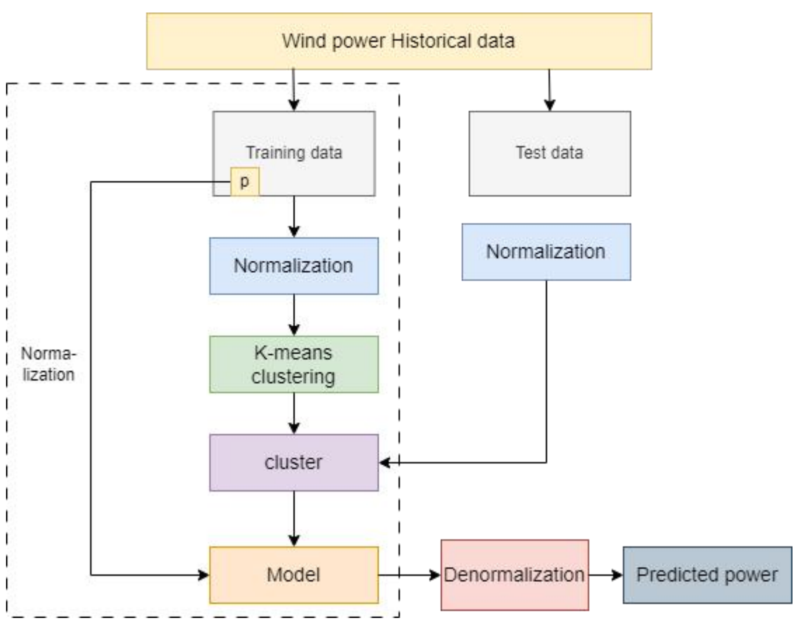


Figure 6. K-Means-LSTM Architecture by Zhou et al. [14]

The experimental results in wind farm A show that the K=6_LSTM model achieves MAE, MAPE, and RMSE of 53.182, 12.490, and 73.017, respectively, demonstrating superior performance compared to other models. Figure 6 shows the basic structure of this approach.

2.3.3 Convolutional Neural Network-Long Short-Term Memory Method

Zhang et al. [15] utilized a CNN-LSTM model for wind power prediction. First, CNN is employed to extract key features from wind power data and its influencing factors. These extracted features are then fed into an LSTM network, which captures temporal dependencies to generate accurate time-series predictions. Figure 7 displayed the structure of CNN-LSTM model.

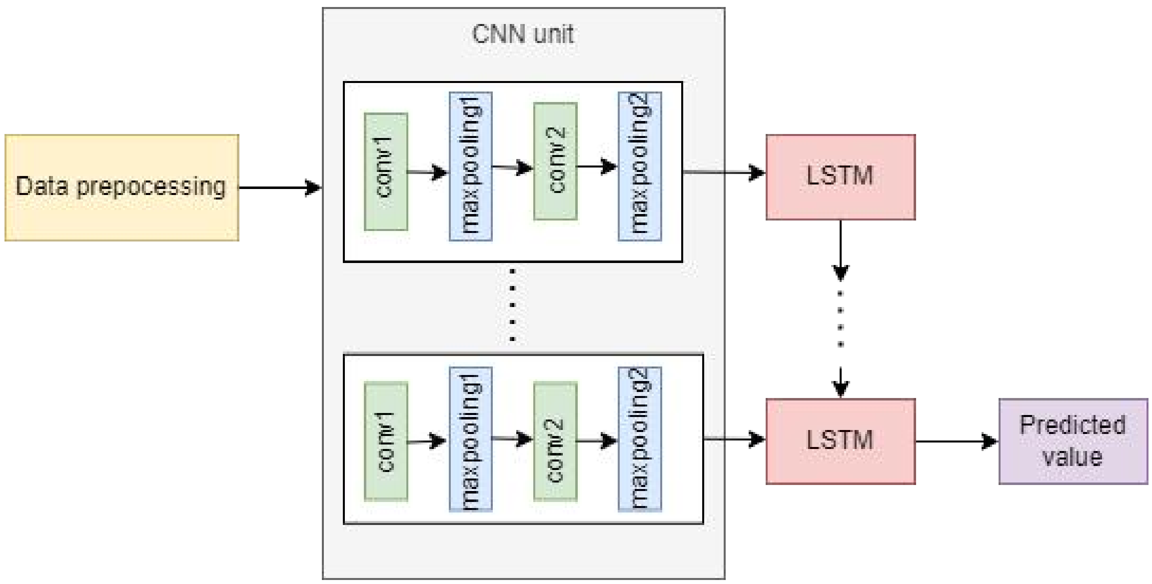


Figure 7. CNN-LSTM architecture by Zhang et al. [15]

Table 1. Summary of related works

Author	Datasets	Model & method	Limitation	Result
Wang et al. [12]	wind farms, historical weather, and generation parameters	physical and statistical approach	High model complexity and substantial training demands. Implementation and upkeep	\

			expenses.	
Kusiak et al. [11]	3476 data points from 100 turbines	10-min Time Series Prediction Model integrate the KNN algorithm.	The KNN model's wind speed predictor is overly sensitive.	MAE of 1.689 MW and a mean relative error of 4.231%.
Hong and Riofloridо [13]	historical wind power generation dataset from the Changgong Wind Farm	hybrid deep learning neural network method based on CNN	It requires more computation and resources for training.	R^2 achieving 0.9.
Zhou et al. [14]	wind power historical dataset	K-Means-long short-term memory	Cluster Dependence and Sensitivity and Computational Complexity.	MAE, MAPE, and RMSE of 53.182, 12.490, and 73.017, respectively
Zhang et al. [15]	full-year data of a wind farm in a certain area of China in 2018	CNN-LSTM	Loss of Temporal Information in CNN Feature Extraction.	The square indicator of R is above 0.84.

Chapter 3 Methodology

3.1 Approach

This section presents the overall proposed model called Attention PowerNet which integrates Power Network, WindFormer Network, and Attention mechanism. This proposed model developed and evaluated to forecast wind power generation. Also, this section describes the dataset used, including the data preprocessing steps, and then followed by the experimental setup, project version control strategy, and performance evaluation metrics.

3.2 Proposed Model Structure

This section presents the architectures of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet, along with a detailed discussion of the attention mechanism employed in Attention-PowerWiNet.

3.2.1 Power Network (PowerNet)

The architecture consists of an input layer, two convolutional layers, two pooling layers, a flattening layer, and an output layer, with each component designed as follows. The Input layer is defined with the shape (the sequence length and the number of features at each time step). Next, two 1D convolutional layers: the first convolutional layer applies 128 convolutional filters, each with a kernel size of 2, a stride of 1, and "same" padding. The ReLU activation function is used to introduce non-linearity. This layer captures local temporal dependencies by extracting features from adjacent time steps. The second convolutional layer reduces the number of filters to 64 while keeping the other parameters identical. This step extracts higher-level temporal features and enhances the model's representational capacity. Following the convolutional layers, two 1D max pooling layers: with a pool size of 2. The pooling operation reduces the dimensionality of the feature maps by down sampling, thereby retaining the most significant information while reducing computational complexity and mitigating overfitting. Afterward, a flattening layer: is applied to convert the multi-dimensional feature maps into a one-dimensional vector. Finally, the output layer: consists of a single neuron, producing the regression output.

The model is compiled using the Adam optimizer. The mean squared error (MSE) is chosen as the loss function, while the mean absolute error (MAE) is used as an evaluation metric. The architecture of the PowerNetwork is shown in Figure 8.

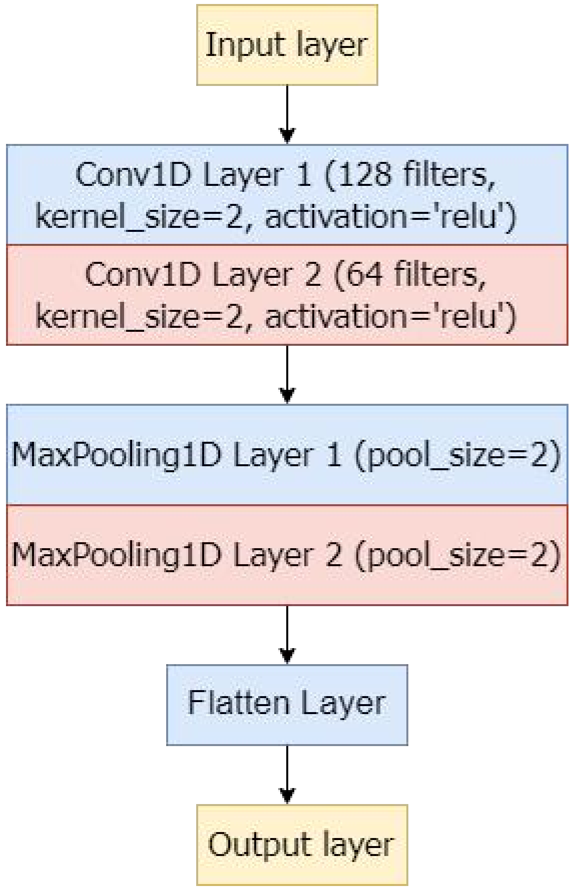


Figure 8. PowerNet architecture

3.2.2 Wind Forecasting Memory Network (WiNet)

The architecture consists of an input layer, two LSTM layers, a dropout layer, and an output layer, with each component designed as follows. The input layer: is defined with the shape (the sequence length and the number of features at each time step). Following this, the first LSTM layer contains 32 units with tanh activation, allowing the model to capture temporal dependencies in the input sequences. To mitigate overfitting and enhance generalization, a dropout layer applies a 30% dropout rate to the outputs of the first LSTM layer to reduce overfitting and improve generalization. Subsequently, the second LSTM layer is introduced, which contains 16 units with tanh activation, outputs only the final hidden state. Finally, the output layer is a dense layer with one neuron and a linear activation function, mapping the extracted temporal features to the final regression output.

The model is compiled using the Adam optimizer. The mean squared error (MSE) is chosen as the loss function, while the mean absolute error (MAE) is used as an evaluation metric. The architecture of the WindFormer Network is shown in Figure 9.

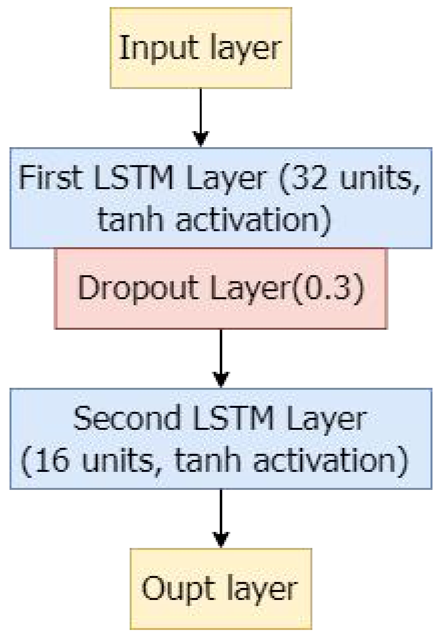


Figure 9. WiNet architecture

3.2.3 Power Wind Forecasting Memory Network (PowerWiNet)

The architecture consists of an input layer, a convolutional layer, a max-pooling layer, two LSTM layers, a dropout layer, and an output layer, with each component designed as follows. The input layer is defined with the shape (the sequence length and the number of features at each time step). Next, the first Convolutional layer contains 64 filters with a kernel size of 2 and ReLU activation. This layer helps capture local temporal patterns in the input sequence. Following the convolution, a 1D max pooling layer with a pool size of 2 is applied to reduce the dimensionality of the feature maps and retain the most salient features. After pooling, the first LSTM layer consists of 32 units and uses the tanh activation function. To prevent overfitting, a dropout layer with a dropout rate of 0.1 is added after the first LSTM layer to prevent overfitting by randomly dropping 10% of the units during training. Then, the second LSTM layer, which contains 16 units and outputs only the final hidden state. Finally, the output layer is a dense layer with one neuron and a sigmoid activation function.

The model is compiled using the Adam optimizer. The mean squared error (MSE) is chosen as the loss function, while the mean absolute error (MAE) is used as an evaluation metric. The architecture of the PowerWiNet model is shown in Figure 10.

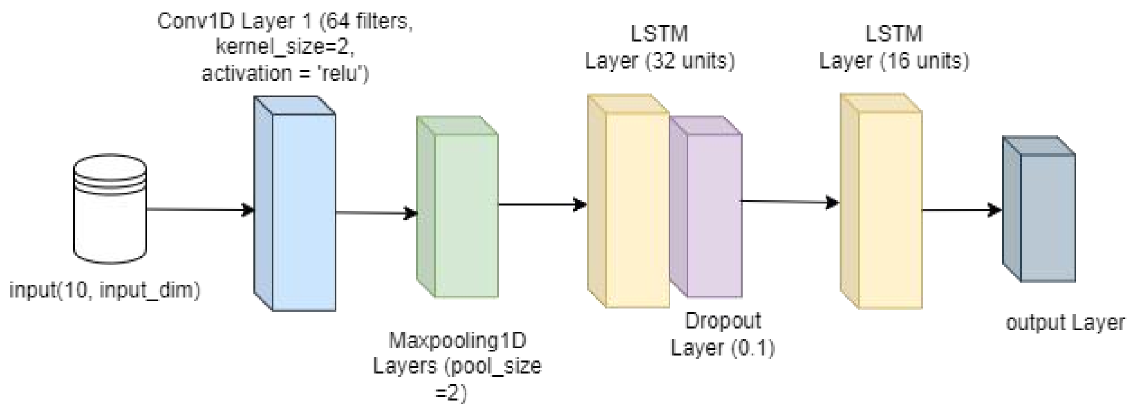


Figure 10. PowerWiNet architecture

3.2.4 Attention-Power Wind Forecasting Memory Network (Attention-PowerWiNet)

This model adds the attention mechanism to the PowerWiNet in 3.1.3. The architecture of the attention mechanism is shown in Figure 11, and the architecture of the overall attention-PowerWiNet is shown in Figure 12.

The attention mechanism consists of two main components designed as follows. First, the Positional Encoding Layer introduces positional information into the input sequence, allowing the model to recognize the order of time steps. It generates a set of fixed sine and cosine values for each position and adds them to the original input vectors. This approach preserves temporal characteristics without using recurrent structures, enhancing the model's ability to understand sequential data. Next, the MultiHeadSelfAttention Layer implements a key component of the Transformer encoder architecture. It first applies TensorFlow’s built-in MultiHeadAttention layer to perform self-attention, allowing the model to capture dependencies between different time steps in the input sequence. Then, it uses a residual connection (adding the original input to the attention output) followed by layer normalization to enhance model stability and convergence. This design effectively models long-range dependencies while preserving original input features.

Overall, the positional encoding and multi-head self-attention mechanism allow the model to sense the sequence order while effectively capturing dependencies between different time steps and stabilizing the training process with residuals linking and normalization.

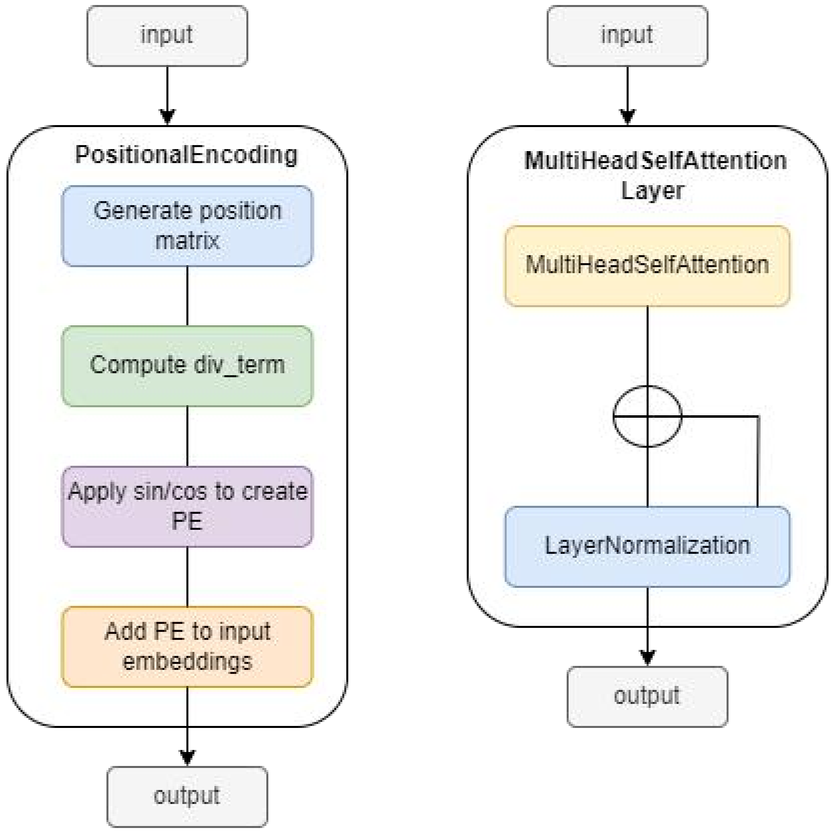


Figure 11. Attention mechanism architecture

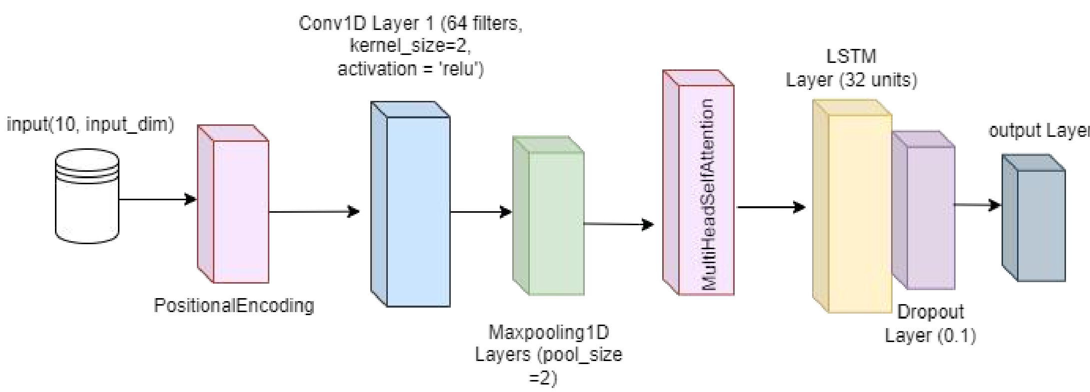


Figure 12. Attention-PowerWiNet architecture

3.3 Dataset

This project uses two historical datasets from wind farms as experimental data to enhance the adaptability and generalization ability of the model. The first dataset is divided into four seasons—spring, summer, autumn, and winter to analyse the impact of seasonal variation on wind energy production. The second dataset preserves the entire time series to assess the model's performance against the overall operational trend. By applying two different data processing strategies, the model's stability and predictive performance can be validated from multiple perspectives, providing more comprehensive experimental support for wind energy forecasting.

3.3.1 Dataset 1

The first dataset contains operational data from a wind farm for the entire year of 2017, including data/time, wind speed, wind direction, temperature, air pressure, humidity, practical power and capacity. The dataset is in CSV format and contains 35,040 rows, the first of which was recorded on January 1, 2017, at 12 am, and a new row was recorded every 15 minutes.

3.3.1.1 Data Split

The historical data measured by a complete year of wind farms are divided into four data groups according to months: spring (March to May), summer (June to August), autumn (September to November) and winter (the remaining months). The four data groups are divided into training, validation, and test data by 0.7, 0.15, and 0.15 respectively. The division into four data groups was done manually, and they were stored separately in four CVS files. The training, validation, and test splits are all done in Python code. The structure of the data separation is shown in Figure 13.

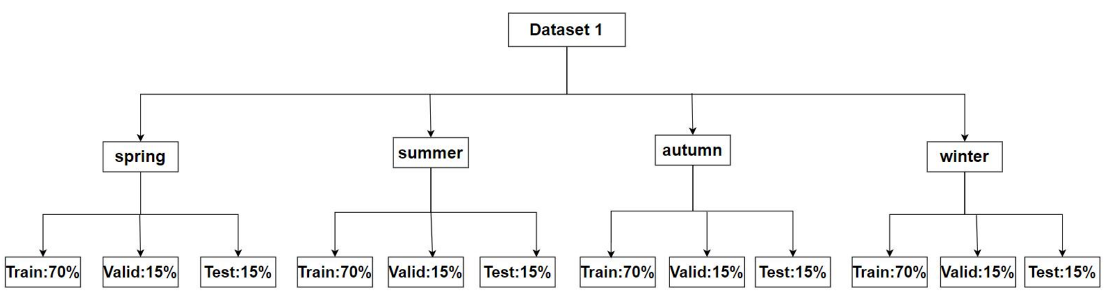


Figure 13. Dataset structure of Season-based Wind Power Prediction

3.3.2 Dataset 2

The second dataset provides time-dependent environmental factors for the wind farm from 2018 to 2020, including cloud cover, dewpoint, humidity, ozone, pressure, temperature, visibility, wind bearing, wind gust, wind speed, and power. This dataset has 14,554 rows in CSV files. The first row was recorded at 3:30 am on June 4, 2018, and a new row was recorded every hour.

3.3.2.1 Data Split

Divide the whole dataset into a training set, validation set, and test set according to the ratio of 0.7, 0.15, and 0.15. The above data-splitting process is implemented using Python code. The structure of the data separation is shown in Figure 14.

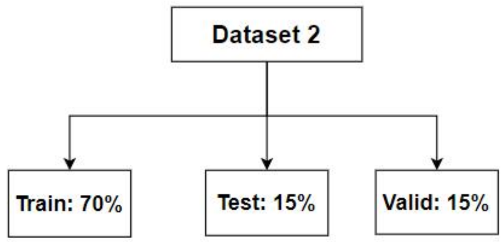


Figure 14. Dataset structure of Wind Power Prediction

3.3.3 Data Preprocessing

First, standardize the training, validation, and test data using the Min-Max normalization method to scale the data to a specified range (default is between 0 and 1). Then, divide the data with a time series window, based on the training set, validation set, and test set, the data is further divided by a sliding window, so that the PowerWiNet model can be trained on the time series to capture the time dependence.

3.4 Experimental Setup and Technology

The parameter settings of the four models, PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet, are shown in the Table 2-5, as well as the technology usage of this project as shown in the table 6.

PowerNetwork Parameters settings:

Table 2. PowerNetwork setting

Network layer	Parameter	Value
Conv1D_1	kernel size	2
	filters	128
Conv1D_2	kernel size	2
	filters	64
MaxPooling1D_1	pool size	2

MaxPooling1D_2	pool size	2
Flatten_1	-	-
Dense_1	units	1

WindFormer Network Parameters settings:

Table 3. WindFormer Network setting

Network layer	Parameter	Value
LSTM_1	units	32
Dropout_1	rate	0.3
LSTM_2	units	16
Dense_1	units	1

PowerWiNet Parameters settings:

Table 4. PowerWiNet setting

Network layer	Parameter	Value
Conv1D_1	kernel size	2
	filters	64
MaxPooling1D_1	pool size	2
LSTM_1	units	32
Dropout_1	rate	0.1
LSTM_2	units	16
Dense	units	1

Attention-PowerWiNet Parameters settings:

Table 5. Attention-PowerWiNet setting

Network layer	Parameter	Value
PositionalEncoding	max_len	TIME_STEPS
	embed_dim	input dim
Conv1D_1	kernel size	2
	filters	64
MaxPooling1D_1	pool size	2
MultiHeadSelfAttention_1	embed_dim	64
	num_heads	4
LayerNormalization_1 (inside Attention)	epsilon	1e-6
LSTM_1	units	32
Dropout_1	rate	0.1
Dense_1	units	1

Table 6. Summary of Relevant Technology

Software	Framework	TensorFlow
	Language	Python
	Libraries	Keras, Numpy
Hardware	Central processing unit (CPU)	AMD Ryzen 7 5800H
	Graphic Processing Unit (GPU)	NVIDIA GeForce GTX 1650

3.5 Project Version Management

The different versions of code modification will be stored in a local file and uploaded to git repositories.

Git repositories link: https://github.com/Nicole-HeShiJia/Project.

Table 7. Version Control Progress

Version Number	Code Name	Content	Results
1	Models code for basic testing of dataset 2.	Data processing, PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet model code, one original dataset.	The training results, and plots of actual versus predicted values.
2	Divide dataset 1 into four groups of data	Data processing, PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet model code. Four spring, summer, autumn, and winter groups datasets.	The training results, and plots of actual versus predicted values

3.6 Evaluation Metrics

This project will evaluate the performance of the model through the following standards.

- MSE (loss): MSE is the average of the squared differences between the predicted and true values and is used to measure how much the model's predictions deviate from the actual values with the expression presented in equation (1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (1)$$

- MAE: MAE is the average of the absolute difference between the predicted and true values. It measures the size of the errors in the predictions without considering their direction. As shown in the equation (2).

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{2}$$

- RMSE: RMSE is the square root of the average of the squared differences between the predicted and actual values. It standardizes the MSE based on the formula provided in the equation (3).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i \quad)^2} \tag{3}$$

- R^2: According to the expression given in Equation (4), R^2 measures the proportion of variance in the dependent variable that is predictable from the independent variables. It reflects how well the model fits the data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i \quad)^2}{\sum_{i=1}^{n} (y_i - \bar{y} \quad)^2} \tag{4}$$

- Line Plot: compares the predicted and actual values from models on the test set, which helps visualize the model's prediction performance,

In summary, to evaluate the model's performance on wind power prediction, this project will conduct a comprehensive evaluation, including MAE, RMSE, MSE, R^2, and line plot

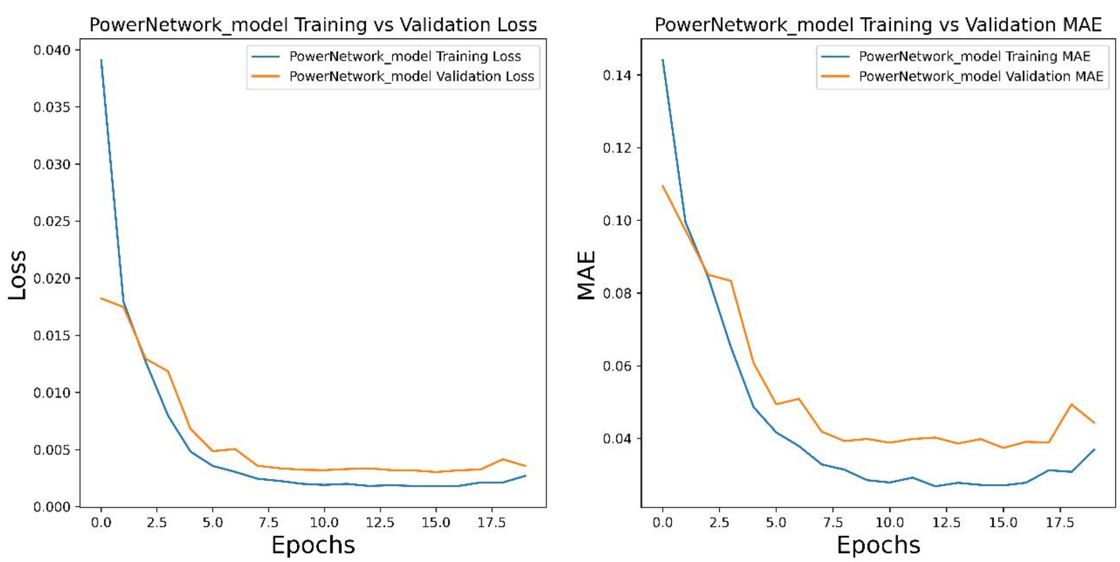
Chapter 4 Experimental Results and Analysis

This project experiment involves the integration of three base models—PowerNetwork, WindFormer Network, and attention mechanism to a robust Attention-PowerWiNet model for the prediction the wind power. The objective is to analyse the performance of four models under different datasets and identify the Attention-PowerWiNet that is expected to deliver the best predictive performance. To be fair in each experiment, all models used MSE as a loss function, and the epochs of training was 20 and the batches size were 64.

4.1 Wind Power Prediction Result-Spring season for dataset 1

4.1.1 PowerNetwork

The assessment focuses on the PowerNetwork model's predictive capabilities over the dataset1(spring). In (a), the training and validation loss decrease rapidly and then stabilize. In (b), the training and validation loss decrease rapidly and then stabilize.



(a) (b)

Figure 15. (a) and (b) present the PowerNetwork train loss and Mae on the first (spring) prediction

In this study, Figure 16 demonstrates the performance of the PowerNetwork in wind power prediction during the spring season. The comparison between predicted and actual values yields an R^2 of 0.80, indicating that the model effectively captures wind power variation trends. This result confirms the high accuracy of the PowerNetwork for spring wind power prediction.

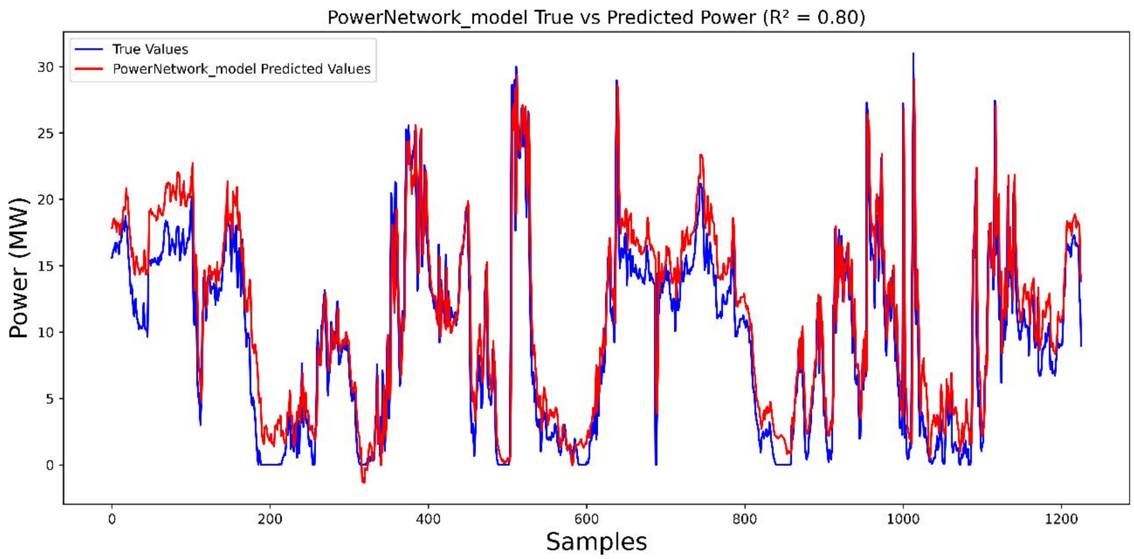
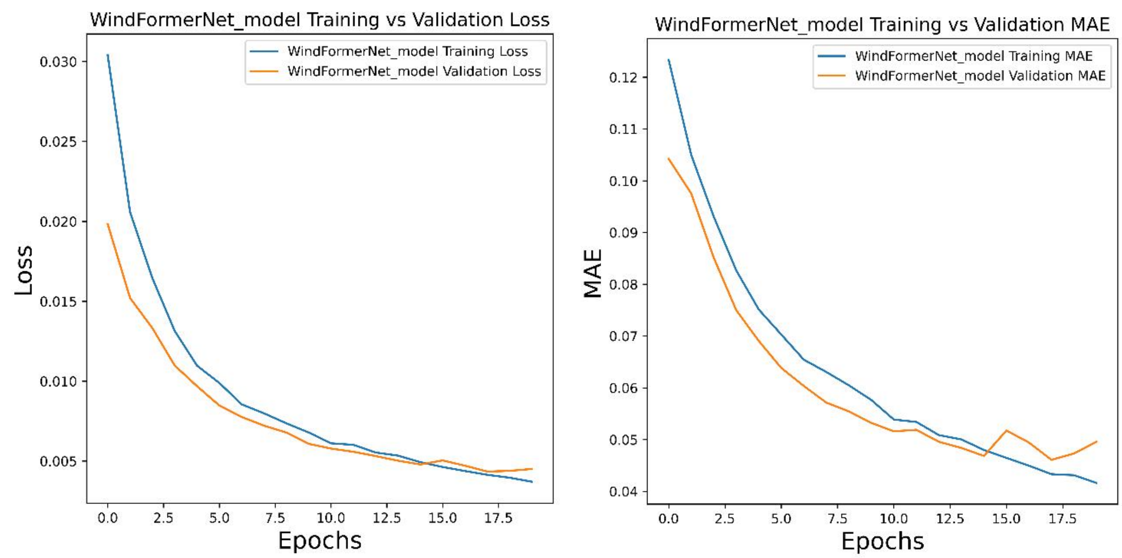


Figure 16. Comparison curve between predicted wind energy value and true value of PowerNetwork

4.1.2 WindFormer Network

The assessment focuses on the WindFormer Network predictive capabilities over the dataset1(spring). In (c), both training and validation loss decrease sharply at the beginning and then level off. In (d), the training and validation MAE follow a similar pattern, with a steep initial drop and later stabilization.



(c) (d)

Figure 17. (c) and (d) present the WindFormer Network train loss and Mae on the first (spring) prediction

Figure 18 shows the comparison between the predicted wind power values from the WindFormerNet model and the actual values. The predicted curve (in red) follows the true values (in blue) not particularly close, indicating that the model can’t effectively capture the overall trend and fluctuations of wind power output.

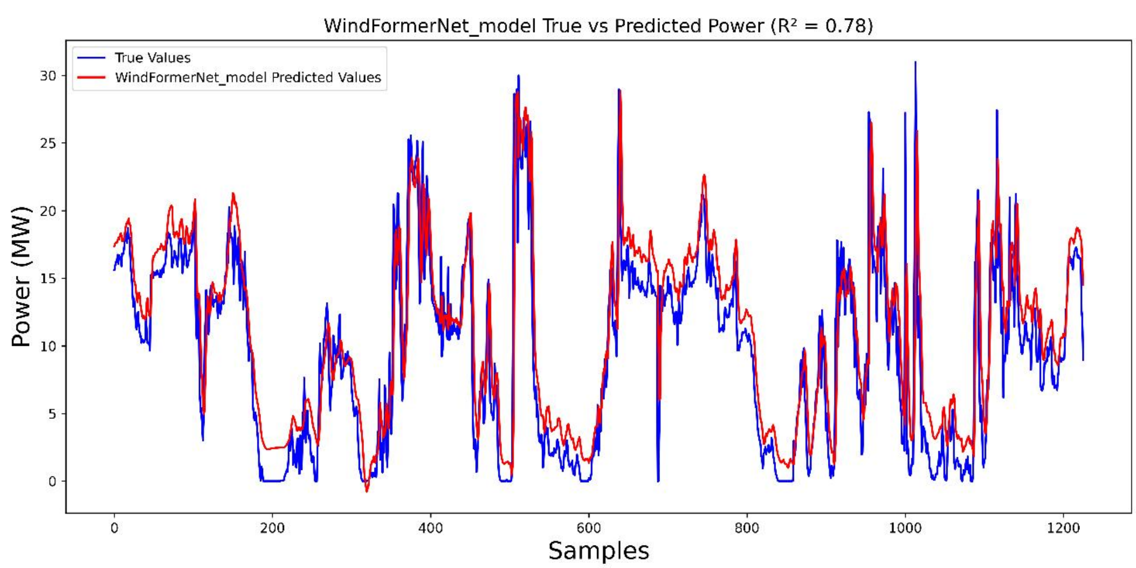
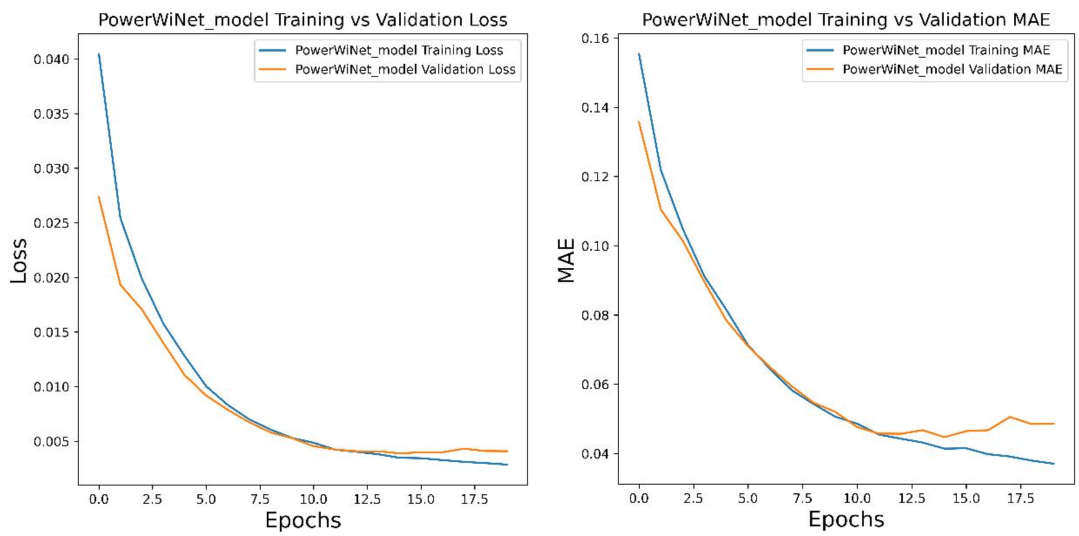


Figure 18. Comparison curve between predicted wind energy value and true value of WindFormer Network

4.1.3 PowerWiNet

The assessment focuses on the PowerWiNet predictive capabilities over the dataset1(spring). In (e), the training and validation loss decrease sharply and then stabilize, indicating effective learning. In (f), the MAE also drops significantly and then levels off.



(e) (f)

Figure 19. (e) and (f) present the PowerWiNet train loss and Mae on the first (spring) prediction

Figure 20 shows the PowerWiNet predictions against actual wind power values during the spring season. The coefficient of determination (R^2) is 0.80, suggesting that the model explains 80% of the variance in the true power output. This indicates good prediction accuracy, though slight deviations appear at some peaks.

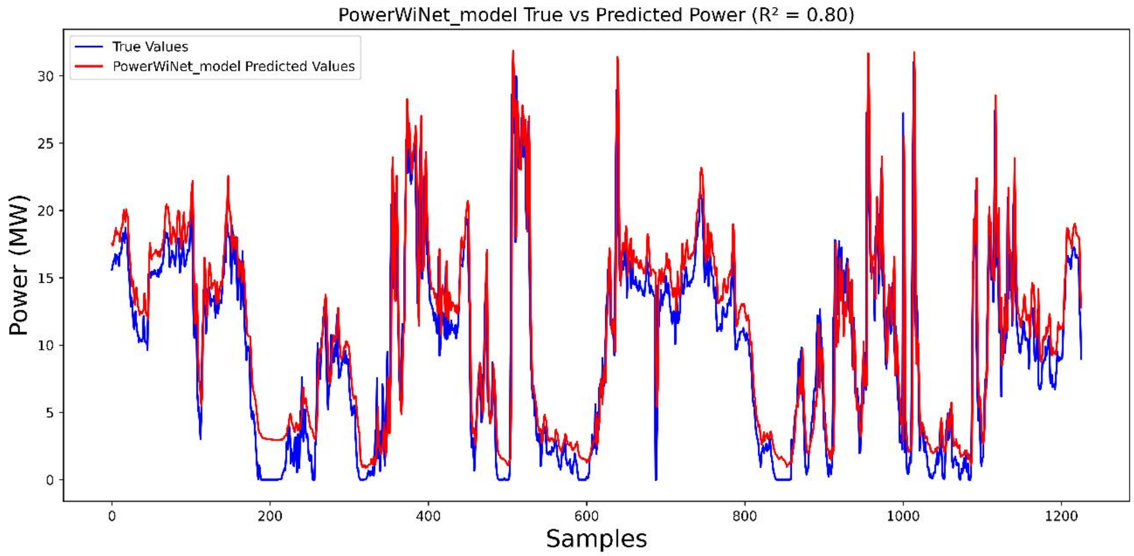
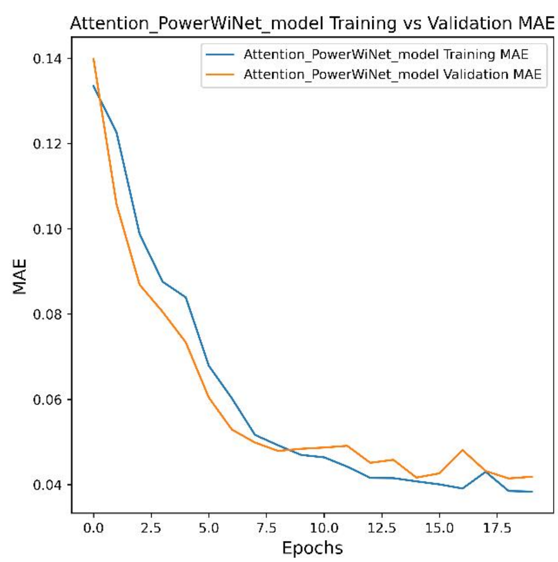


Figure 20. Comparison curve between predicted wind energy value and true value of PowerWiNet

4.1.4 Attention-PowerWiNet

The assessment focuses on the Attention-PowerWiNet predictive capabilities over the dataset1(spring). In (g), the loss continues to decrease steadily with both training and validation loss converging closely. In (h), the Mean Absolute Error (MAE) also shows a consistent decrease, with the training MAE being lower than the validation MAE, indicating the model is generalizing well.

(g)



(h)

Figure 21. (g) and (h) present the Attention-PowerWiNet train loss and Mae on the first (spring) prediction

Figure 22 shows the Attention-PowerWiNet predictions against actual wind power values during the spring season, with an R^2 of 0.84. The blue curve represents the true power values, while the red curve indicates the model's predictions. The predicted curve (in red) follows the true values (in blue) closely, indicating that the model can effectively capture the overall trend and fluctuations of wind power output.

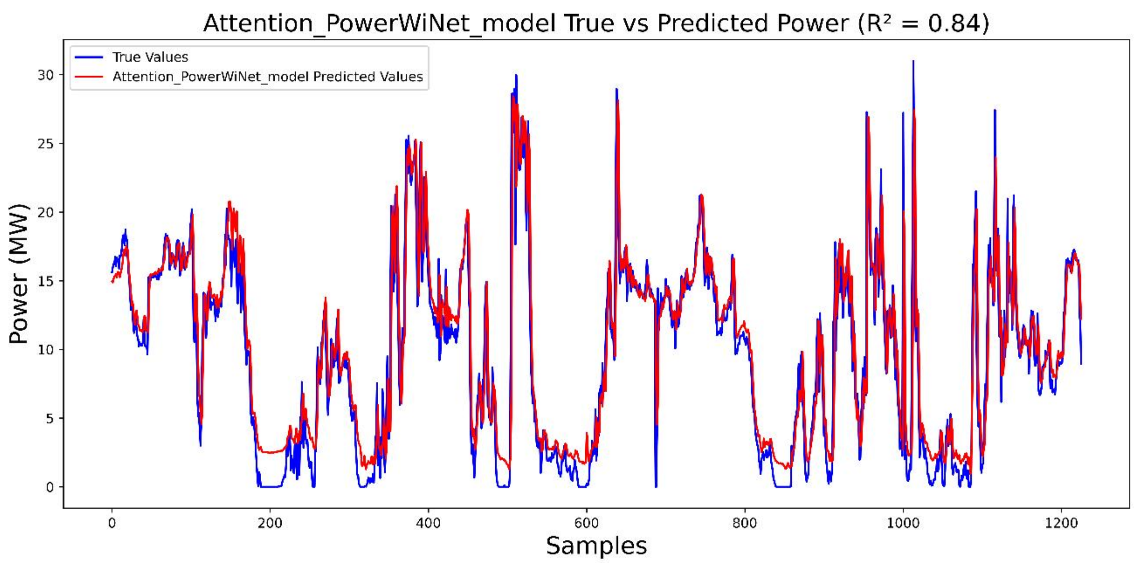


Figure 22. Comparison curve between predicted wind energy value and true value of Attention-PowerWiNet

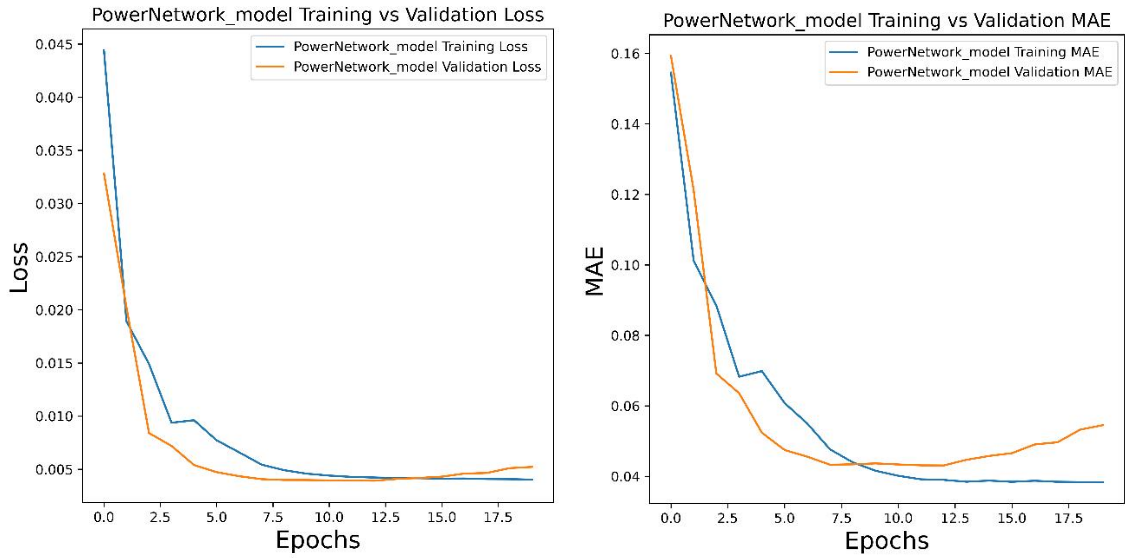
4.1.5 Comparison of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet (spring)

In general, among PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet, Attention-PowerWiNet performs the best, with an R squared of 0.84. In contrast, the WindFormer Network performed the worst with an R squared of 0.78.

4.2 Wind Power Prediction Result-Summer season for dataset 1

4.2.1 PowerNetwork

The assessment focuses on the PowerNetwork predictive capabilities over the dataset1(summer). In (i), both training and validation loss decrease significantly and then stabilize. In (j), the training and validation MAE also drop sharply and then level off.



(i) (j)

Figure 23. (i) and (j) present the PowerNetwork train loss and Mae on the first (summer) prediction

Figure 24 shows the PowerNetwork model's predictions against actual wind power values during the summer season, with an R^2 of 0.89. The blue curve represents the true power values, while the red curve indicates the model’s predictions. The close alignment between the two curves indicates that the PowerNetwork model demonstrates a strong ability to track the temporal dynamics and fluctuations of the actual wind power.

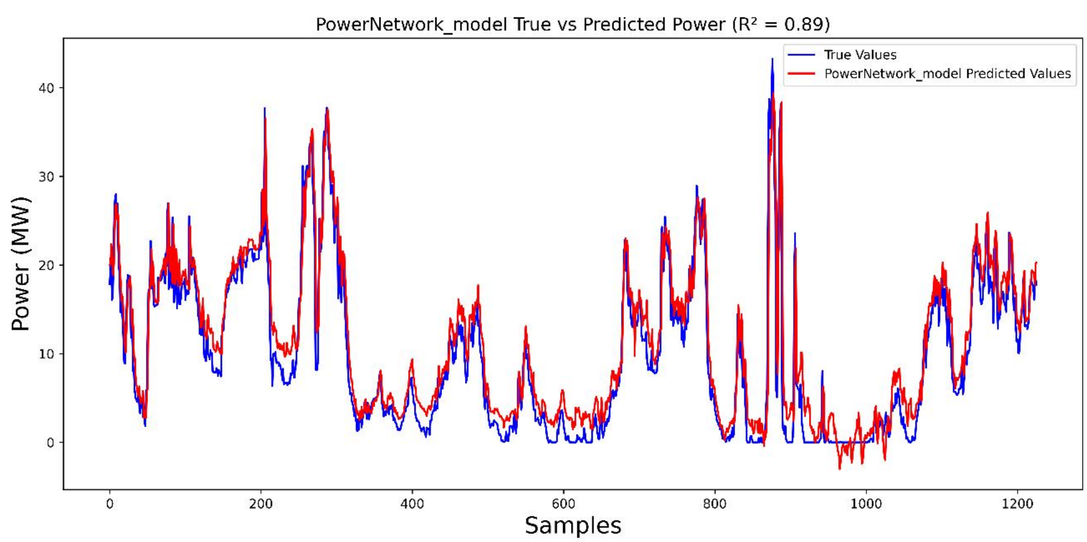
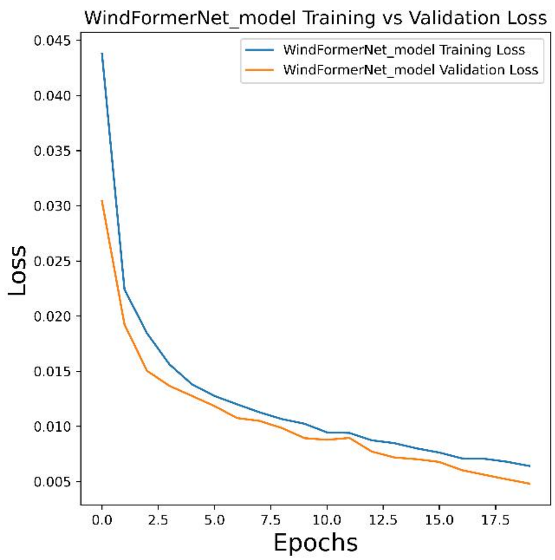


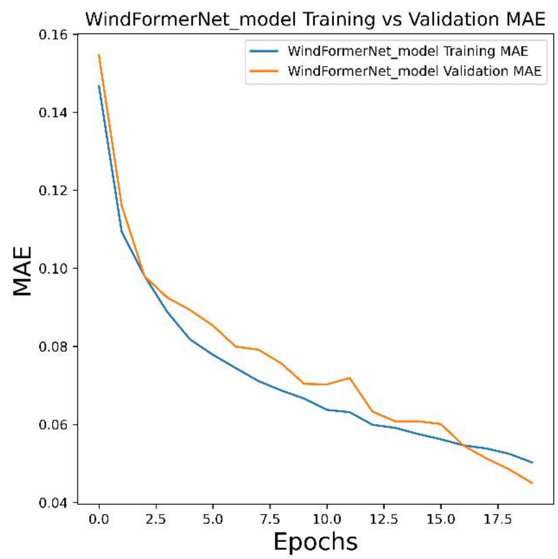
Figure 24. Comparison curve between predicted wind energy value and true value of PowerNetwork

4.2.2 WindFormer Network

The assessment focuses on the WindFormer Network predictive capabilities over the dataset1(summer). Both graphs show a steep initial decline, followed by a plateau, suggesting the model learns quickly and maintains a stable performance.



(k)



(l)

Figure 25. (k) and (l) present the WindFormer Network train loss and Mae on the first (summer) prediction

Figure 26 shows the WindFormer Network model's predictions against actual wind power values during the summer season. The blue curve represents the true power values, while the red curve indicates the model’s predictions. The coefficient of determination (R^2) is 0.88, suggesting that the model explains 88% of the variance in the true power output. Although some deviations are shown, particularly during peak values or rapid transitions, the model maintains a generally reliable performance in predicting wind energy output.

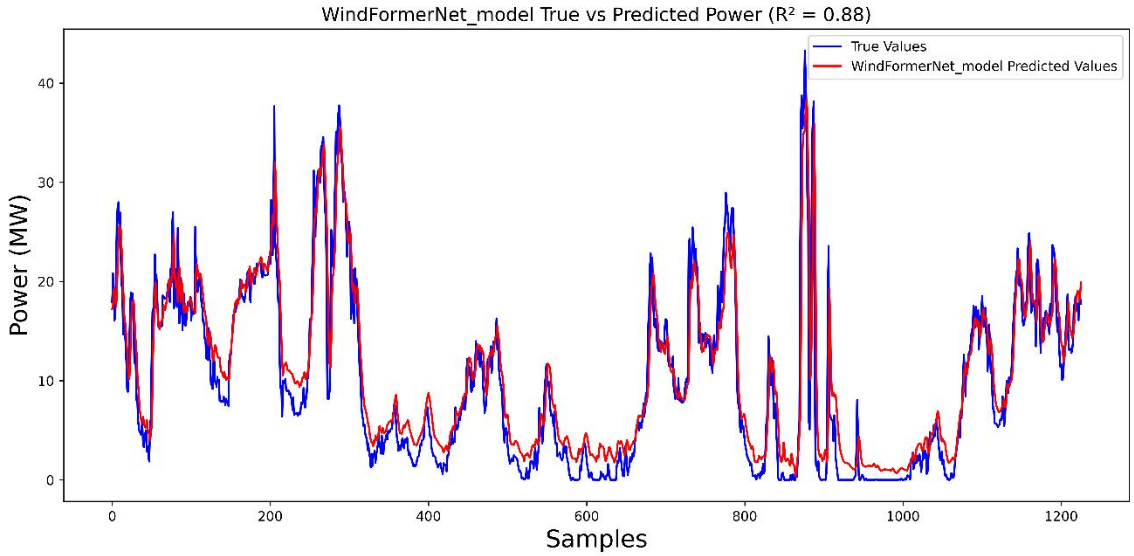
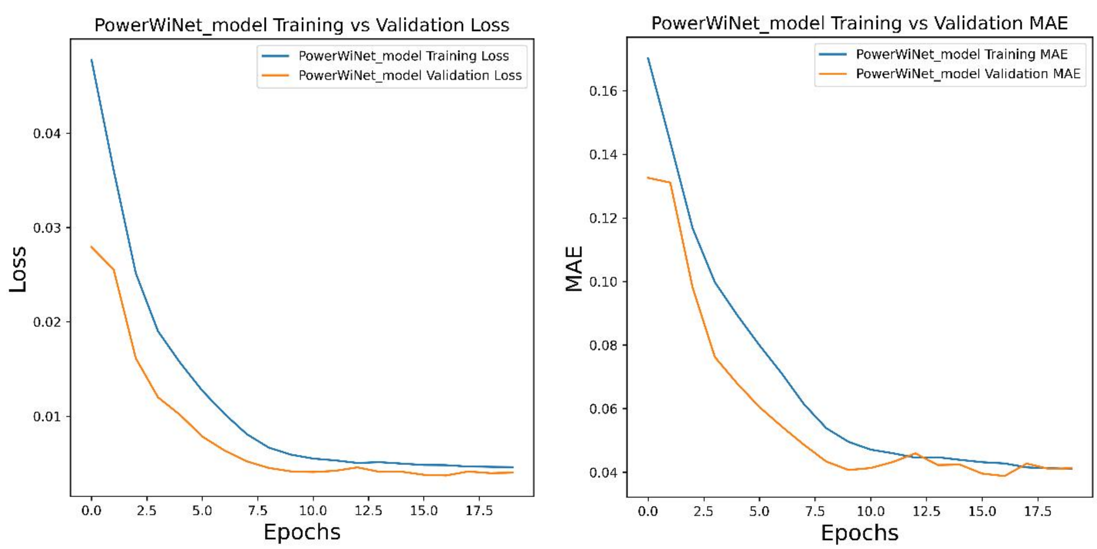


Figure 26. Comparison curve between predicted wind energy value and true value of WindFormer Network

4.2.3 PowerWiNet

The assessment focuses on the PowerWiNet predictive capabilities over the dataset1(summer). In (m), both training and validation loss decrease rapidly and then level off. In (n), the training and validation MAE also drop sharply and stabilize.



(m) (n)

Figure 27. (m) and (n) present the PowerWiNet train loss and Mae on the first (summer) prediction

Figure 28 shows the PowerWiNet predictions against actual wind power values during the summer season. The blue curve represents the true power values, while the red curve indicates the model's predictions. The R^2 value of 0.89 indicates a high level of prediction accuracy and a strong correlation between the predicted and actual power outputs.

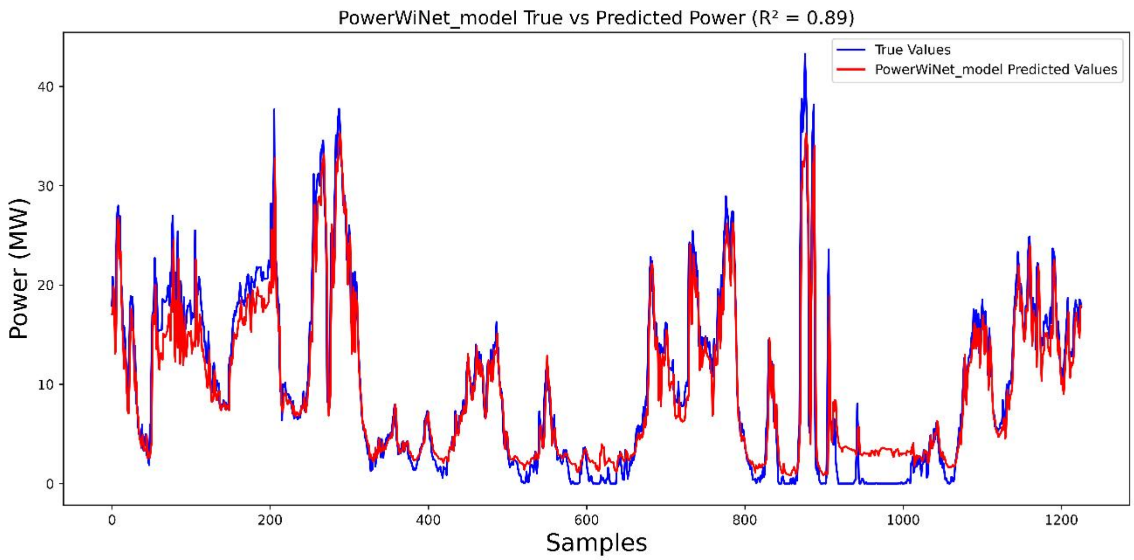
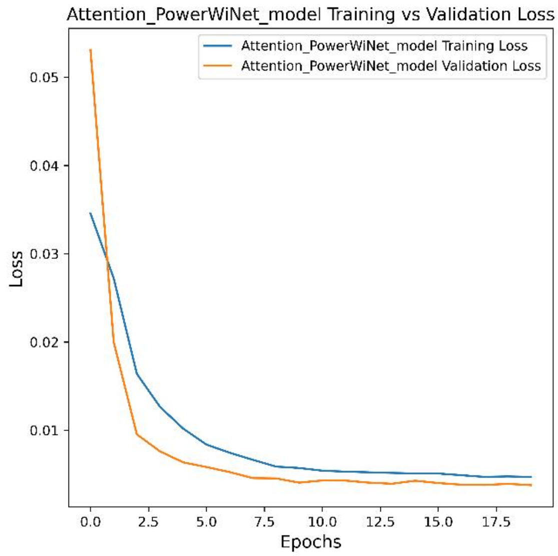
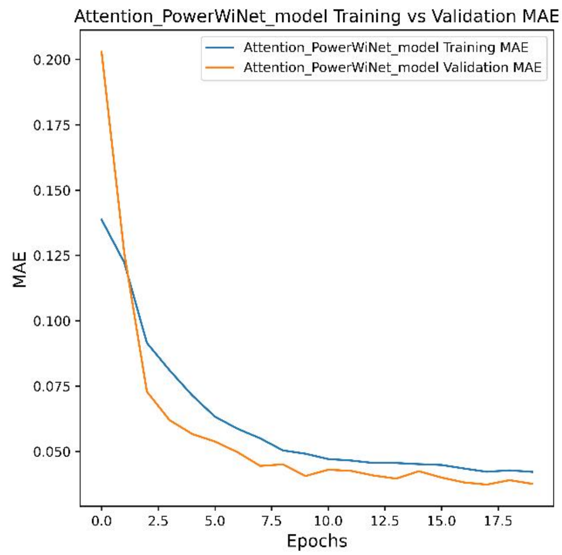


Figure 28. Comparison curve between predicted wind energy value and true value of PowerWiNet

4.2.4 Attention-PowerWiNet

The assessment focuses on the Attention-PowerWiNet predictive capabilities over the dataset1(summer). The loss curves (o) demonstrate that both training and validation losses decrease steadily, indicating excellent model convergence. The MAE curves (p) show synchronized reduction to approximately 0.04, confirming the model's reliability for summer wind power prediction.





(o) (p)

Figure 29. (o) and (p) present the Attention-PowerWiNet train loss and Mae on the first (summer) prediction

Figure 30 shows the Attention-PowerWiNet predictions against actual wind power values during the summer season, with an R^2 of 0.90. The blue curve represents the true power values, while the red curve indicates the model's predictions. Compared to PowerNetwork, WindFormerNet, and PowerWiNet shows better tracking performance.

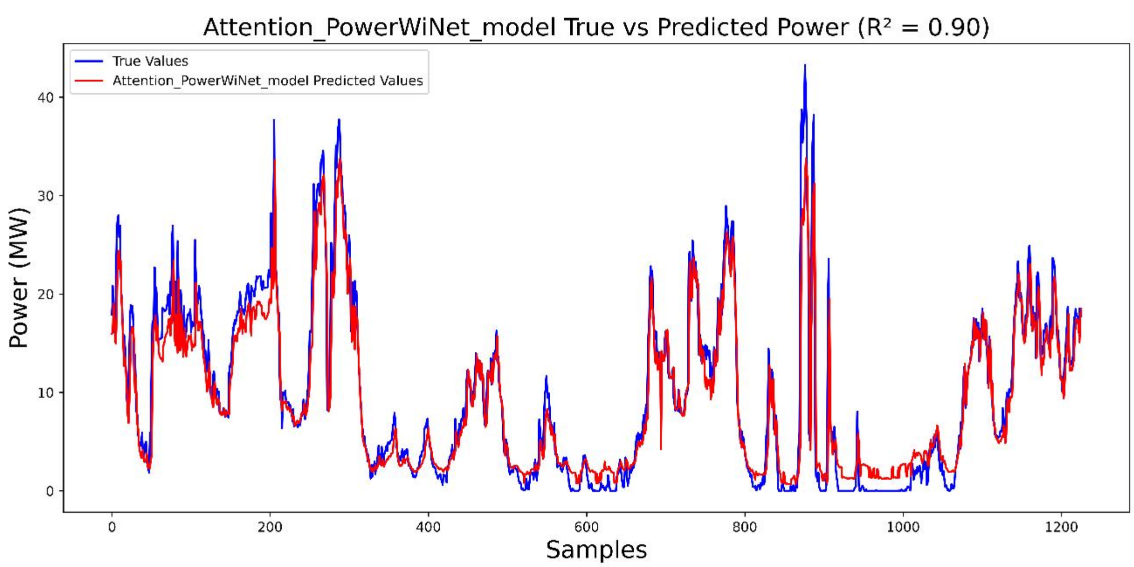


Figure 30. Comparison curve between predicted wind energy value and true value of Attention-PowerWiNet

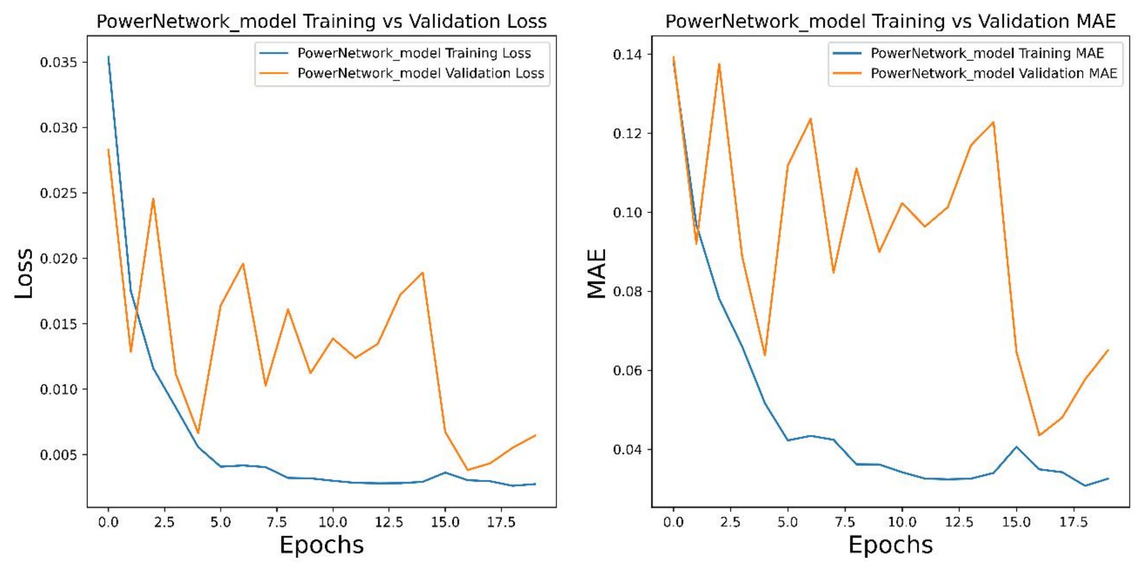
4.2.5 Comparison of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet (summer)

Through comparative experiments, this study demonstrates that among the four models (PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet), the Attention-PowerWiNet achieves the best performance with an R-squared value of 0.90. Moreover, its loss and MAE curves during training exhibit the most stable convergence behaviour, indicating superior model robustness and predictive capability.

4.3 Wind Power Prediction Result-Autumn season for dataset 1

4.3.1 PowerNetwork

The assessment focuses on the PowerNetwork predictive capabilities over the dataset1(autumn). In (q), shows a clear trend of decreasing training loss and fluctuating validation loss, which suggests that the model is learning from the training data but may be experiencing some overfitting as the validation loss does not consistently decrease. In (r), the training MAE drops significantly and levels off, whereas the validation MAE decreases and then exhibits variability.



(q) (r)

Figure 31. (q) and (r) present the PowerNetwork train loss and Mae on the first (autumn) prediction

Figure 32 shows the PowerNetwork predictions against actual wind power values during the autumn season. The blue curve represents the true power values, while the red curve indicates the model's predictions. The R^2 value of 0.79 indicates a moderate level of prediction accuracy, suggesting that while the model can capture general patterns in the data.

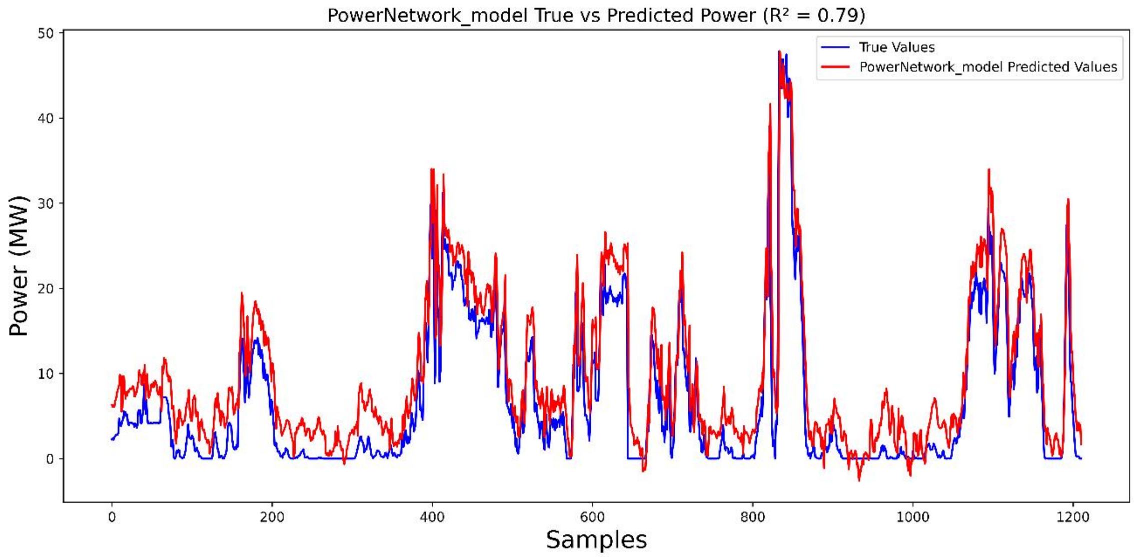
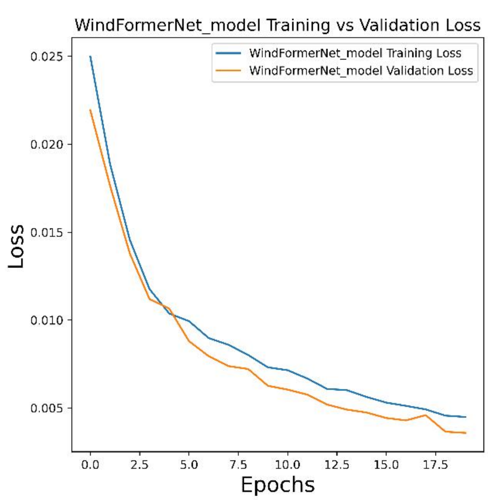


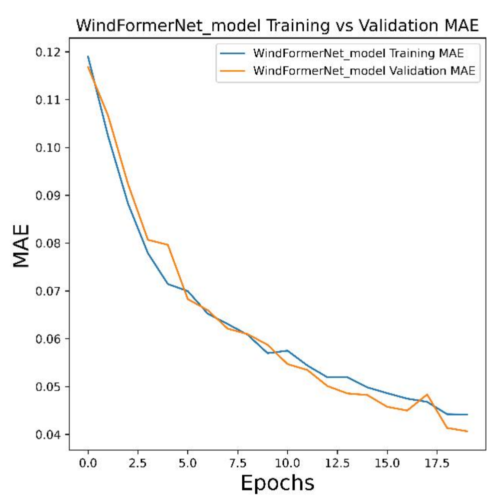
Figure 32. Comparison curve between predicted wind energy value and true value of PowerNetwork

4.3.2 WindFormer Network

The assessment focuses on the WindFormer Network predictive capabilities over the dataset1(autumn). In (s), the training loss decreases sharply and then stabilizes, while the validation loss shows a general downward trend. In (t), the training and validation MAE drops significantly and then levels off.



(s)



(t)

Figure 33. (s) and (t) present the WindFormer Network train loss and Mae on the first (autumn) prediction

Figure 34 shows the WindFormer Network predictions against actual wind power values during the autumn season. The blue curve represents the true power values, while the red curve indicates the model's predictions. The R^2 value of 0.90 indicates a high level of prediction accuracy and a strong correlation between the predicted and actual power outputs.

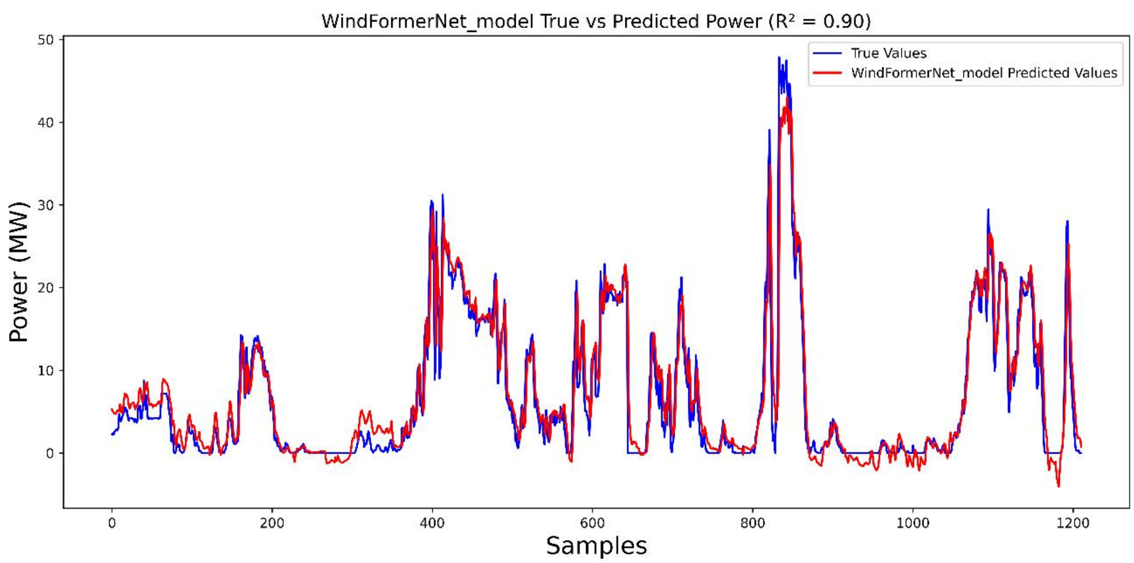
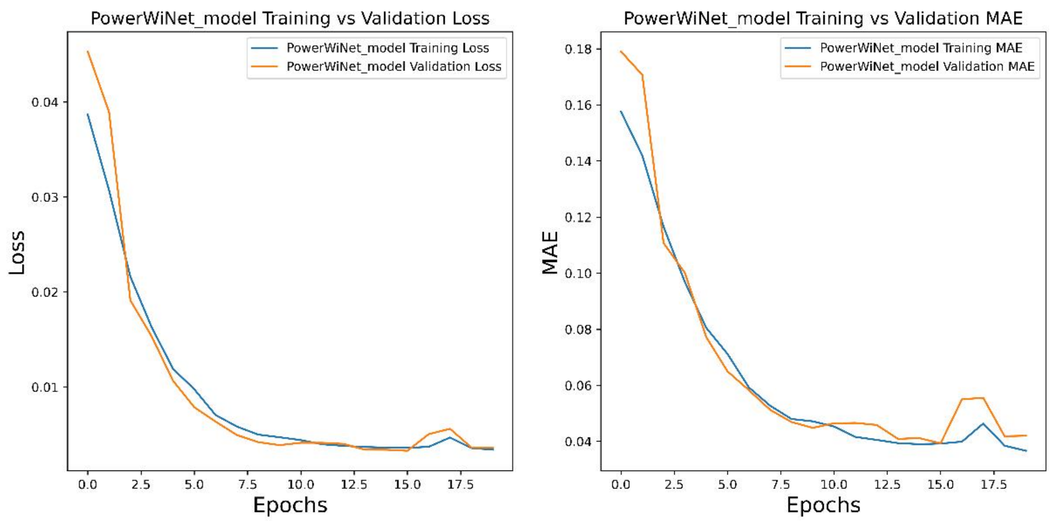


Figure 34. Comparison curve between predicted wind energy value and true value of WindFormer Network

4.3.3 PowerWiNet

The assessment focuses on the PowerWiNet predictive capabilities over the dataset1(autumn). In (u), both training and validation loss decrease sharply and then stabilize, indicating efficient learning. In (v), the training and validation MAE also drop significantly and then level off, confirming the model's strong predictive accuracy.



(u) (v)

Figure 35. (u) and (v) present the PowerWiNet train loss and Mae on the first (autumn) prediction

Figure 36 shows the PowerWiNet predictions against actual wind power values during the autumn season. The blue curve represents the true power values, while the red curve indicates the model's predictions. The R^2 value of 0.90 indicates a high level of prediction accuracy and a strong correlation between the predicted and actual power outputs.

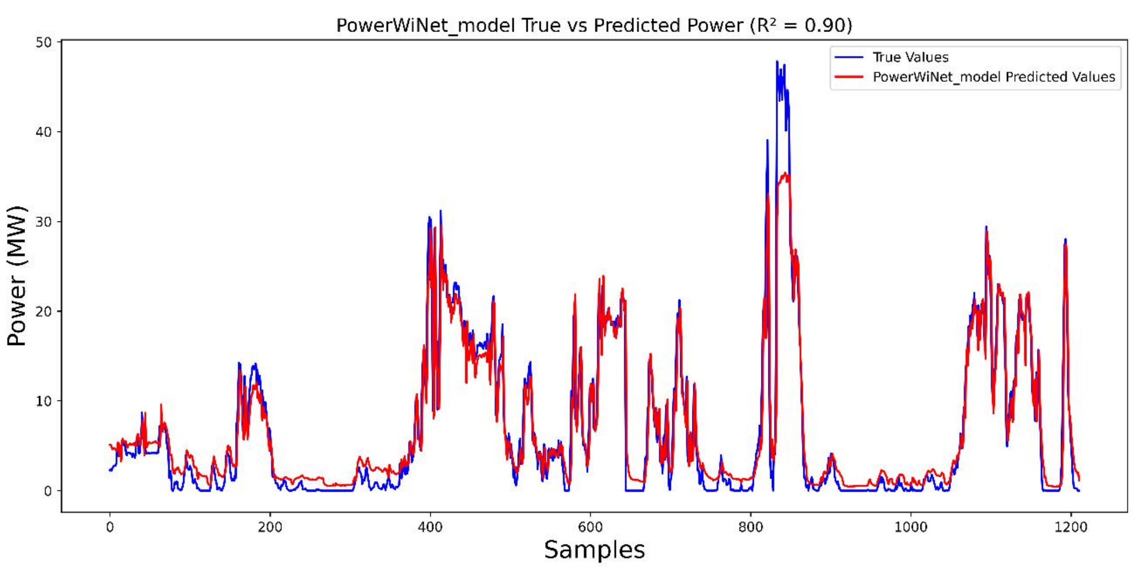
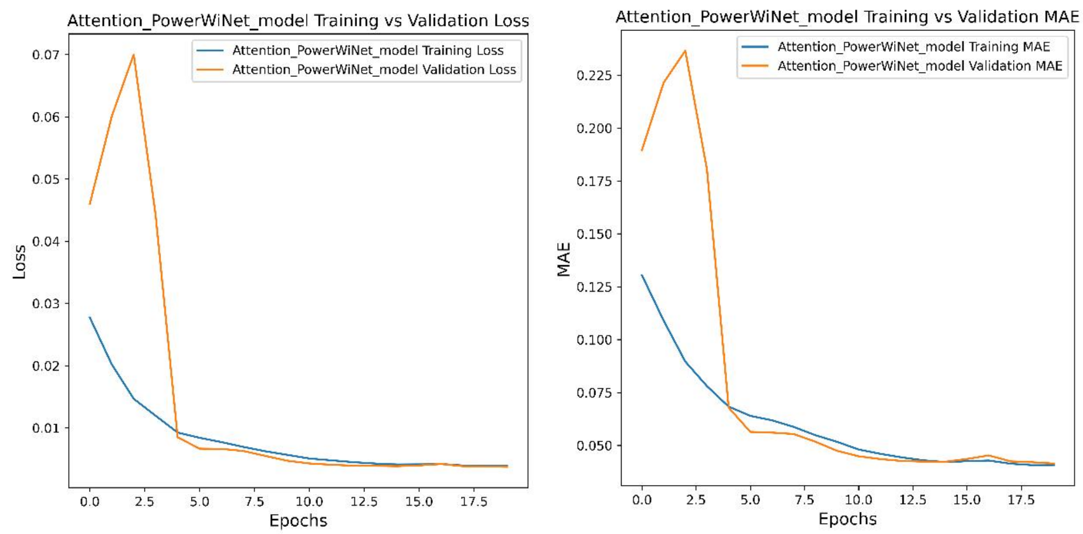


Figure 36. Comparison curve between predicted wind energy value and true value of PowerWiNet

4.3.4 Attention-PowerWiNet

The assessment focuses on the Attention-PowerWiNet predictive capabilities over the dataset1(autumn). The loss curves (w) demonstrate that both training and validation losses decrease steadily with epochs, eventually converging to approximately 0.003. Similarly, the MAE curves (x) show a consistent downward trend, with training and validation MAE stabilizing around 0.025.



(w) (x)

Figure 37. (w) and (x) present the Attention-PowerWiNet train loss and Mae on the first (autumn) prediction

Figure 38 shows the Attention-PowerWiNet predictions against actual wind power values during the autumn season. The blue curve represents the true power values, while the red curve indicates the model’s predictions. The coefficient of determination (R^2) is 0.90, suggesting that the model explains 90% of the variance in the true power output. The model maintains a generally reliable performance in predicting wind energy output.

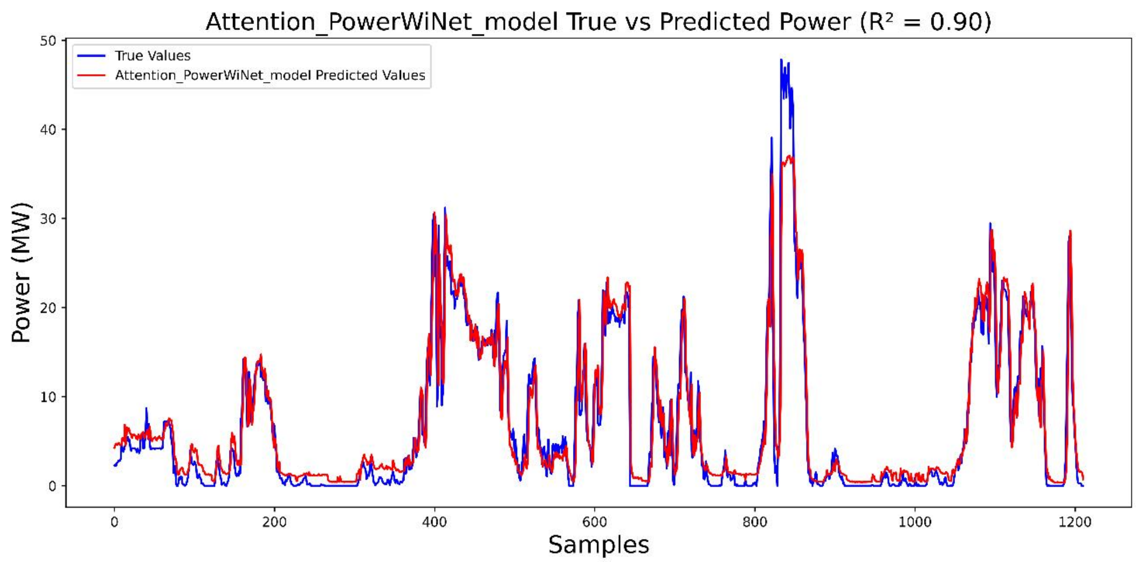


Figure 38. Comparison curve between predicted wind energy value and true value of Attention-PowerWiNet

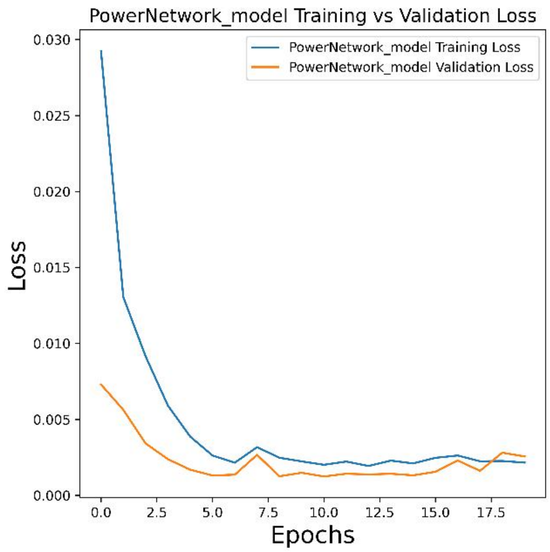
4.3.5 Comparison of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet (autumn)

In general, Comparison of WindFormer Network, PowerWiNet, and Attention-PowerWiNet all perform well in the autumn dataset. In contrast, PowerNetwork performed the worst because the square of R was 0.79.

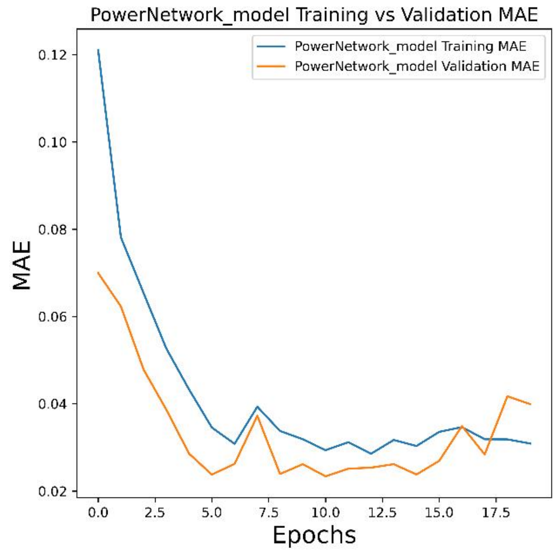
4.4 Wind Power Prediction Result-Winter season for dataset 1

4.4.1 PowerNetwork

The assessment focuses on the PowerNetwork predictive capabilities over the dataset1(winter). The left plot shows a decreasing trend in the loss function over epochs, indicating successful model convergence. The right plot presents the MAE, where both training and validation curves generally follow a similar pattern. Notably, a temporary fluctuation is observed around the 7.5th epoch.



(y)



(z)

Figure 39. (y) and (z) present the PowerNetwork train loss and Mae on the first (winter) prediction

Figure 40 shows the PowerNetwork predictions against actual wind power values during the winter season, with an R^2 of 0.88. The predicted curve (in red) follows the true values (in blue) closely, indicating that the model can effectively capture the overall trend and fluctuations of wind power output.

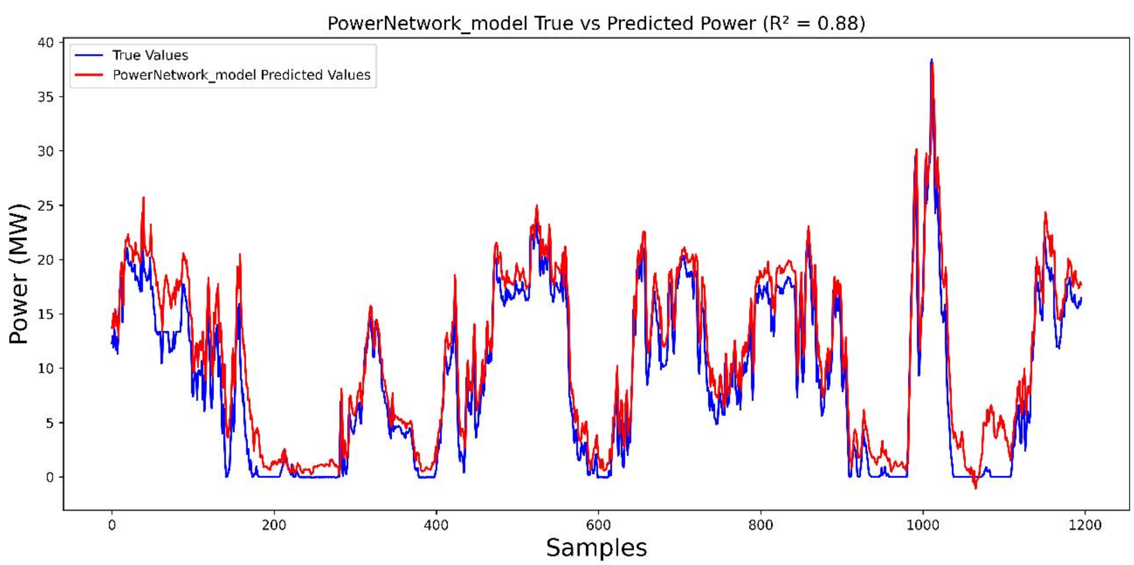
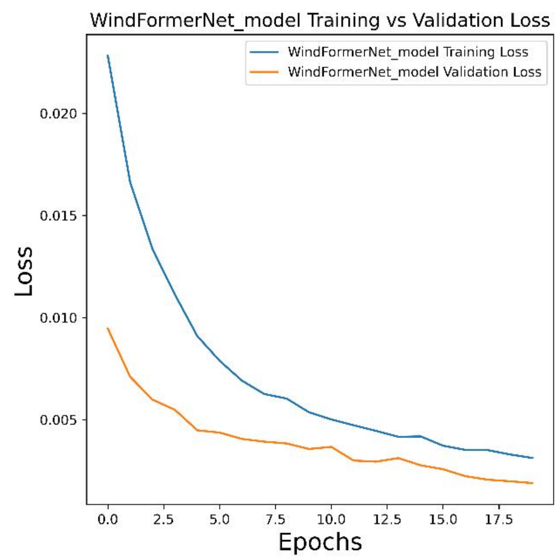


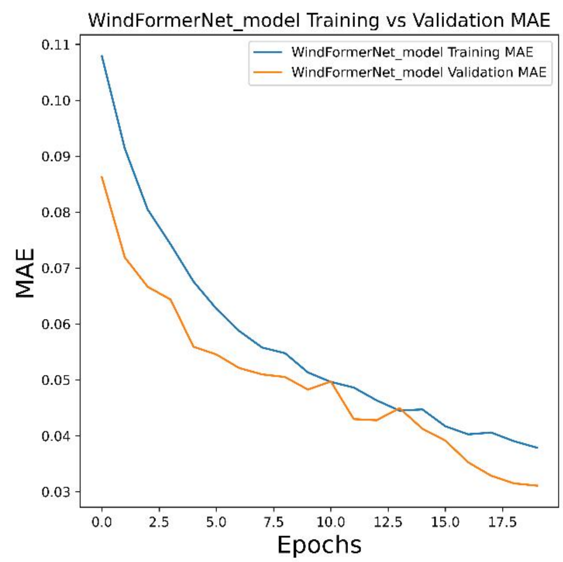
Figure 40. Comparison curve between predicted wind energy value and true value of PowerNetwork

4.4.2 WindFormer Network

The assessment focuses on the WindFormer Network predictive capabilities over the dataset1(winter). In (a1), both training and validation loss decrease significantly and then level off, indicating effective learning and a good fit. In (b1), the training and validation MAE follow a similar pattern, with a sharp initial drop and later stabilization, confirming the model's strong predictive accuracy.



(a1)



(b1)

Figure 41. (a1) and (b1) present the WindFormer Network train loss and Mae on the first (winter) prediction

Figure 42 shows the WindFormer Network predictions against actual wind power values during the winter season, with an R^2 of 0.90. The blue curve represents the true power values, while the red curve indicates the model's predictions. The predicted curve follows the overall trend of the true values, showing a strong correlation between the predicted and actual power outputs.

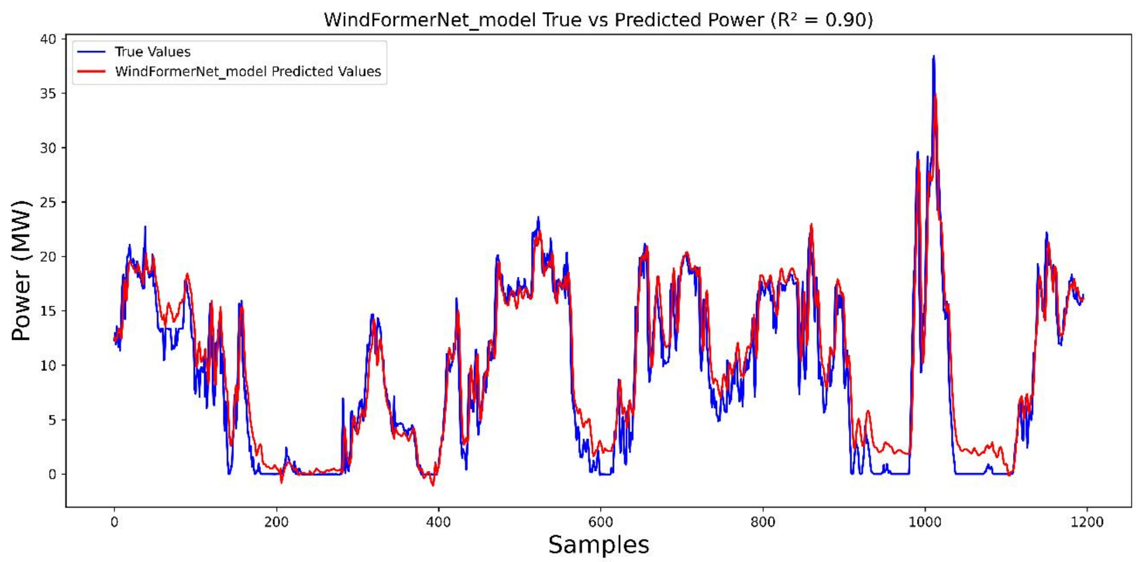
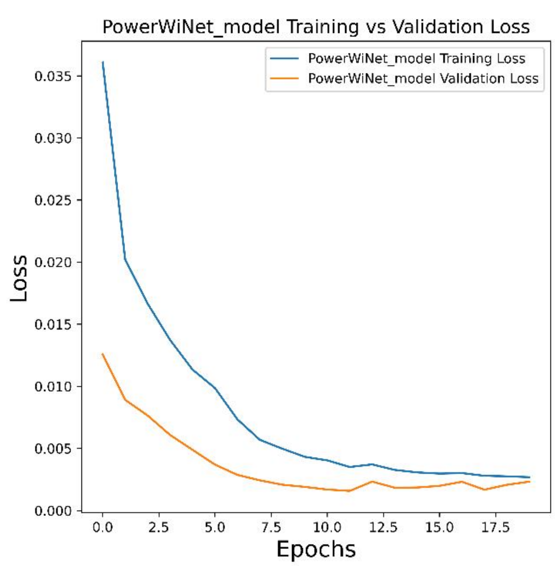


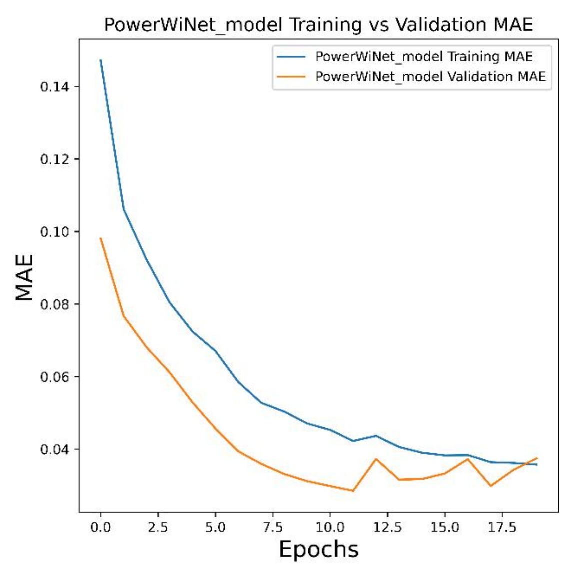
Figure 42. Comparison curve between predicted wind energy value and true value of WindFormer Network

4.4.3 PowerWiNet

The assessment focuses on the PowerWiNet predictive capabilities over the dataset1(winter). In (c1), the training and validation loss decrease rapidly and then stabilize, indicating effective learning. In (d1), the training and validation MAE also drop significantly and then level off, confirming the model's strong predictive accuracy.



(c1)



(d1)

Figure 43. (c1) and (d1) present the PowerWiNet train loss and Mae on the first (winter) prediction

Figure 44 shows the PowerWiNet predictions against actual wind power values during the winter season, with an R^2 of 0.90. The blue curve represents the true power values, while the red curve indicates the model's predictions. The red predicted curve generally reflects the changing trend of the true values shown in blue, show the stable performance of this model.

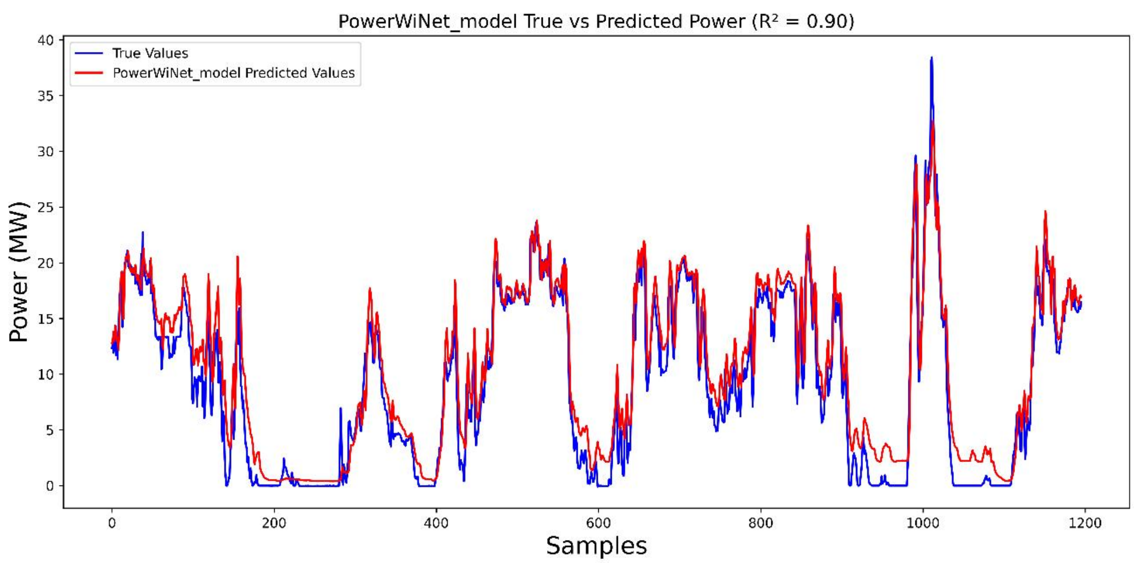
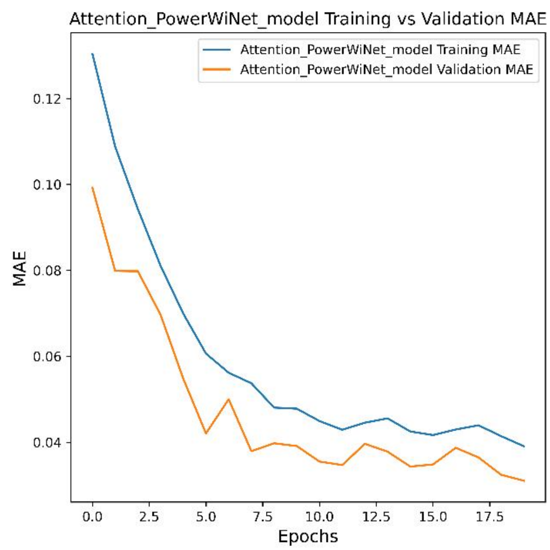


Figure 44. Comparison curve between predicted wind energy value and true value of PowerWiNet

4.4.4 Attention-PowerWiNet

The assessment focuses on the Attention-PowerWiNet predictive capabilities over the dataset1(winter). In the left plot (e1), the loss steadily decreases throughout training, with the validation loss consistently lower than the training loss and stabilizing after around epoch 10. The right plot (f1) shows MAE, which also decreases significantly over time. However, minor fluctuations are observed in the validation MAE around epochs 5 and 15, possibly due to model sensitivity to specific sample patterns.



(e1) (f1)

Figure 45. (e1) and (f1) present the Attention-PowerWiNet train loss and Mae on the first (winter) prediction

Figure 46 shows the Attention-PowerWiNet predictions against actual wind power values during the winter season. The predicted curve (red) shows a strong alignment with the true values (blue) across most sample points, capturing both the overall trend and detailed fluctuations. The high R^2 (0.91) indicates its strong capability in wind power forecasting.

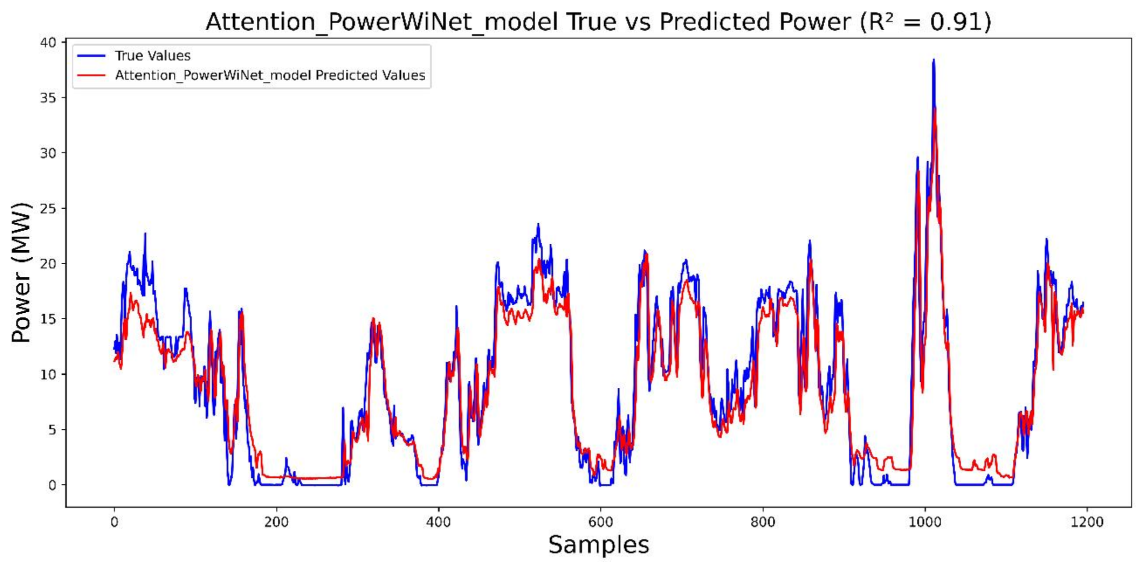


Figure 46. Comparison curve between predicted wind energy value and true value of Attention-PowerWiNet

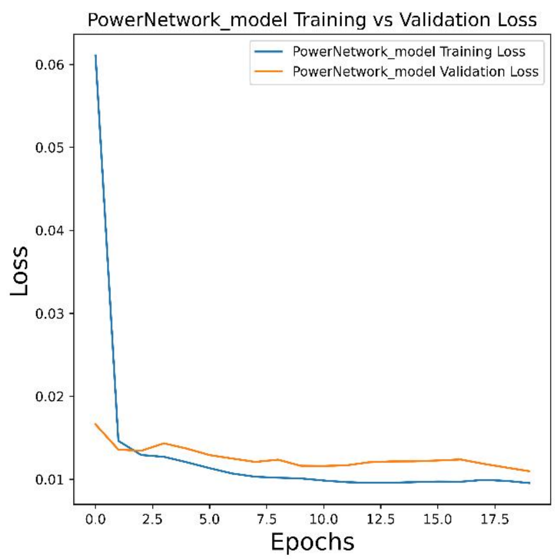
4.4.5 Comparison of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet (winter)

In general, PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet all perform well in the winter dataset, but Attention-PowerWiNet performs better because its r square is the highest, which is 0.91

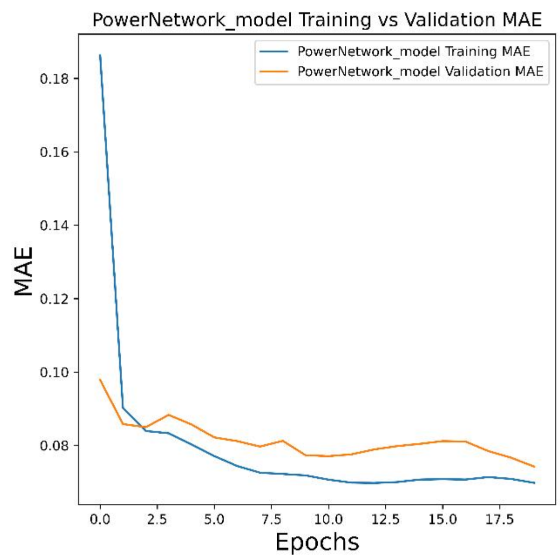
4.5 Wind Power Prediction Result for Dataset2

4.5.1 PowerNetwork

The assessment focuses on the PowerNetwork predictive capabilities over the dataset2. In the left plot (g1), the loss decreases rapidly during the initial epochs and then stabilizes, with the training loss consistently lower than the validation loss, indicating good convergence. The right plot (h1) demonstrates a clear decreasing trend in MAE.



(g1)



(h1)

Figure 47. (e1) and (f1) present the PowerNetwork train loss and Mae on the second prediction

Figure 48 shows the PowerNetwork predictions against actual wind power values of the dataset2, with an R^2 of 0.87. The predicted curve (in red) generally follows the true values (in blue) with good accuracy, effectively capturing the overall trends and most of the variations.

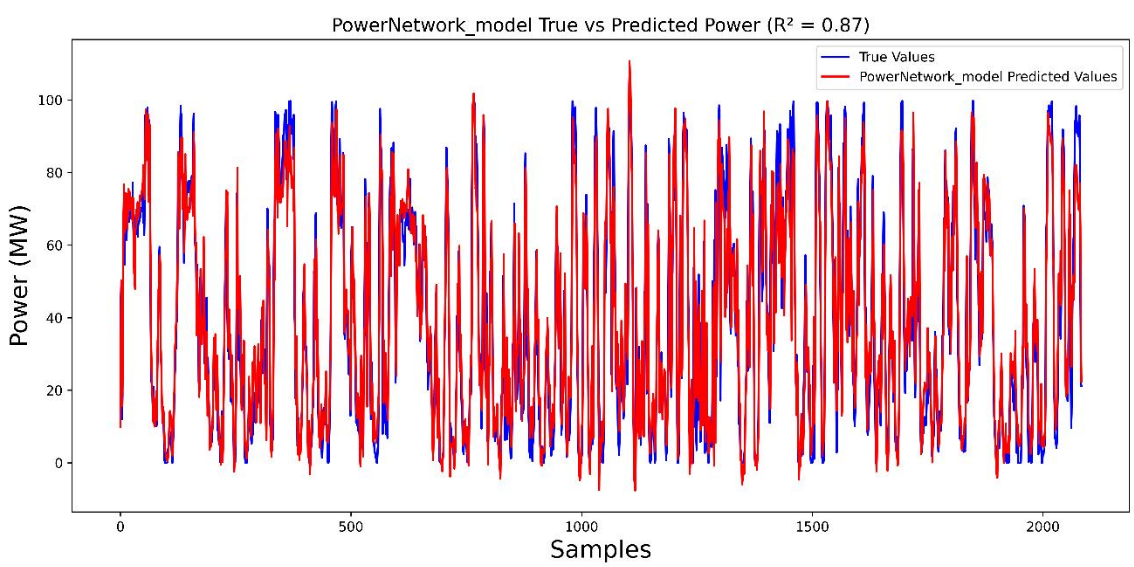
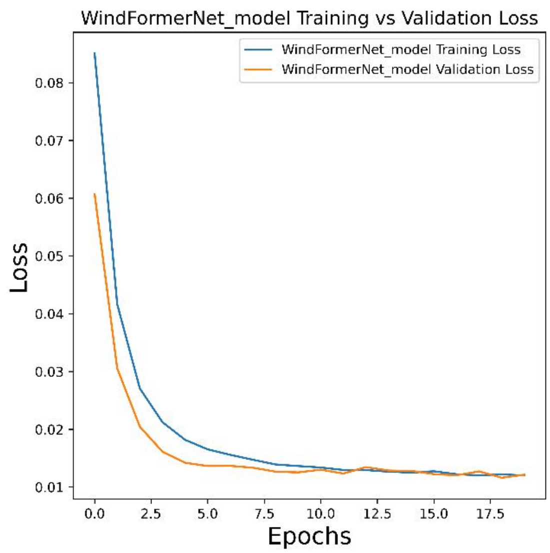


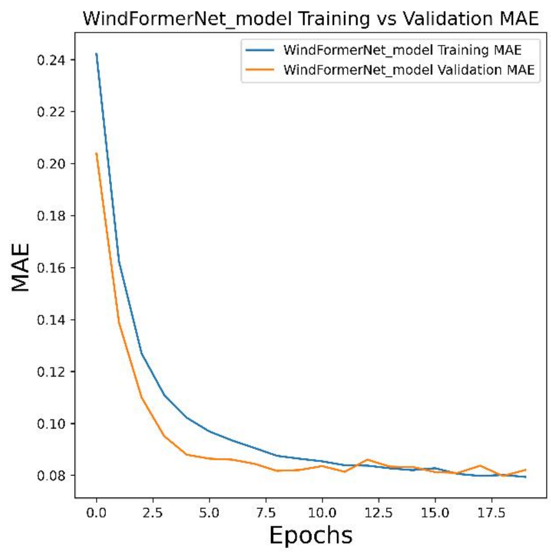
Figure 48. Comparison curve between predicted wind energy value and true value of PowerNetwork

4.5.2 WindFormer Network

The assessment focuses on the WindFormer Network predictive capabilities over the dataset2. In (i1), both the training loss (blue line) and validation loss (orange line) decrease sharply in the initial epochs and then level off. In (j1), like the loss graph, both the training MAE (blue line) and validation MAE (orange line) drop significantly at the beginning and then stabilize.



(i1)



(j1)

Figure 49. (i1) and (j1) present the WindFormer Network train loss and Mae on the second prediction

Figure 50 shows the WindFormer Network predictions against actual wind power values of the dataset2, with an R^2 of 0.87. The predicted curve (in red) generally follows the true values (in blue) with good accuracy, effectively capturing the overall trends and most of the variations.

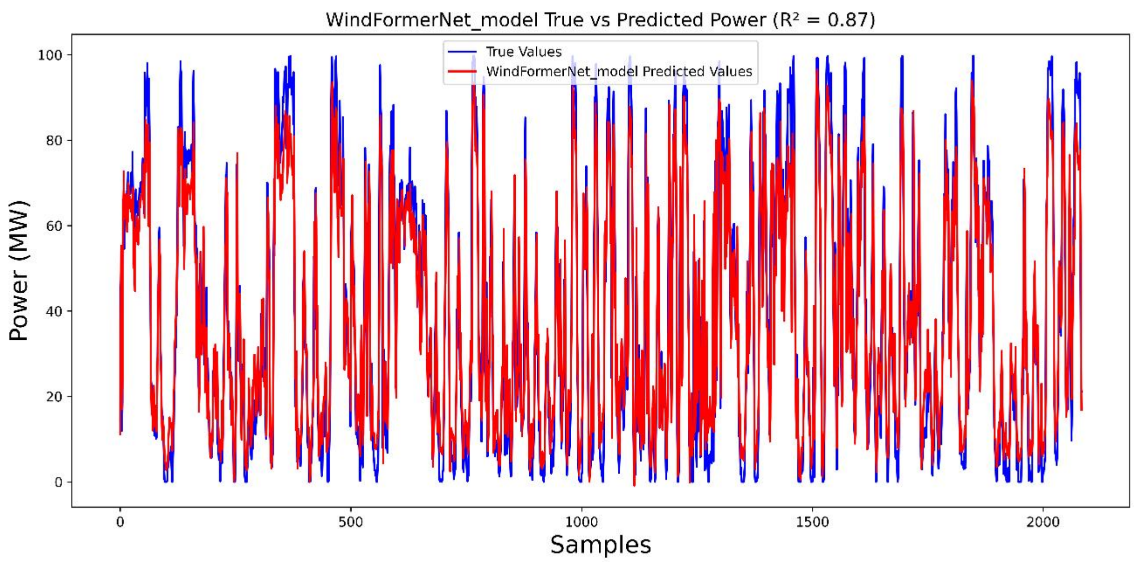
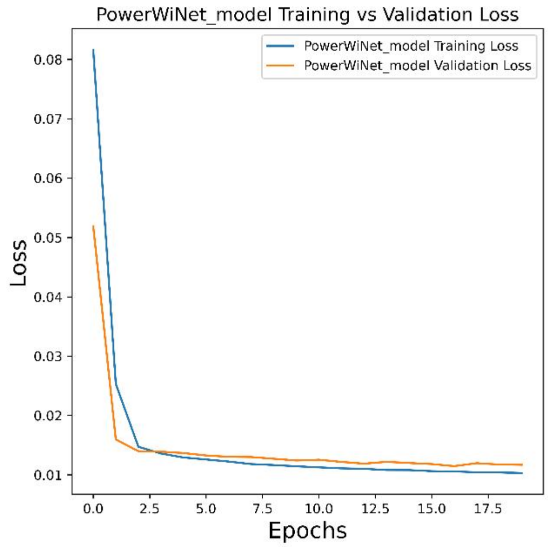


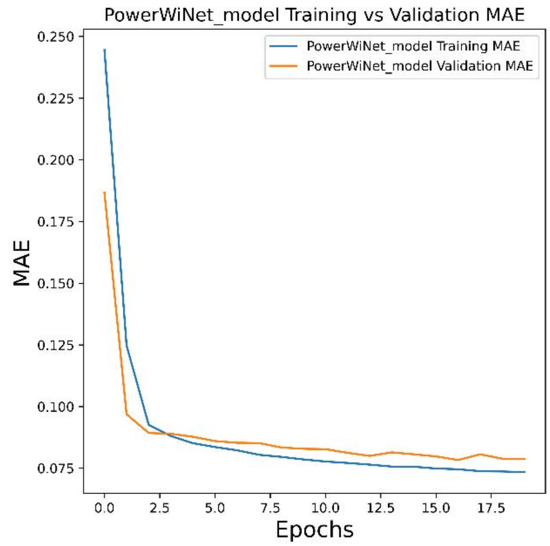
Figure 50. Comparison curve between predicted wind energy value and true value of WindFormer Network

4.5.3 PowerWiNet

The assessment focuses on the PowerWiNet predictive capabilities over the dataset2. In (k1), the training loss (blue line) starts high and drops sharply, then continues to decrease steadily before plateauing. The validation loss (orange line) follows a similar trend but at a slower rate. In (l1), the training MAE (blue line) decreases rapidly at first and then more gradually. The validation MAE (orange line) also decreases and stabilizes.



(k1)



(l1)

Figure 51. (k1) and (l1) present the PowerWiNet train loss and Mae on the second prediction

Figure 52 shows the PowerWiNet predictions against actual wind power values of the dataset2 with an R^2 of 0.87, showing a strong correlation between the predicted (red) and actual (blue) wind power values. Demonstrating the model strong capability in wind power forecasting.

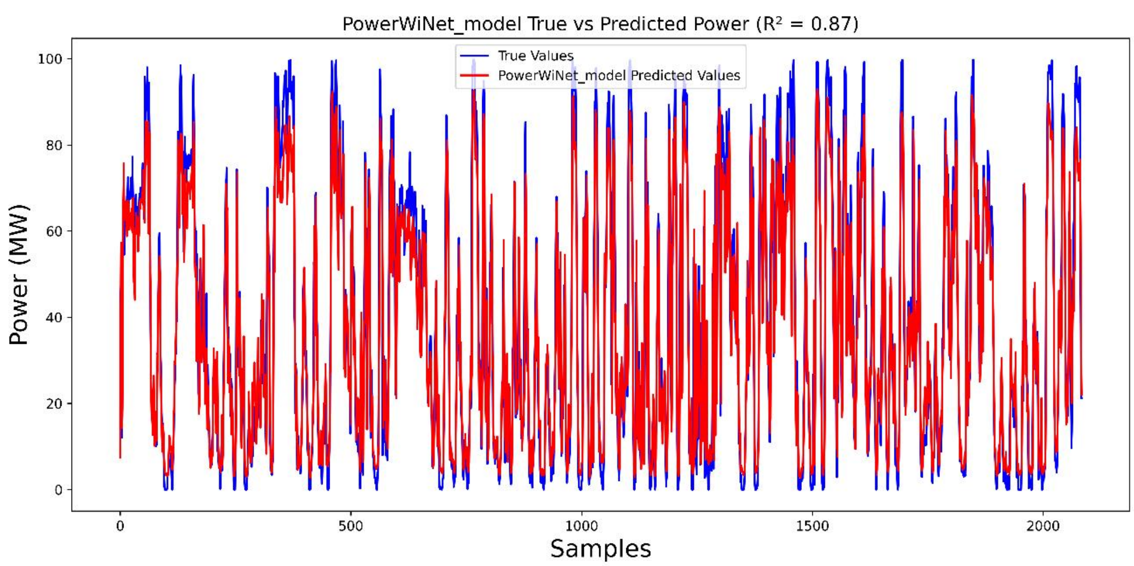
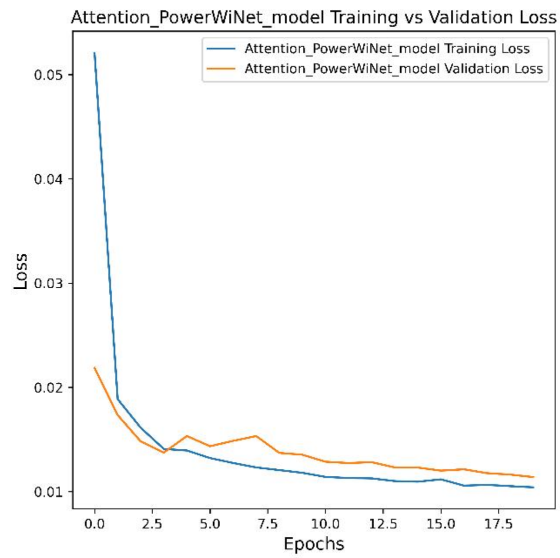


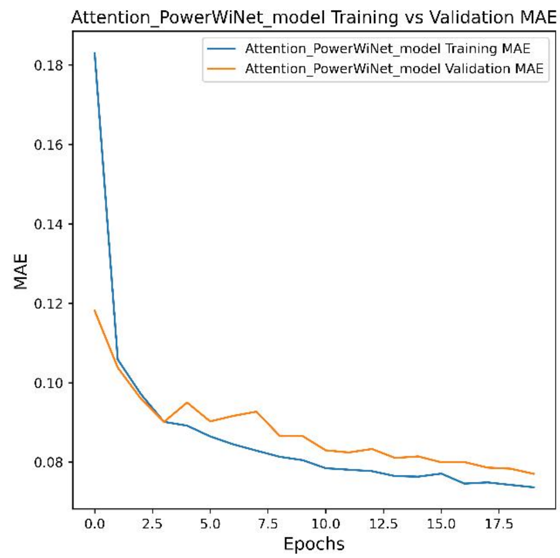
Figure 52. Comparison curve between predicted wind energy value and true value of PowerWiNet

4.5.4 Attention-PowerWiNet

The assessment focuses on the Attention-PowerWiNet predictive capabilities over the dataset2. In (m1), both training and validation loss decrease sharply in the initial epochs and then level off. In (n1), the Mean Absolute Error (MAE) for both training and validation decreases significantly and then stabilizes.



(m1)



(n1)

Figure 53. (m1) and (n1) present the Attention-PowerWiNet train loss and Mae on the second prediction

Figure 54 shows the Attention-PowerWiNet predictions against actual wind power values of the dataset2, with an R^2 of 0.88. The red predicted curve generally reflects the changing trend of the true values shown in blue. Overall, the prediction results demonstrate that the model is capable of providing accurate and reliable forecasts for most time periods.

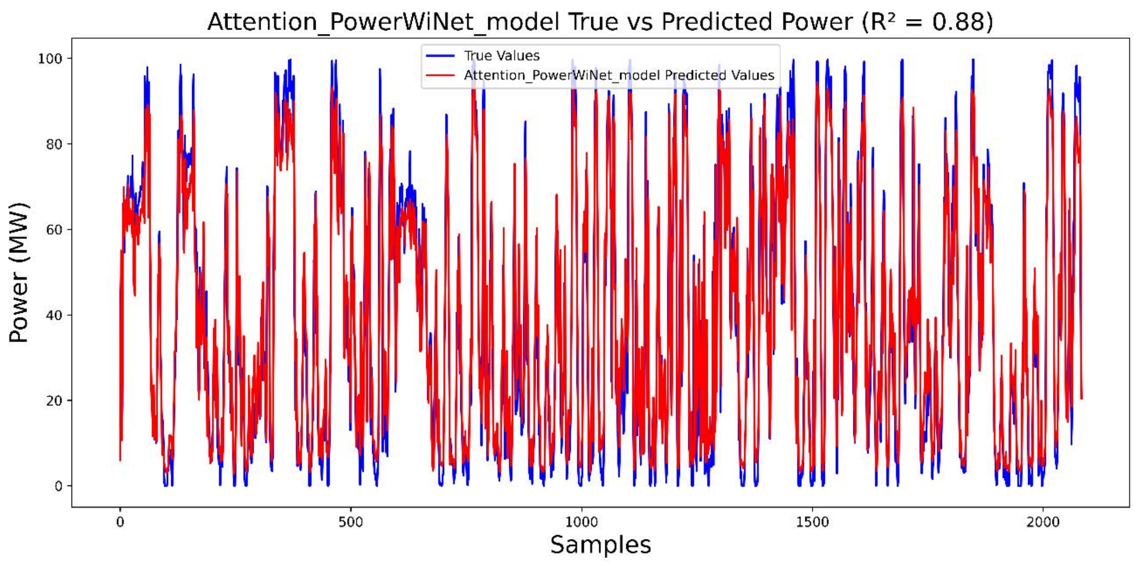


Figure 54. Comparison curve between predicted wind energy value and true value of Attention-PowerWiNet

4.5.5 Comparison of PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet

In general, PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet all perform well in the dataset2 as the square values of R are all relatively good.

4.6 Comparison of the two predictions

Table 8. Summary of comparison indicators

	Model	MAE	RMSE	R^2
Dataset 1 (Spring)	PowerNetwork	0.047	3.079	0.796
	WindFormer Network	0.049	3.227	0.776
	PowerWiNet	0.045	3.045	0.801
	Attention-PowerWiNet	0.037	2.707	0.842
Dataset 1 (Summer)	PowerNetwork	0.045	2.969	0.886
	WindFormer Network	0.044	3.104	0.875
	PowerWiNet	0.040	2.874	0.893

	Attention-PowerWiNet	0.036	2.725	0.904
Dataset 1 (Autumn)	PowerNetwork	0.069	4.054	0.790
	WindFormer Network	0.036	2.786	0.901
	PowerWiNet	0.036	2.756	0.903
	Attention-PowerWiNet	0.035	2.739	0.904
Dataset 1 (Winter)	PowerNetwork	0.043	2.642	0.878
	WindFormer Network	0.036	2.393	0.899
	PowerWiNet	0.038	2.384	0.900
	Attention-PowerWiNet	0.035	2.314	0.906
Dataset 2	PowerNetwork	0.080	10.949	0.869
	WindFormer Network	0.083	10.883	0.871
	PowerWiNet	0.798	10.873	0.871
	Attention-PowerWiNet	0.077	10.417	0.882

In wind power generation forecasting, the Attention-PowerWiNet demonstrates remarkable superiority, consistently outperforming standalone PowerNetwork, WindFormer Network, and PowerWiNet across both the overall dataset and seasonal subsets. Specifically, on Dataset 1 (Winter), the Attention-PowerWiNet achieves the lowest MAE (0.035), as well as the highest R^2 (0.906). Attention-PowerWiNet maintains top performance across all four seasonal subsets, particularly excelling in the autumn, summer and winter data with significantly lower MAE and RMSE values compared to other models. Even on the more challenging Dataset 2, the Attention-PowerWiNet shows robust performance, achieving a lower MAE (0.077) and higher R^2 (0.882) than PowerNetwork and WindFormer Network. These results suggest that the Attention-PowerWiNet is a comparatively suitable option for wind power generation forecasting tasks.

4.7 Model Explainability

In recent year, the use of Deep learning and Machine Learning models has increased dramatically across a wide range of areas, such as finance and business, health care. In order for users to understand the predictive logic of the models and build trust, the models need to be clearly interpretable, making the decision-making process transparent and easy to understand [16]. As a result, the emerging field of Interpretable Artificial Intelligence (XAI) has emerged to enhance model transparency and allow users to understand their decision-making mechanisms, thereby building trusted AI systems [17]. As a feature interpretation method based on game theory, SHapley Additive exPlanations (SHAP) can provide intuitive feature importance interpretation for supervised learning without changing the structure of the model [18].

The project uses five SHAP diagrams as the model explainability of the best model (Attention-PowerWiNet). The five SHAP graphs are SHAP Decision Plot, Dependence Plot, Force Plot, Global Importance Bar, SHAP Summary. The corresponding SHAP plots for Dataset 2 will be interpreted as follows.

4.7.1 SHAP Decision Plot

Figure 55 present the SHAP Decision Plot of the proposed model. The vertical axis is arranged in descending order of feature importance, with the most influential feature at the top [19]; The horizontal axis shows the output range of the model (0.558-0.566). As can be seen in the figure, ‘cloudCover’ and ‘windSpeed’ are at the top, and their corresponding line segments vary the most, indicating that these two features play a decisive role in the prediction results. In contrast, features such as ‘dewPoint’ (dew point temperature) and ‘windBearing’ (wind direction) at the bottom have a more limited impact on the forecast results.

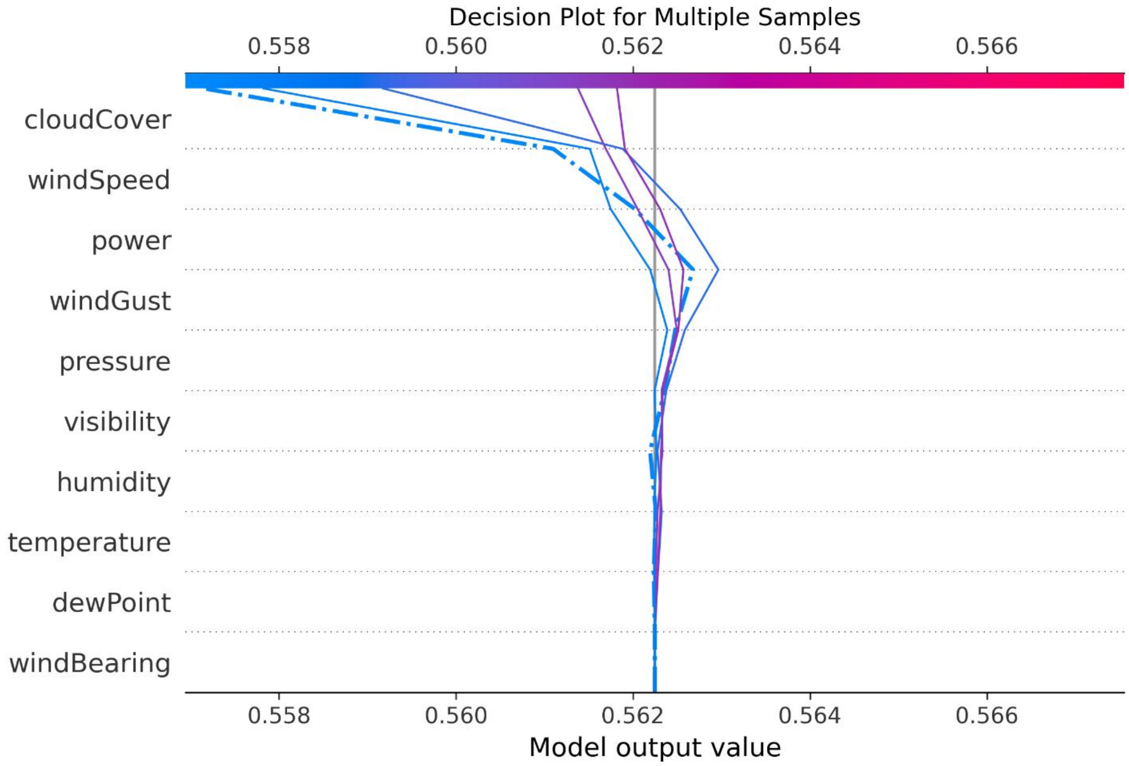


Figure 55. SHAP Decision Plot

4.7.2 Dependence Plot

Figure 56 present the SHAP Dependence Plot of the proposed model. It clearly shows the impact of 'cloudCover' characteristics on model predictions and their interaction with 'windSpeed'. The SHAP dependence plot reveals a nonlinear relationship where cloud cover has a slightly positive effect at low levels but becomes significantly negative at higher concentrations. Importantly, wind speed modulates this effect - high wind speeds amplify cloud cover's negative impact while low winds neutralize it. All in all, the plot not only shows the influence of single features, but also highlights the interaction effect between features.

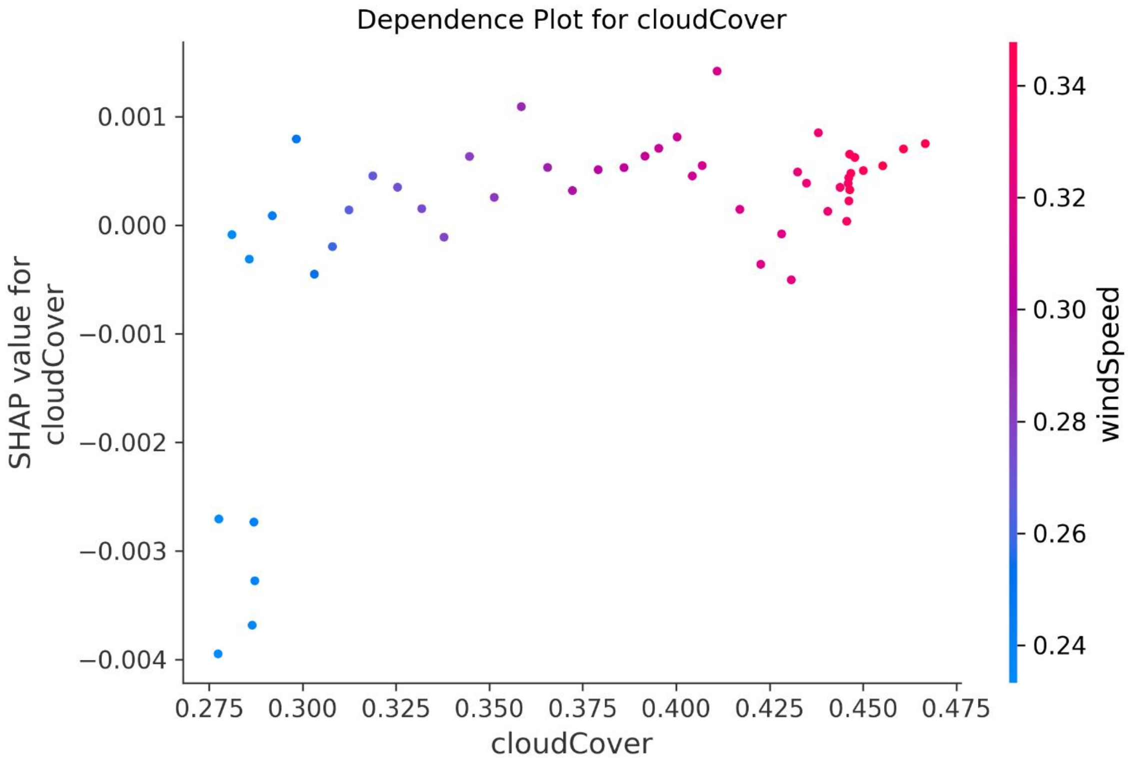


Figure 56. Dependence Plot

4.7.3 Force Plot

Figure 57 present the SHAP Force Plot of the proposed model. It explains the prediction for a single sample (Sample 0). The base value is approximately 0.562 and the final prediction is around 0.556, showing that the model predicted a slightly below average result. The red arrows show features hat decrease the prediction, contributing negatively. In contrast, the blue arrows show features that increase the prediction, contributing positively. The final predicted value is the result of the model combining all the features.

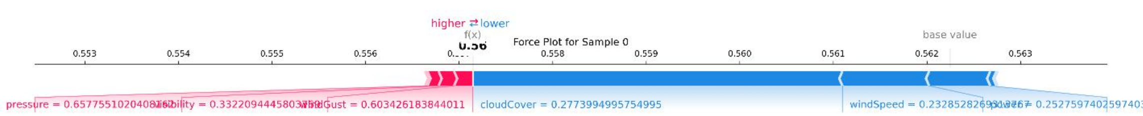


Figure 57. Force Plot

4.7.4 Global Importance Bar

Figure 58 present the Global Importance Bar of the proposed model. It display the global feature importance determined by each feature's mean absolute SHAP values [20]. The most significant feature is ‘cloudCover’, while ‘temperature’ and ‘windBearing’ have little effect on the model’s result.

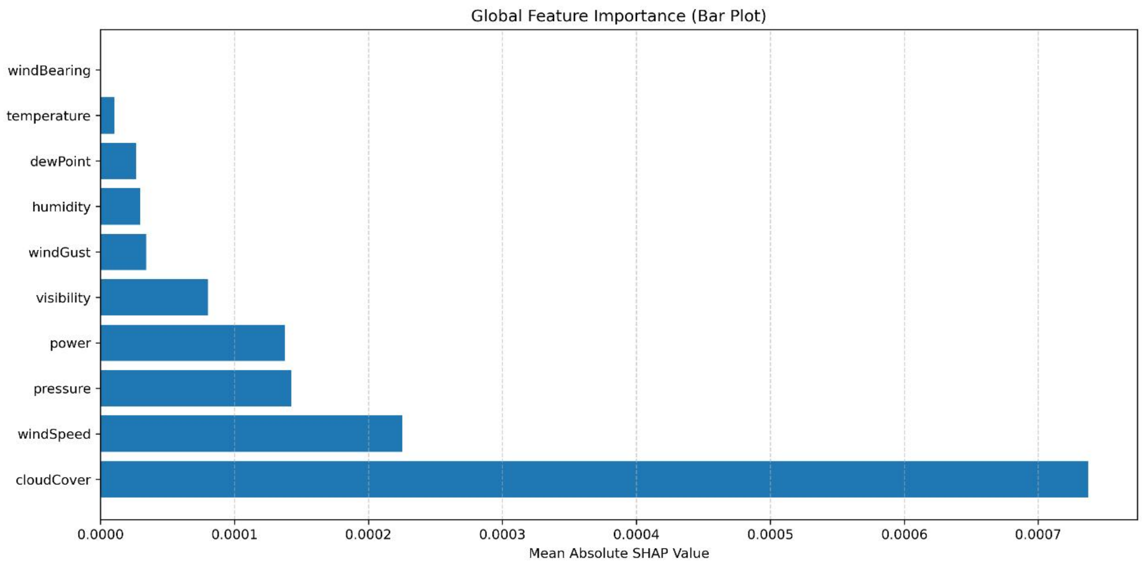


Figure 58. Global Importance Bar

4.7.5 SHAP Summary

Figure 59 demonstrates present the SHAP Summary Plot of the proposed model. It shows that ‘cloudCover’, ‘windSpeed’, and ‘pressure’ are the most influential features on the model's output. In general, predictions are increased by high cloudCover values and decreased by low ones. Features such as ‘wind Bearing’ and ‘temperature’ that are lower on the plot have less effect.

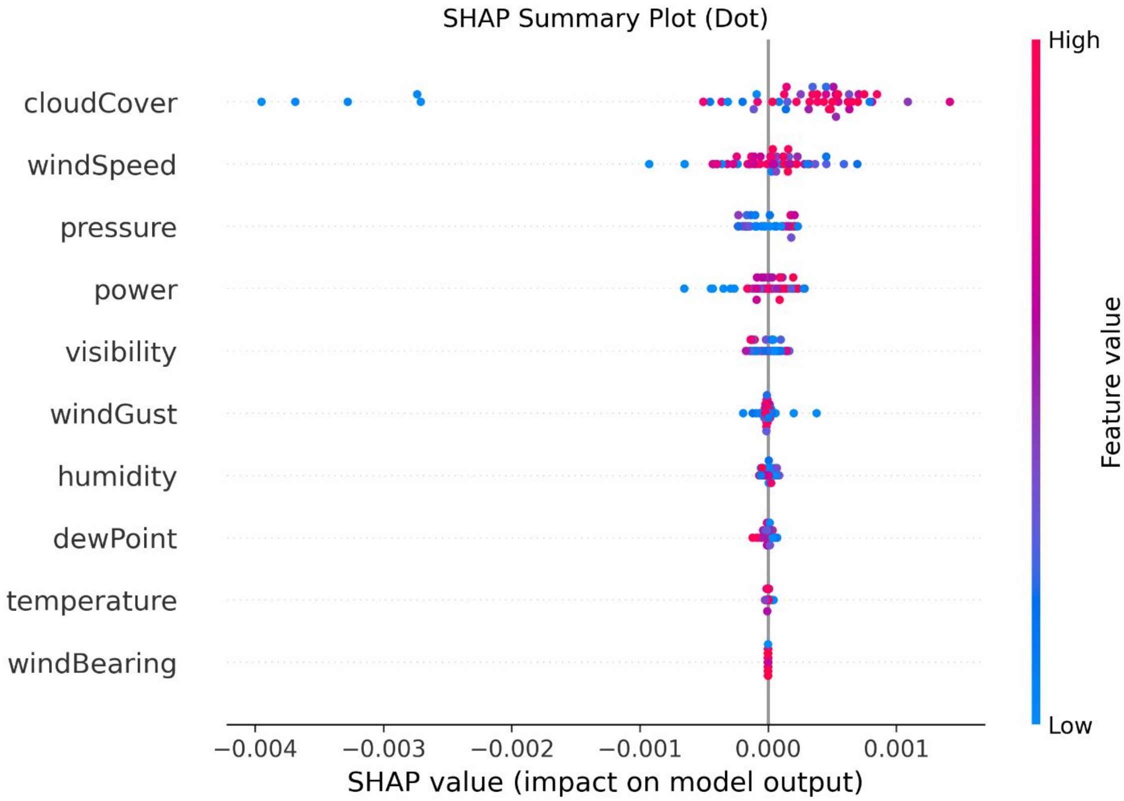


Figure 59. SHAP Summary

4.8 GUI design

This project develops an interactive wind power forecasting platform to realize the visual application of wind energy output prediction through a user-friendly web interface. The platform provides two core features: support for uploading wind farm operation data in CSV format, and power generation prediction based on the Attention-PowerWiNet.

- Figures 60-61 show the home page of the page, which provides basic information about the project and contains two buttons that jump to the prediction screen. Finally, users can jump to the web page to see the dataset used.

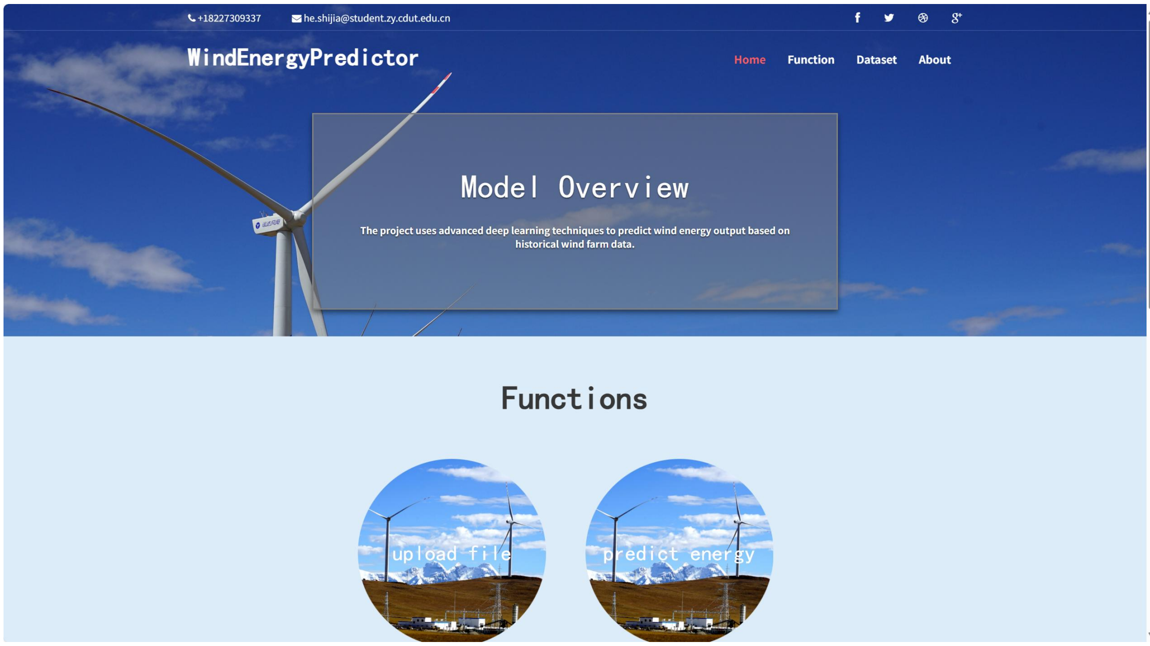


Figure 60. Home page 1

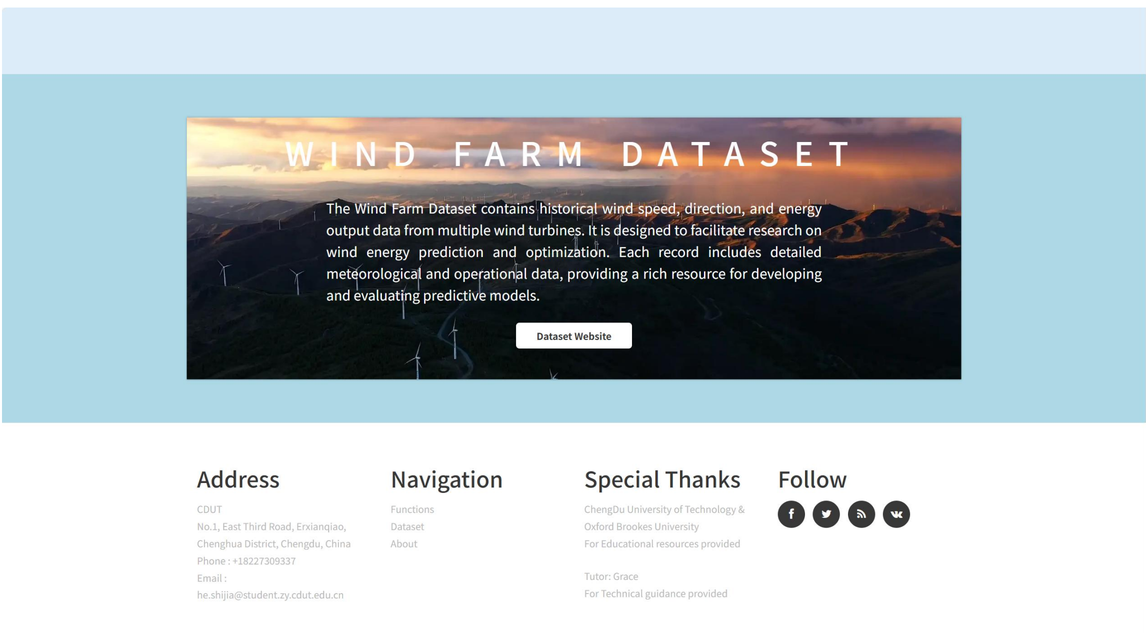


Figure 61. Home page 2

- Figure 62 shows the prediction interface, which is divided into two modules, the left module for the user to upload the csv file, and the right module for getting the prediction results.

Figure 62. Predict page

- Figure 63-64 shows the prediction interface. After the user uploads the document and clicks the 'Predict Energy' button, the result will be obtained in the following module with five SHAP plots.

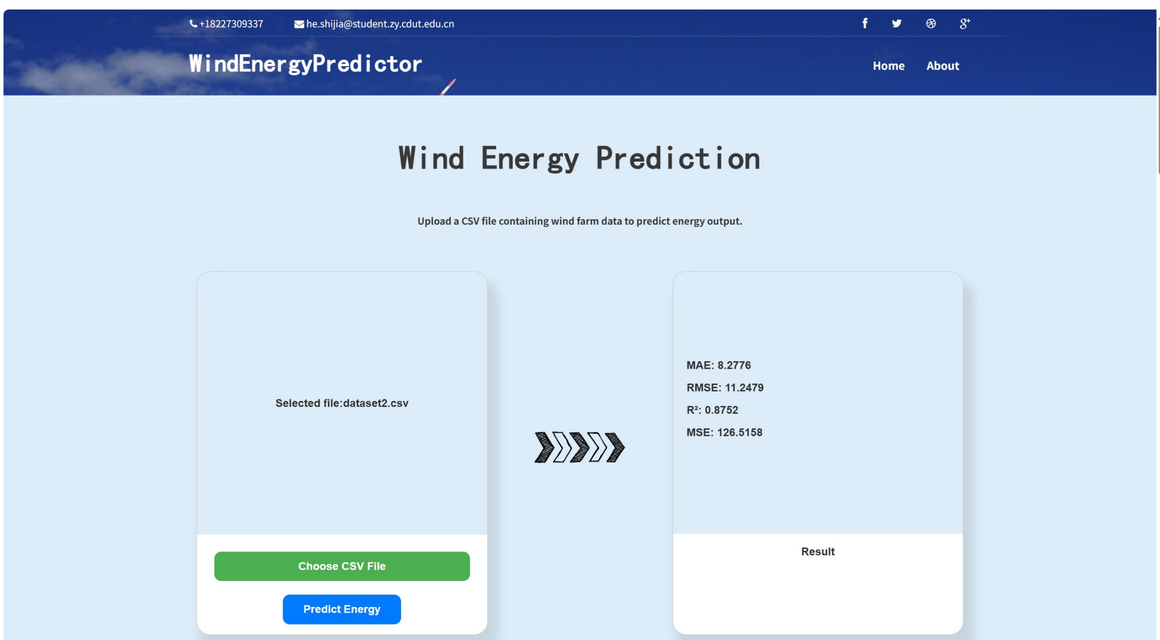


Figure 63. Result page 1

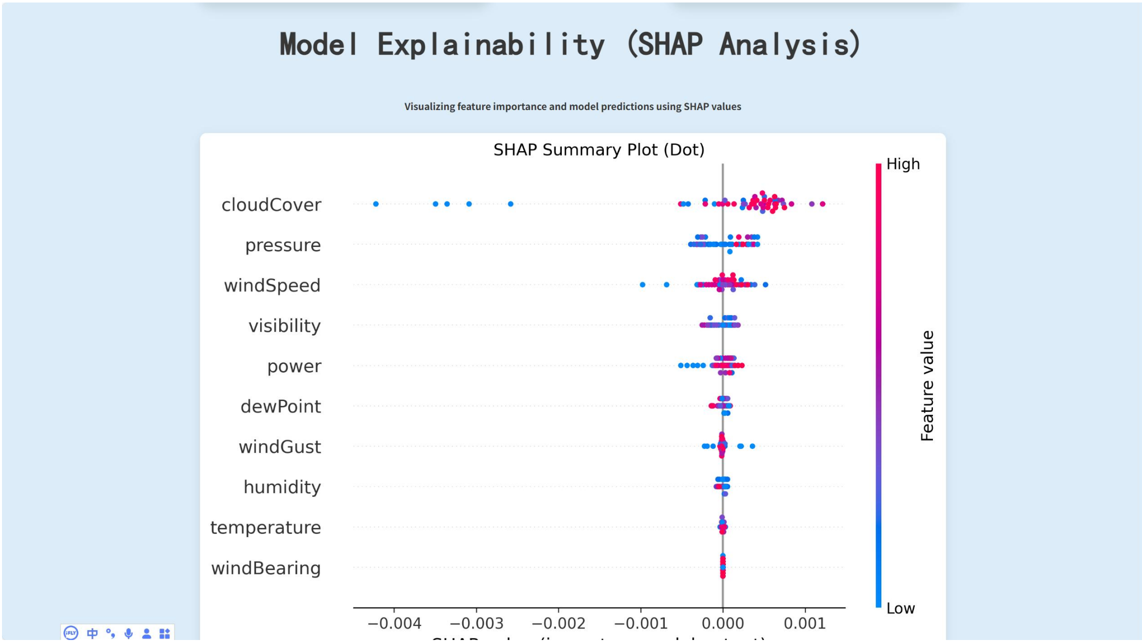


Figure 64. Result page 2

Chapter 5 Professional Issues

5.1 Project Management

In chapter 5, presented the professional issues related to this project. This section covers project management, risk analysis, and ethical, legal, social, and environmental considerations.

5.1.1 Activities

Table 9. The steps of each task

Phase	Objectives	State
1. Preparation	1. Understand the basics of the Proposed Titles. 2. Identify key issues and challenges. 3. Seek relevant solutions and previous research. 4. Choose a suitable prediction model.	Completed
2. Model Exploration	1. Study the CNN and LSTM models. 2. Understand loss functions, optimizers, and training techniques.	Completed
3. Data collection	1. Gather at least 2-3 publicly available wind farm history data datasets. 2. Decide on the training, test data, and valid data ratio.	Completed
4. Development and Implementation	1. Data process. 2. Build the PowerNetwork, WindFormer Network, PowerWiNet model architectures and define metrics for model evaluation. 3. Search for papers related to deep learning attention mechanism, and formulate corresponding attention mechanism for PowerWiNet model. 4. Train, analyse, and compare model	Completed

	performance; adjust hyperparameters for optimization.	
5. Testing	1. Change another dataset to evaluate this model. 2. Analyse the outcome and gather the all work.	Completed
6. Design a GUI	1. Understand the definition of GUI 2. Design the architecture for the GUI. 3. Identify the technologies used by the GUI, such as the flak framework. 4. Implement the code for the GUI.	Completed
7. Project Summary and Presentation	1. Write the final project report and demonstrate the achievement using slides. 2. prepare the presentation.	Completed

5.1.2 Schedule

The Gantt chart of time planning as below:

Table 10. Gantt chart

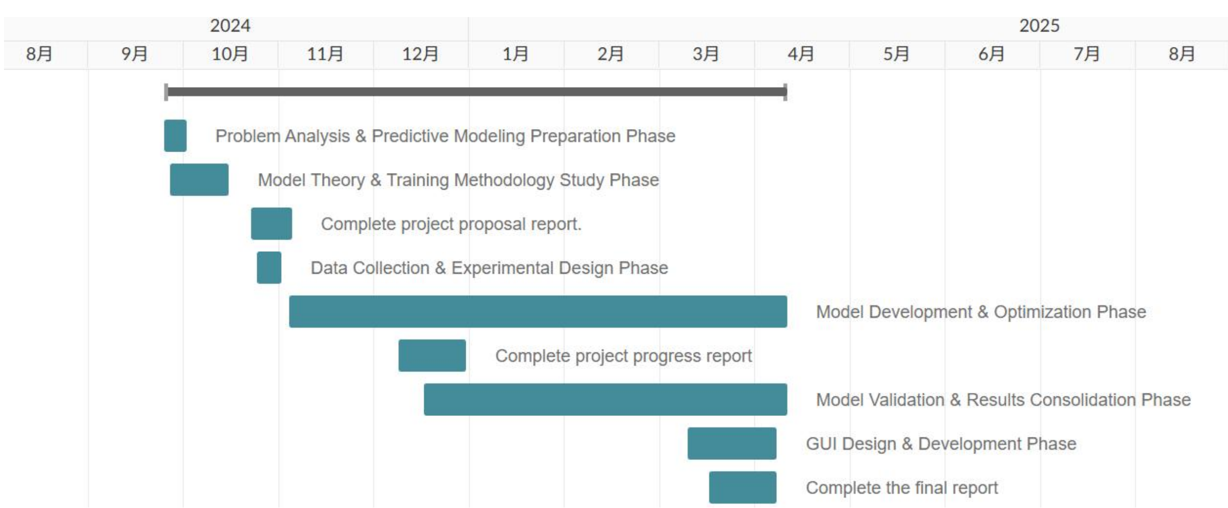


Table 11. Timetable

Task	Start date	End date	Duration
Problem Analysis & Predictive Modeling Preparation Phase	2024/9/25	2024/10/1	7 days
Model Theory & Training Methodology Study Phase	2024/9/27	2024/10/15	18 days
Complete project proposal report.	2024/10/23	2024/11/4	13 days
Data Collection & Experimental Design Phase	2024/10/25	2024/11.1	7 days
Model Development & Optimization Phase	2024/11/4	2025/4.10	160 days
Complete project progress report	2024/12/9	2024/12/30	22 days
Model Validation & Results Consolidation Phase	2024/12/17	2025/4.10	116 days
GUI Design & Development Phase	2025/3/20	2025/4/7	19 days
Complete the final report	2025/3/17	2025/4/7	22 days

5.1.3 Project Data Management

Resources including datasets, project model codes, progress reports, related literature, and so on will be stored in a local file and uploaded to git repositories.

Git repositories link: https://github.com/Nicole-HeShiJia/Project.

5.1.4 Project Deliverables

1. Datasets

2. All model codes

3. PowerNetwork, WindFormer Network, PowerWiNet, and Attention-PowerWiNet models' diagram

4. All model evaluation plots documents.

5. Ethics Forms

6. Reports (Proposal, Progress, Final)

7. Presentation slides

8. Literature and references

9. poster

5.2 Risk Analysis

The risks analysed as the project progresses are shown in Table 12.

Table 12. Risk Analysis

Risk ID	Potential Risk	Potential Causes	Severity	Risk	Mitigation
R1.1	Experimental results using the first dataset requirements (satellite images, weather data) are poor	Two datasets with time series could not be found.	4	6	Find relevant papers and use datasets from the literature.
R 1.2	Loss of project data	Physical hardware destruction	4	7	Protect the hardware and save the project data on the hard drive.
R 1.3	Miss the deadline	sick	3	3	Take care of my body.
		The student website is having problems	3	3	Ask Tina for help.

5.3 Professional Issues

Legal Issues: Historical wind energy data may include sensitive information about the wind farm's performance or operations. If the data originates from the European Union or similar frameworks in other regions with strict data protection laws, ensuring compliance with data protection regulations like General Data Protection Regulation (GDPR) is crucial.

Social Issues: The ability to predict wind energy accurately can improve the efficiency and reliability of renewable energy distribution, benefiting communities by reducing power outages and ensuring equitable energy access. However, to achieve this, the project must consider stakeholder engagement by involving wind farm operators and local communities, ensuring their needs are addressed and fostering trust in the technology.

Ethical Issues: Transparency and responsibility are important when developing a predictive model, as the results can impact wind energy distribution decisions and operational planning. The model should aim to be accurate and produce predictions that are easy to understand. This helps build trust in the system and ensures accountability if decisions based on the model lead to unexpected results.

Environmental Issues: The project can help optimize renewable energy use by improving wind energy predictions, making it easier to integrate renewable energy into the grid. This reduces reliance on fossil fuels and lowers greenhouse gas emissions, supporting efforts to fight climate change.

Using wind energy prediction models involves key legal, social, ethical, and environmental issues, such as data privacy, stakeholder engagement, responsible use, transparency, and energy efficiency. These concerns must be carefully managed to ensure the project is conducted responsibly and ethically.

Chapter 6 Conclusion

This project explored the application of deep learning techniques to improve wind power prediction accuracy using historical wind farm data. In this study, an innovative PowerWiNet hybrid neural network architecture based on attention mechanism enhancement is proposed which named Attention-PowerWiNet to solve the problems of time dependence and feature extraction in wind power prediction. In the experiment, three models (PowerNetwork, WindFormer Network, PowerWiNet) and Attention-PowerWiNet were used to compare the wind power prediction performance. The experimental results show that the four models of PowerNetwork, WindFormer Network, PowerWiNet and Attention-PowerWiNet all show good prediction ability on different datasets. The Attention-PowerWiNet model performs the best among them, particularly when using summer, autumn and winter datasets as the R^2 values and its mean absolute error (MAE) and root mean square error (RMSE) are lower, demonstrating its great generalization ability and robustness. In addition, The SHAP method was used to analyse the model prediction mechanism, and the key input features of the model (such as cloud cover and wind speed) were visually analysed, which verified the rationality of the model decision and improved the reliability of the prediction results. Furthermore, also developed a visual web interface to support users to upload historical wind farm data, real-time output forecast results and interpretable charts, providing a convenient and efficient data support tool for wind farm operators, grid schedulers and energy management departments.

However, the existing work has some limits. First, the data in this study mainly rely on historical environmental variables (wind speed, temperature, pressure, etc.) to predict wind energy. But wind energy output is also related to other factors, such as topography, fan height, and so on. Incomplete factors may limit a more comprehensive understanding of the actual wind farm. Second, model is not evaluated under abnormal data. In actual wind farm applications, there are often various abnormal disturbances, such as extreme weather and equipment damage, which will affect the prediction performance of the model.

In the future, this research will be expanded to the data and model level. First, at the data level, more factors related to wind energy output will be introduced, which can help the model to understand the wind power generation mechanism more comprehensively. Secondly, can introduce simulated noise, random missing eigenvalues or other ways to construct abnormal data scenes to improve the stability and practicality of the model in the real environment. Overall, this project provides some insights into wind power forecasting that can help wind farms operate better.

References

[1] 'Changes in the Global Structure of Energy Consumption and the Energy Transition Process'. Accessed: Mar. 15, 2025. [Online]. Available: https://www.researchgate.net/publication/385773688_Changes_in_the_Global_Structure_of_Energy_Consumption_and_the_Energy_Transition_Process

[2] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, 'The role of renewable energy in the global energy transformation', *Energy Strategy Reviews*, vol. 24, pp. 38–50, Apr. 2019, doi: 10.1016/j.esr.2019.01.006.

[3] T.-Z. Ang, M. Salem, M. Kamarol, H. S. Das, M. A. Nazari, and N. Prabaharan, 'A comprehensive study of renewable energy sources: Classifications, challenges and suggestions', *Energy Strategy Reviews*, vol. 43, p. 100939, Sep. 2022, doi: 10.1016/j.esr.2022.100939.

[4] 'Renewable energy as a solution to climate change: Insights from a comprehensive study across nations | PLOS One'. Accessed: Mar. 23, 2025. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0299807

[5] F. Perera, 'Pollution from Fossil-Fuel Combustion is the Leading Environmental Threat to Global Pediatric Health and Equity: Solutions Exist', *Int J Environ Res Public Health*, vol. 15, no. 1, p. 16, Jan. 2018, doi: 10.3390/ijerph15010016.

[6] S. A. Vargas, G. R. T. Esteves, P. M. Maçaira, B. Q. Bastos, F. L. Cyrino Oliveira, and R. C. Souza, 'Wind power generation: A review and a research agenda', *Journal of Cleaner Production*, vol. 218, pp. 850–870, May 2019, doi: 10.1016/j.jclepro.2019.02.015.

[7] P. Veers *et al.*, 'Grand challenges in the design, manufacture, and operation of future wind turbine systems', *Wind Energy Science*, vol. 8, no. 7, pp. 1071–1131, Jul. 2023, doi: 10.5194/wes-8-1071-2023.

[8] O. Summerfield-Ryan and S. Park, ‘The power of wind: The global wind energy industry’s successes and failures’, *Ecological Economics*, vol. 210, p. 107841, Aug. 2023, doi: 10.1016/j.ecolecon.2023.107841.

[9] P. Maegaard, ‘Wind energy development and application prospects of non-grid-connected wind power’, in *2009 World Non-Grid-Connected Wind Power and Energy Conference*, Sep. 2009, pp. 1–3. doi: 10.1109/WNWEC.2009.5335847.

[10] Y. Huo, C. Xu, Q. Li, Q. Li, and M. Yin, ‘Maximizing wind farm production through pitch control using graph neural networks and hybrid learning methods’, *IET Renewable Power Generation*, Oct. 2024, doi: 10.1049/rpg2.13133.

[11] A. Kusiak, H. Zheng, and Z. Song, ‘Short-Term Prediction of Wind Farm Power: A Data Mining Approach’, *IEEE Transactions on Energy Conversion*, vol. 24, no. 1, pp. 125–136, Mar. 2009, doi: 10.1109/TEC.2008.2006552.

[12] X. Wang, P. Guo, and X. Huang, ‘A Review of Wind Power Forecasting Models’, *Energy Procedia*, vol. 12, pp. 770–778, Jan. 2011, doi: 10.1016/j.egypro.2011.10.103.

[13] Y.-Y. Hong and C. L. P. P. Rioflorido, ‘A hybrid deep learning-based neural network for 24-h ahead wind power forecasting’, *Applied Energy*, vol. 250, pp. 530–539, Sep. 2019, doi: 10.1016/j.apenergy.2019.05.044.

[14] ‘Wind Power Prediction Based on LSTM Networks and Nonparametric Kernel Density Estimation | IEEE Journals & Magazine | IEEE Xplore’. Accessed: Mar. 11, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/8894464

[15] ‘Wind power prediction based on CNN-LSTM | IEEE Conference Publication | IEEE Xplore’. Accessed: Mar. 11, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/9713238

[16] B. Moreira Cunha and S. Diniz Junqueira Barbosa, ‘Evaluating the Effectiveness of Visual Representations of SHAP Values Toward Explainable Artificial Intelligence’, in *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*, in IHC ’24. New York, NY, USA: Association for Computing Machinery, Dec. 2024, pp. 1–11. doi: 10.1145/3702038.3702093.

[17] A. M. Salih *et al.*, 'A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME', *Advanced Intelligent Systems*, vol. 7, no. 1, p. 2400304, 2025, doi: 10.1002/aisy.202400304.

[18] A. V. Ponce-Bobadilla, V. Schmitt, C. S. Maier, S. Mensing, and S. Stodtmann, 'Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development', *Clin Transl Sci*, vol. 17, no. 11, p. e70056, Oct. 2024, doi: 10.1111/cts.70056.

[19] W. Chu, C. Zhang, H. Li, L. Zhang, D. Shen, and R. Li, 'SHAP-powered insights into spatiotemporal effects: Unlocking explainable Bayesian-neural-network urban flood forecasting', *International Journal of Applied Earth Observation and Geoinformation*, vol. 131, p. 103972, Jul. 2024, doi: 10.1016/j.jag.2024.103972.

[20] A. V. Ponce-Bobadilla, V. Schmitt, C. S. Maier, S. Mensing, and S. Stodtmann, 'Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development', *Clin Transl Sci*, vol. 17, no. 11, p. e70056, Oct. 2024, doi: 10.1111/cts.70056.

Appendices

The project is upload on GitHub, the URL is followed:

https://github.com/Nicole-HeShiJia/Project

The dataset 1 link:

https://github.com/jlian2/Tensorflow-Wind-Power-Prediction/blob/master/raw_data.csv

The dataset 2 link:

https://aistudio.baidu.com/datasetdetail/92257