School of Information Technology and
Engineering at the ADA University

School of Engineering and Applied Science
at the George Washington University

TITLE OF THESIS

EXAMINATION OF THE WORKING MODES OF THE 240 MW WIND POWER PLANT AI-BASED SHORT-TERM WIND POWER FORECASTING SYSTEM MULTI-SENSOR WIND SPEED PREDICTION AND HYBRID ENSEMBLE APPROACH

A Thesis
Presented to the Graduate Program of Electrical and Power Engineering
of the School of Information Technology and Engineering
ADA University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Electrical and Power Engineering
ADA University

By
Aliyev Rasim

December, 2025

THESIS ACCEPTANCE

This Thesis by: Rasim Aliyev
Entitled: *Title of Thesis in Italics*

has been approved as meeting the requirement for the Degree of Master of Science in Electrical and Power Engineering of the School of Information Technology and Engineering, ADA University.

Approved:

(Adviser)	(Date)
(Program Director)	(Date)
(Dean)	(Date)

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my industrial supervisor at AzerEnergy OJSC, Mr. Yusif Qarayev, First Assistant to the Vice President of AzerEnergy OJSC, for his continuous support and valuable guidance throughout this work. Despite his demanding schedule, he generously devoted time to discussions, shared his practical experience, and provided essential resources that greatly contributed to the successful completion of this research.

Furthermore, I would like to extend my deepest thanks to my academic supervisor at ADA University, Prof. Burcu Ramazanli, for her professional supervision, constructive feedback, and constant encouragement during all stages of this thesis. Her scientific insights, high academic standards, and patience have been invaluable in shaping the quality and direction of this study.

Finally, I am grateful to my family, friends, and colleagues for their moral support and understanding during the preparation of this thesis.

ABSTRACT

The rapid growth of large-scale wind power in Azerbaijan requires accurate short-term forecasting tools to support secure and economical power system operation. This thesis develops and evaluates an AI-based short-term wind speed and power forecasting framework using high-resolution measurements from a 120-m meteorological mast located near a 240 MW wind power plant. The dataset consists of one year of 1-second observations from multiple anemometers, wind vanes, temperature, humidity, and pressure sensors, which are preprocessed, quality-controlled, and aggregated to operationally relevant time scales.

The methodology combines advanced feature engineering, data decomposition, and machine-learning models. Physics-informed features such as wind shear, veer, turbulence intensity, and stability indicators are derived from the multi-level measurements. Tree-based algorithms (XGBoost and Random Forest) are used as fast and robust baselines, while a deep learning architecture based on a CNN–BiLSTM network with an attention mechanism is implemented to capture complex temporal dependencies. A regime-based design is adopted in which XGBoost provides accurate forecasts over the full operating range, and a Random Forest specialist is activated for high-wind conditions that are critical for turbine loading and grid security. Forecasted wind speeds are converted into active power using the manufacturer's power curve of the Envision 171/6.5 MW turbines.

The results show that the proposed framework achieves high accuracy and good generalization for 15-minute-resolution forecasts up to 24 hours ahead, with substantial error reductions in the high-wind regime compared to a single global model. The study demonstrates that combining multi-sensor measurements, physics-aware features, and hybrid AI models can deliver operationally meaningful wind power forecasts, supporting day-ahead scheduling, reserve allocation, and more reliable integration of large-scale wind farms into the Azerbaijani power system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Explanation
ABL	Atmospheric Boundary Layer
ADMM	Alternating Direction Method of Multipliers
AEP	Annual Energy Production
ANN	Artificial Neural Network
BiLSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
GRU	Gated Recurrent Unit
IMF	Intrinsic Mode Function
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
NWP	Numerical Weather Prediction
REWS	Rotor Equivalent Wind Speed
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
R^2	Coefficient of Determination
TI	Turbulence Intensity
TPE	Tree-structured Parzen Estimator
TSC	Time Series Classification
TSE	Time Series Evaluation (context-specific)
TCN	Temporal Convolutional Network

CHAPTER ONE
INTRODUCTION

In the last decade, the growth in wind power capacity has rapidly altered the operation and planning of modern power systems. By increasing the depth of variable renewable energy penetration, grid operators and market participants depend on highly accurate wind power forecasts to maintain system balance while reducing reserve requirements or optimizing dispatch decisions. Traditional forecasting methods based on statistical time series models and numerical weather prediction normally have difficulty capturing the highly nonlinear, multi-scale, and turbulent nature of wind dynamics, particularly in complex terrain and within the atmospheric boundary layer.

The advance of artificial intelligence and the availability of high-resolution measurement data have opened new opportunities for more accurate and robust wind power forecasting. To be specific, ultra-high-frequency met mast measurements allow for the development of hybrid models that combine physical understanding of the atmosphere with data-driven learning. Nevertheless, there is still a lack of systematic research on how to fully exploit 1-second resolution multi-height wind measurements and advanced decomposition-based ensemble architectures for short-term wind power forecasting.

1.1 Problem Statement

Despite the remarkable achievements achieved so far in AI-based wind power forecasting, a number of challenges still remain. First, most of the existing works are based on 10-min SCADA data or coarser resolutions, which cannot resolve high-frequency turbulence and rapid wind regime transitions. Second, most models are based on hub-height wind speed only and do not account for the vertical structure of the wind field and its effects on turbine rotor-equivalent inflow. Third, very few works have integrated the physics-informed features like wind shear, turbulence intensity, atmospheric stability, and REWS into state-of-the-art modern deep learning architectures. Finally, most of the decomposition-based hybrid ensemble models have been tested on small data sets, and there is limited analysis across different operating conditions, forecast horizons, and atmospheric regimes.

Consequently, this calls for an integrated methodology using ultra-high-resolution met mast data and physics-informed features with advanced hybrid ensemble architectures to enhance performance in short-term wind power forecasting.

1.2 Research Objectives

This thesis concentrates on developing, implementing, and evaluating an AI-based hybrid ensemble framework for short-term wind power forecasting using ultra-high-resolution (1-second) multi-height met mast data.

The stated specifics are:

Design a multi-resolution data processing and feature engineering pipeline to leverage 1-second, 10-second, 1-minute, and 10-minute resolutions in the representation of the wind field,

including physics-informed features such as REWS, wind shear, turbulence intensity, atmospheric stability, air density, and wind veer.

Apply and compare baseline forecasting models including persistence, Random Forest, XGBoost, LSTM, and CNN-BiLSTM with attention using multi-height met mast data.

Assess the proposed framework using a comprehensive set of error metrics and statistical tests for different forecast horizons, wind speed ranges, atmospheric stability regimes, and seasons. Quantify the contribution of ultra-high-resolution data, physics-informed features, decomposition methods, and ensemble strategies by performing systematic ablation studies.

1.3 Scope and Limitations

The present work focuses on wind speed and power forecasting for a single prospective wind farm site instrumented with one 120 m met mast, considering up to one full year of met mast data in the analysis. It also does not consider turbine SCADA data from an operating wind farm. Forecasting horizons from 1 min up to 6 h ahead are covered. Long-term forecasting (day-ahead and beyond) and direct coupling with numerical weather prediction models are outside the scope of this thesis. Additionally, proposed models are trained and validated using only data from a single geographic location, and generalization to other sites is not explicitly pursued within this thesis.

CHAPTER TWO

LITERATURE REVIEW

2.1 AI-Driven Wind Power Forecasting Methods

Accurate wind power forecasting reduces the risk of uncertainty and allows for better grid planning and integration of wind into power systems. The power system operators have to predict changes in the wind power production in order to schedule the spinning reserve capacity and to manage the grid operations. To reduce the reserve capacity and increase the wind power penetration, accurate forecasting of wind power is needed. In addition, wind power forecasting plays an important role in the allocation of balancing power. Besides, wind power forecasting is used for the day-ahead scheduling of conventional power plants and trading of electricity on the spot market. Forecasting models for wind power can be divided into two overall groups. The first group is based upon analysis of historical time series of wind, and a second group uses forecasted values from a numerical weather prediction (NWP) model as an input [24]. However, wind power forecasting is generally described in terms of physical methods, traditional statistical or 'black box' methods and more recently the so-called learning approaches, artificial intelligence or 'gray box' methods [21]. Hybrid methods can involve some aspects of all of these. The models in the first group use the statistical approach to forecast mean hourly wind speed or to directly forecast electric power production. The models in the second group use explanatory variables (mainly hourly mean wind speed and direction) derived from a meteorological model of the wind dynamics to predict wind power N-steps ahead. The models of the first group provide good results, in the majority of cases, there are three steps in wind power forecasting: firstly, determining wind speed from a model; then calculating the wind power output forecast or prediction [28]; and finally regional forecasting or upscaling or down scaling, which may be applied over different time horizons. Very short-term forecasting models are usually statistically based. For statistical and learning methods a large amount of historical time series data is essential [19].

In forecasting method Weather Observation is the initial conditions. Then the Numerical Weather Prediction models forecast the evolution of the weather system. Statistical models convert the wind into power output and correct the error pattern. Finally, the actual wind farm production data is feedback into the statistical models to improve the forecast. There are several different methods that can be used for wind power forecasting such as Physical Approach, and Statistical Approach. Physical Approach, which use parameterization based on the detailed physical description of the atmosphere, usually wind speed given by a weather service on a coarse grid is transformed to the onsite conditions at the location of the wind farm. Numerical Weather Prediction is a Physical Approach to wind power forecasting. Statistical Approach is based on training with measurement data and uses difference between the predicted and actual wind speeds in immediate past to tune the model parameters [1]. There are two types of statistical approaches. They are Time series-based models and neural network-based models.
. Forecasting in this context focuses on predicting power generation rather than load. AI and machine learning models have emerged as powerful tools for short-term forecasting [41,42,14,62]. These include artificial recurrent neural networks (ARNNs) [13,7], fuzzy logic controllers (FLCs) [2,4], support vector machines (SVMs), and hybrid models combining these methods.

Fuzzy logic models, particularly the Takagi-Sugeno-Kang (TSK) fuzzy inference systems [11], offer unique advantages in this context. These models interpret linguistic variables such as "low," "medium," or "high" wind speed and direction, and use IF-THEN rules to derive predictions [26]. The fuzzy inference system handles uncertainty in input data and uses spatial correlation between wind measuring stations to improve prediction accuracy. Training is conducted using a genetic algorithm, which optimizes both the fuzzy membership functions and rule parameters [23]. This approach has demonstrated superior performance compared to traditional persistence-based models, especially for forecasting horizons from 30 minutes to 2 hours ahead [20].

By improving wind speed forecast precision, grid operators in Azerbaijan can better anticipate fluctuations, optimize dispatch planning, and reduce reserve margins. This is particularly important as wind farms begin feeding into a grid currently dominated by thermal generation.

2.2 Grid Stability Challenges with High Wind Penetration

The injection of wind power can affect not only the connected substations but also the performance of the broader zonal transmission network. An important consideration in evaluating the grid impact of a wind power plant is the set of heuristics and regulatory rules used to prioritize renewable energy generation. Each country has its own framework for incentivizing and prioritizing the dispatch of renewable energy.

The impact of wind power on the electric power system depends on several factors, including the wind power penetration level, the size of the grid, and the generation mix within the system. When the penetration level is less than 5%, it typically does not pose any issues for the grid operator. However, when penetration exceeds 10%, grid adaptation and remedial measures become necessary. At penetration levels above 20%, strengthening of the existing grid becomes essential. Integrating large volumes of wind power into the grid can compromise both dynamic (frequency) and static (voltage) stability [6]. Inverter-based wind farms contribute little to system inertia, making frequency regulation more difficult during disturbances. Reduced inertia accelerates frequency deviations, requiring advanced frequency response mechanisms [15,59,35]. In such cases the LFC (Load Frequency Control) comes into play and the bulk generators of the grid adjust their generation through their speed governors and correct the variation to zero. In doing so it is possible that the flow pattern changes and the LDC (Load Dispatch Centre) will decide whether the system is in a normal state or alert state and will initiate suitable preventive control if needed. If the variation is large, which the governors cannot take care of, the emergency or extremis state is reached, and suitable procedure is to be followed by the LDC.

Similarly, voltage stability can be impacted by rapid changes in wind generation. This is especially problematic in weaker sections of the grid or when using older fixed-speed turbines that lack voltage control capabilities. If the voltage drops by more than 20% during a fault, the WPPs may disconnect from the grid. But with greater grid penetration of WECS (Wind Energy Conversion System), grid operators want the WPP to stay connected even during the faults and continue to support the power system. Such capability of the WPP is called Fault-Ride-Through capability. For example, WPPs have to stay connected to the grid for 100ms for a three-phase fault, for 100ms for two phase fault with or without involvement of ground followed by an auto re-closure after 300ms to 500ms. To maintain grid reliability and stability, Azerbaijan will require to implement control mechanisms on wind turbines, reinforce transmission

infrastructure, and deploy reactive power compensation devices. To deal with these challenges, strategies like the AZURE project [3] can be shown, which include constructing new 330 kV substations and transmission lines.

2.3 Role of FACTS and Grid-Support Technologies

As wind power continues to play a larger role in modern energy systems, it brings with it some important challenges especially when it comes to keeping the grid stable, maintaining power quality, and handling faults smoothly. To deal with these issues, researchers have explored using FACTS devices like STATCOM, SSSC, and UPFC. These technologies have shown strong technical potential in improving grid performance. However, most of the research so far has mainly focused on how well they work technically, without giving much attention to whether they make sense financially. Several studies have worked on optimizing the placement and operation of FACTS devices. For example, one recent paper used a Modified Whale Optimization Algorithm to find the best locations for FACTS devices with the goal of reducing transmission losses. Although these and similar studies showed promising technical results, they generally did not account for costs or long-term financial returns, which are critical for real-world implementation.

STATCOM, in particular, has received a great deal of attention. It has been widely recognized for its fast and reliable voltage regulation, especially in systems with high levels of renewable energy. A comprehensive review of STATCOM technology highlighted its various modeling methods, control strategies (including intelligent and adaptive controllers), and real-world applications. However, even this broad survey noted that there is still room for innovation especially when it comes to integrating technical performance with economic and environmental considerations.

When it comes to wind farms, the unpredictable nature of wind creates unique challenges. While many technical solutions have been proposed to manage faults and maintain grid stability, very few studies have taken a step further to evaluate how financially viable these solutions are. Most comparisons between devices like STATCOM, UPFC, or SSSC have been purely technical, without considering important factors such as installation cost, operational savings, and revenue from energy retention or carbon credits [9].

By introducing a novel framework that combines technical and economic performance through dedicated indices, the study allows for a more balanced and practical assessment of FACTS devices in wind power applications. It introduces three key indices: one for technical performance (Improved Voltage Index), one for economic evaluation (Cost-Related Index), and a combined Techno-Economic Index that helps decision-makers weigh their priorities.

2.4 Control Strategies for Managing Wind Penetration

Modern wind turbines support advanced control strategies, including synthetic inertia, primary frequency response, and voltage regulation. By adjusting output based on grid frequency or voltage, wind turbines can actively stabilize the system. Curtailment strategies are also used to limit steep power ramps, protecting the grid from overloading [8].

Forecast-informed dispatch planning, which uses wind predictions to optimize reserve commitments, is another key strategy. Integration of wind farm control systems with energy storage and conventional generation allows for smoother transitions and better system reliability. Azerbaijan can adopt similar coordinated control approaches to integrate new wind farms while maintaining operational resilience.

2.5 Role of Energy Storage Systems in Balancing Wind Variability

Energy Storage systems have enormous potential to increase the effectiveness of energy-conversion equipment use and for facilitating large-scale fuel substitutions in the world's economy. Energy storage is the capture of energy produced at one time for use at a later time [16]. Energy comes in multiple forms including radiation, chemical, gravitational potential, electrical potential, electricity, elevated temperature, latent heat and kinetic. Energy storage involves converting energy from forms that are difficult to store to more conveniently or economically storable forms. Traditional energy sources like coal, oil and natural gas-based power systems have to be turned on and off as demand fluctuates and are almost never operating at peak performance. This means that energy not only costs more but pollutes more than is necessary to meet our energy needs. And the slow ramp up time of these bulk generation facilities means they cannot respond to spikes in demand in real time, potentially leading to brownouts and poor power quality [27]. With the widespread adoption of renewable energy resources, energy storage is very significant. As is often noted, these energy sources are intermittent in nature, producing energy when the sun is shining and the wind is blowing. By storing the energy produced and delivering it on demand, these clean technologies can continue to power our grid even when the sun has set and the air is still-leveling out jumps in output to create a continuous, reliable stream of power throughout the day. Storage technologies also improve the quality of power through frequency regulation, allow us to produce power when it is cheapest and most efficient, and provide an uninterruptible source of power for critical infrastructure and services.

Energy demand in the commercial, industrial, public, residential, and utility sectors varies on a daily, weekly, and seasonal basis. Therefore, Energy Storage provides an alternative method of supplying peak energy demands. Likewise, Energy Storage systems can improve the operation of cogeneration, solar, wind, and run-of-river hydro facilities [62].

2.6 Contribution of Spinning Reserve to Wind Energy Integration

Spinning reserve (SR) plays a key role in supporting the integration of wind energy by acting as a buffer that quickly responds to sudden drops in wind generation or unexpected demand changes. As wind power is naturally variable and difficult to predict, maintaining power system stability becomes more complex. This method improves how SR is allocated by accounting for multiple uncertainties such as wind fluctuations, load forecast errors, equipment failures, and the condition dependent failure rates of generators and lines. By using a risk-based approach through Expected Energy Not Served (EENS), the method ensures enough reserve is in place to reduce outages and wind curtailment. Unlike traditional models that rely on slow, iterative processes, this approach uses efficient mathematical tools to speed up calculations and works even in large systems with high wind penetration. It also adapts SR levels to real-time grid conditions, avoiding both under and overallocation. Overall, the method balances reliability

and cost, offering a practical and scalable solution for integrating more wind power into the grid.

2.7 CONCLUSION

The transition toward a low-carbon electricity sector hinges on the successful integration of variable renewable energy (VRE) sources chiefly wind and solar into existing power systems. The literature reviewed here highlights six inter-locking themes that must be addressed in parallel if countries such as Azerbaijan are to meet ambitious renewable-capacity targets while safeguarding reliability and power quality.

Azerbaijan's pursuit of large-scale wind integration presents both opportunities and challenges. Leveraging AI-based forecasting, advanced control systems, FACTS devices, and energy storage technologies will be vital to maintain grid stability and power quality. Early investments in grid modernization and strategic planning such as the AZURE project and industrial-scale BESS demonstrate a readiness to adapt best practices from global leaders [62]. With coordinated policy, regulation, and infrastructure development, Azerbaijan is well positioned to integrate its wind resources efficiently, strengthening energy security and contributing to decarbonization goals.

Forecast accuracy is foundational. AI-driven wind-power forecasting models spanning recurrent neural networks, support-vector machines, and hybrid fuzzy-logic approaches consistently outperform persistence and classical statistical methods for horizons from minutes to a day ahead. Better forecasts translate directly into lower reserve margins, reduced curtailment, and more economical dispatch planning.

Grid stability constraints intensify with penetration**.** Once wind energy surpasses roughly 10 % of system demand, both dynamic (frequency) and static (voltage) limits emerge. The weaker network sections around Baku underscore the need for synthetic inertia, primary frequency response, and robust fault-ride-through (FRT) capabilities in modern turbines.
FACTS and other grid-support technologies must be judged on techno-economic grounds. Devices such as STATCOMs, SSSCs and UPFCs demonstrably improve voltage regulation and damping, yet deployment decisions should be guided by combined technical-performance and life-cycle-cost indices not by technical benefit alone.

Energy-storage systems are the linchpin for balancing variability. Batteries, flywheels, and other storage media can absorb surplus production, supply fast frequency response, and shift renewable output to peak periods, thereby smoothing the net-load profile and easing pressure on conventional generators.

Spinning-reserve allocation needs to be risk-based and adaptive. Methods that explicitly model wind, load, and equipment-failure uncertainties using metrics such as Expected Energy Not Served prevent both under- and over-provisioning, keeping reliability high without excessive cost.

Holistic integration studies remain essential. Every decision plant siting, network reinforcement, control strategy, market rule should be stress-tested through steady-state,

dynamic, and stochastic simulations that reflect realistic fuel prices, CO_2 costs, and demand-response capabilities.

METHODOLOGY

3.1 Overview

This chapter presents the methodology for AI-based short-term wind power forecasting using ultra-high-resolution meteorological tower (met mast) data and hybrid ensemble approaches. The framework combines state-of-the-art deep learning architectures with physics-informed feature engineering and multi-resolution data processing. The methodology is organized into five phases: first data acquisition and preprocessing, second multi-resolution feature engineering, third Baseline models were developed using XGBoost, Random Forest, and a CNN–BiLSTM with attention, four hybrid ensemble implementation, and five evaluation and validation.

The methodology is organized into five main phases:

1. **Data acquisition and preprocessing**
2. **Multi-resolution feature engineering**
3. **Baseline model development** using XGBoost, Random Forest, and a CNN–BiLSTM model with attention
4. **Hybrid ensemble implementation** based on decomposition and component-specific models
5. **Evaluation and validation** using multi-horizon error metrics and stratified analysis

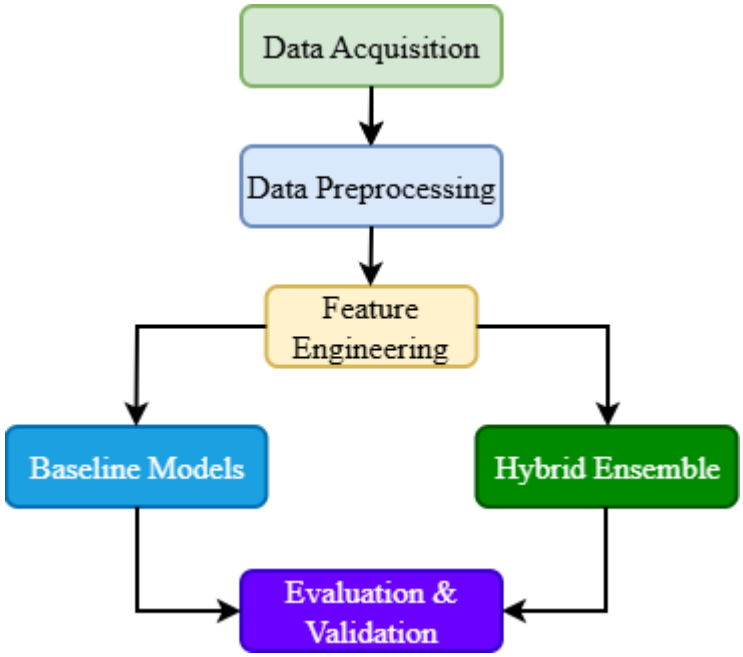


Figure 3.1 shows the overall methodological framework.

The study uses data from a 120-meter meteorological tower with sensors installed at strategically selected heights to capture the full vertical wind profile. The sensor configuration is provided in Table 3.1. Meteorological Tower Sensor Configuration

Sensor Type	Quantity	Heights (m)	Orientation	Measurement Range
Anemometer	4	60, 80, 100, 120	240°	0-50 m/s
Anemometer	2	120, 80	60°	0-50 m/s
Wind Vane	3	60, 100, 117	60°	0-360°
Thermometer	2	10, 116	-	-30 to +70°C
Hygrometer	2	10, 116	-	0-100%
Barometer	1	10	-	800-1100 hPa

Table 3.1 The Meteorological Tower Sensor Configuration

The dual-anemometer configuration at heights of approximately 80 m and 120 m, with sensors mounted on opposite booms oriented at 60° and 240°, provides redundancy for data validation and enables the detection of boom wake interference effects (Kaimal & Finnigan, 1994). This setup complies with IEC 61400-12-1 standards for power performance measurements (IEC, 2017).

3.2 Dataset Characteristics

The dataset consists of one full year of continuous measurements (7/8/2024 12:00:00 AM–9/7/2025 11:59:00 PM) recorded at a 1-second temporal resolution, resulting in approximately 31.5 million observations per sensor. The complete dataset structure is shown in Table 3.2, and Table 3.2.1

Date/time	ANE_1;win	ANE_2;win	ANE_3;win	ANE_4;win	ANE_5;win	ANE_6;win	WD_1;win	WD_2;win	WD_3;win	TH_1;hum	TH_2;hum	TH_1;temp	TH_2;temp	BAR_1;air_
7/24/2024 19:41	16.8661	16.8152	16.3066	15.71	15.6235	15.4722	70.3125	68.9062	62.2266	69.8578	69.455	24.3924	24.81873	1020.557
7/24/2024 19:41	16.8661	17.0448	16.4902	16.2603	16.1741	15.3346	70.3125	68.5547	63.6328	69.8242	69.455	24.41254	24.81873	1020.522
7/24/2024 19:41	16.912	17.0907	16.3066	16.2603	16.3576	15.0592	72.4219	68.5547	67.5	69.8242	69.5221	24.39575	24.82208	1020.33
7/24/2024 19:41	16.7283	16.8611	16.4902	16.0769	16.1741	15.6099	71.0156	69.6094	63.9844	69.8242	69.5557	24.40582	24.82208	1020.47
7/24/2024 19:41	16.7742	16.8152	16.5361	16.3062	16.2199	15.4263	73.8281	69.6094	65.3906	69.8242	69.5557	24.41589	24.82208	1020.54
7/24/2024 19:41	16.8661	16.907	16.6279	16.4437	16.1741	15.2887	73.8281	71.3672	64.3359	69.8242	69.5221	24.38904	24.82544	1020.54
7/24/2024 19:42	16.8201	16.7233	16.5361	16.2144	16.1741	15.564	73.8281	71.3672	68.5547	69.8242	69.3878	24.42261	24.82544	1020.278
7/24/2024 19:42	16.8661	16.6774	17.0411	16.4896	16.7705	15.977	74.1797	71.0156	64.3359	69.8242	69.2535	24.40582	24.82544	1020.522
7/24/2024 19:42	17.0498	16.9989	16.7657	16.8106	16.8622	16.5735	71.3672	70.3125	63.9844	69.7235	69.1864	24.42596	24.83551	1020.522
7/24/2024 19:42	17.0498	16.9989	16.8116	16.7189	17.1375	16.8489	72.4219	69.9609	66.7969	69.7571	69.1864	24.40247	24.82544	1020.487
7/24/2024 19:42	16.8661	17.0448	16.9952	16.7648	16.7705	16.7571	73.4766	69.6094	65.0391	69.7571	69.1528	24.43268	24.8288	1020.365
7/24/2024 19:42	16.9579	17.0907	16.7657	16.352	16.2199	16.2523	72.7734	71.0156	63.9844	69.7906	69.1528	24.40918	24.83551	1020.435
7/24/2024 19:42	16.8661	16.907	16.3525	16.031	14.8436	15.2887	74.5312	73.4766	64.3359	69.7235	69.1528	24.42932	24.83215	1020.47
7/24/2024 19:42	16.6823	16.6774	16.582	15.1137	15.6235	14.1415	72.7734	72.4219	63.2812	69.7235	69.1528	24.45618	24.83215	1020.435
7/24/2024 19:42	16.6823	16.4019	16.5361	15.1137	15.6694	14.5086	73.125	72.4219	69.6094	69.6899	69.1528	24.41254	24.83887	1020.278
7/24/2024 19:42	16.223	16.4937	16.3984	15.1137	15.7612	14.1874	74.1797	72.4219	71.7188	69.6564	69.1528	24.44275	24.83215	1020.418
7/24/2024 19:42	16.4986	16.9989	16.6279	14.9303	15.5776	14.4168	73.8281	71.3672	70.3125	69.6564	69.1528	24.45953	24.83215	1020.522
7/24/2024 19:42	16.223	16.6315	16.2147	15.4348	15.5318	13.5908	75.5859	69.9609	77.3438	69.5892	69.1528	24.44611	24.83887	1020.435
7/24/2024 19:42	16.1312	16.2182	15.6638	15.1137	15.3024	12.9943	72.7734	69.2578	69.9609	69.5557	69.22	24.46625	24.83551	1020.33

Table 3.2.1- The complete dataset structure

Parameter	Specification
Temporal Coverage	12 months (365 days)
Sampling Rate	1 Hz (1-second intervals)
Total Observations	31,536,000 per sensor
Total Data Points	~441 million (14 sensors)
Data Volume	~15 GB (raw format)
Missing Data Rate	<2% (to be verified)

Table 3.3 - The discrete temporal variables

The 1-second temporal resolution represents a 600-fold increase in data density compared to the standard 10-minute SCADA data typically used in wind forecasting research (Lin & Liu, 2020; Wang et al., 2020). This ultra-high resolution enables:

Turbulence characterization:

Turbulence characterization is performed at scales relevant to wind turbine structural loading. This includes modeling short-term fluctuations for ultra-short-term forecasting (on the order of seconds to minutes) and extracting high-frequency features that capture rapid transitions between wind regimes. These steps improve the accuracy of derived turbulence intensity and related statistical parameters.

Target Variable Definition:

The primary target variable for forecasting is defined according to the Rotor Equivalent Wind Speed (REWS) concept specified in IEC 61400-12-1 (IEC, 2017). REWS offers a more accurate representation of the wind field experienced by the rotating turbine rotor compared to single-point hub height measurements.

For a turbine with hub height $H$and rotor radius R, the REWS is calculated as:

Equation 3.1: Rotor Equivalent Wind Speed (REWS)

The Rotor Equivalent Wind Speed (REWS) is expressed mathematically as:

$$V_{eq} = \left(\frac{1}{A} \sum_{i=1}^{n_h} A_i \, V_i^3 \right)^{\frac{1}{3}} \quad \text{REWS} = \left[\frac{\iint_A V^3(r,\theta)}{A} \right]^{\frac{1}{3}} \tag{3.1}$$

where:

REWS: Rotor Equivalent Wind Speed in m/s.

A: The total swept area of the wind turbine rotor in m², which may be computed as πR^2, considering R to be the blade length.

v(r, θ): The incoming wind speed at a given point on the rotor plane defined in polar coordinates (r, θ), distance from the hub, and angle.

dA: An infinitesimal element of area A.

$\iint_a \ldots dA$: A double integral over the whole rotor swept area representing the total kinetic energy flux across that area.

1/3: The cubic root, normalizing the power-weighted average wind speed back to a representative single wind speed value, since power is proportional to the cube of wind speed, $P \propto v^3$.This formulation accounts for spatial variations in wind speed over the rotor area, providing a more realistic measure of the wind field experienced by the turbine blades.

The rotor equivalent wind speed model based on equivalent power can be expressed as follows:

The wind turbine's rotor, or wind wheel, is a key component responsible for converting wind energy into mechanical energy. According to aerodynamic principles, the turbine's mechanical output power and torque can be derived as:

$$P = \frac{1}{2}\rho A V^3 C_P(\lambda, \beta) \tag{3.2}$$

$$T = \frac{P}{\Omega} = \frac{1}{2}\rho A V^2 C_T(\lambda, \beta) R \tag{3.3}$$

where:

- V— wind speed at hub height (m/s)
- ρ— air density (kg/m³)
- A— rotor swept area (m²)
- R— rotor radius (m)
- β— pitch angle (°)
- $\lambda = \frac{\Omega R}{V}$— tip speed ratio
- Ω— rotor angular velocity (rad/s)
- $C_P(\lambda, \beta)$— power coefficient, which according to Betz's law does not exceed 0.593

This formulation provides the foundation for calculating the rotor equivalent wind speed (REWS) based on equivalent power, allowing for a more accurate assessment of the effective wind field acting on the turbine rotor.

Discrete REWS Approximation:

For discrete measurement heights from the met mast, a sector-based weighted average is used:
Equation 3.4: Discrete REWS Approximation

$$\text{REWS} = \left[\frac{\sum_i V_i^3 A_i}{A_{\text{total}}}\right]^{\frac{1}{3}} \tag{3.4}$$

where,

- V_i is the measured wind speed at height z_i,
- A_i is the rotor disk sector represented by measurement i, and
- A_{total} is the total rotor swept area.

Three REWS variants are computed to reflect different turbine configurations:

- REWS_{80}: hub height = 80 m, rotor diameter = 120 m (typical 2–3 MW turbine)
- REWS_{100}: hub height = 100 m, rotor diameter = 140 m (typical 3–4 MW turbine)
- REWS_{120}: hub height = 120 m, rotor diameter = 170 m (typical 6–6.5 MW turbine)

3.3 Data Preprocessing

Data preprocessing plays a crucial role in converting raw met-mast sensor readings into a dependable dataset for short-term wind power forecasting. The original data, sampled at 1-second intervals from multiple sensors, often includes common field measurement issues such as missing data, sensor noise, outliers, and occasional equipment malfunctions. To handle these challenges and align with global standards, a structured preprocessing workflow is employed, following MEASNET (2016) and IEC 61400-12-1 protocols. This process ensures each variable is physically plausible, identifies and filters out anomalous readings, and synchronises sensor outputs in both time and units.

In addition to cleaning, preprocessing also aims to explore statistical patterns among the variables and validate the dataset's internal consistency. Once quality control is complete, a correlation analysis is performed to assess how wind speed, direction, and thermodynamic factors interact across sensor levels. The Pearson correlation matrix offers a concise view of these relationships, confirming the data's physical credibility and guiding the selection of features**,** dimensionality reduction strategies, and overall model architecture.

Figure 3.2 shows the Pearson correlation matrix for all met-mast variables after quality control was applied. The six anemometers display very strong positive correlations with each other ($\rho \approx 0.93$–1.00), indicating highly coherent airflow across the mast and suggesting reliable sensor performance, with no apparent faults or anomalies. Wind direction measurements are moderately correlated with one another ($\rho \approx 0.74$–0.87), while their relationship with wind speed is moderately negative ($\rho \approx -0.19$ to -0.33), reflecting expected sectoral shifts and directional veering under varying flow conditions. Thermodynamic variables show weaker but still notable correlations with wind speed: humidity and temperature are only weakly linked to the anemometers ($\rho \approx 0.10$–0.23), whereas air pressure shows a somewhat stronger positive relationship ($\rho \approx 0.31$–0.32), suggesting an influence from broader weather patterns. Strong internal correlations are also evident, such as between the two humidity sensors ($\rho \approx 0.68$) and

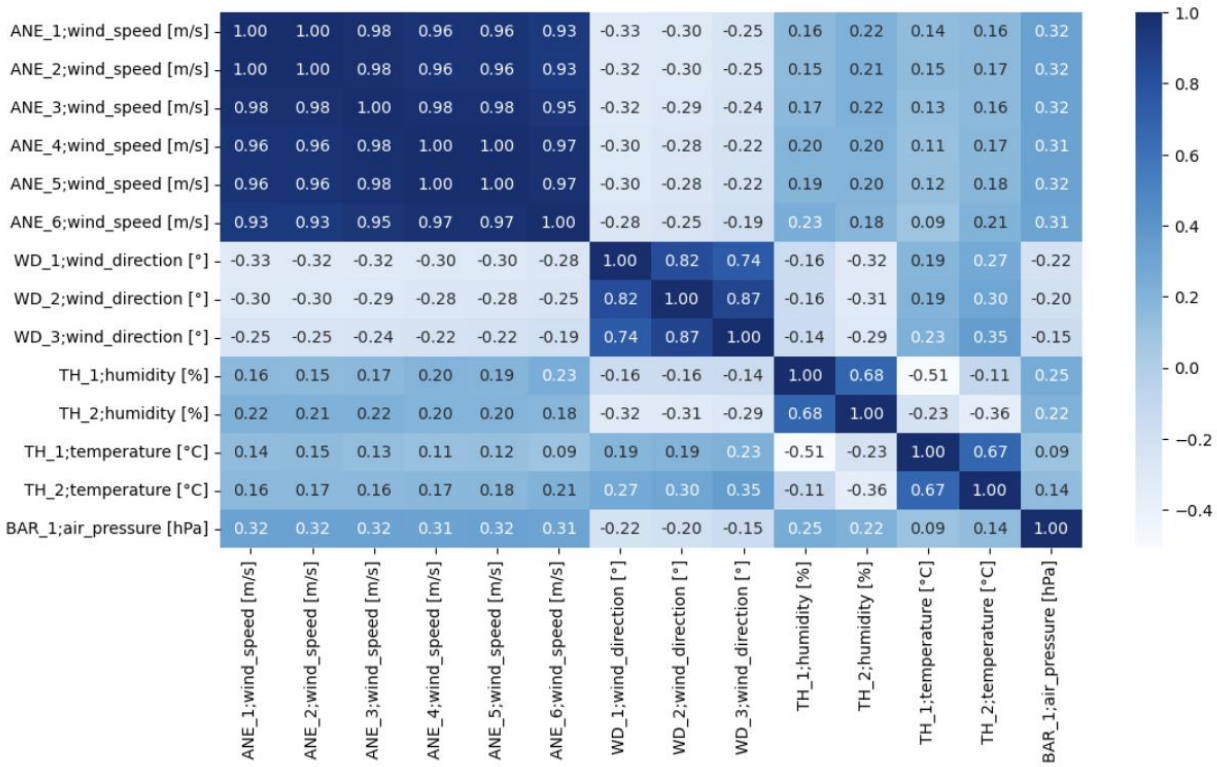


Figure 3.2 Correlation Structure of Met-Mast Variables

the two temperature sensors ($\rho \approx 0.67$), reinforcing the internal consistency of the data. Overall, the correlation analysis confirms the physical soundness and redundancy of the dataset, offering useful insight for feature selection and dimensionality reduction in later forecasting applications.

3.3.1 Quality Control and Anomaly Detection

A multi-stage quality control procedure is applied following MEASNET (2016) and IEC 61400-12-1. The pipeline consists of five sequential stages.

Stage 1: Range Testing

Physical limits are used to identify sensor malfunctions.

Equation 3.5: Range Validation

$$\text{Valid}(x) = \begin{cases} 1, & x_{\min} \leq x \leq x_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

Specified ranges by sensor type:

- Wind speed: $[0,50]$m/s
- Wind direction: $[0,360]$degrees
- Temperature: $[-40,50]$°C
- Relative humidity: $[0,100]$%
- Barometric pressure: $[950,1050]$hPa

Stage 2: Statistical Outlier Detection

Tukey's fence with a conservative threshold is used to flag outliers while preserving extreme events.

Equation 3.6: Interquartile Range (IQR) Method

$$\text{Outlier}(x) = \begin{cases} 1, & x < Q_1 - k \cdot \text{IQR or } x > Q_3 + k \cdot \text{IQR} \\ 0, & \text{otherwise} \end{cases} \tag{3.6}$$

with $Q_1, Q_3$the first and third quartiles, $\text{IQR} = Q_3 - Q_1$, and $k = 3$.

Stage 3: Dual-Sensor Cross-Validation

For paired anemometers at 80 m and 120 m, cross-validation detects boom interference or sensor faults.

Equation 3.7: Dual-Sensor Consistency Check

$$\delta = | V_1 - V_2 |, \quad \text{Flag} = \begin{cases} 1, & \delta > \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{3.7}$$

where $V_1, V_2$are simultaneous readings and $\text{threshold} = \max\big(0.5 \text{ m/s}, 0.1 \cdot \text{mean}(V_1, V_2)\big)$. When inconsistency is detected, the sensor closer to adjacent-height values is retained; if both appear valid, their mean is used.

Stage 4: Temporal Consistency Check

Physically implausible rates of change indicate transmission errors.

Equation 3.8: Rate-of-Change Test

$$\frac{\Delta V}{\Delta t} = \frac{| V(t) - V(t-1) |}{\Delta t}, \quad \text{Flag} = \begin{cases} 1, & \frac{\Delta V}{\Delta t} > \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

For 1-second data, the threshold is 5 m/s.

Stage 5: Missing Data Imputation

Gaps are treated by duration:

- Short gaps (< 10 s): linear interpolation
- Medium gaps (10 s–5 min): cubic spline interpolation
- Long gaps (> 5 min): marked as missing; models are trained to handle missing values.

Equation 3.9: Cubic Spline Interpolation

$$S(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, x \in [x_i, x_{i+1}] \quad (3.9)$$

3.3.2 Multi-Resolution Down sampling

To manage computational constraints and support multi-scale temporal modeling, the raw 1-second data is resampled into multiple time resolutions. This down sampling process maintains key statistical properties while significantly reducing data volume.
Resolution Hierarchy:

- 1-second (raw): retained only for turbulence intensity analysis
- 10-second: primary high-resolution dataset (631,152 observations per sensor per year)
- 1-minute: standard short-term forecasting resolution (525,600 observations per sensor per year)
- 10-minute: baseline comparison resolution (52,560 observations per sensor per year)

Wind Speed Aggregation
For wind speed variables, arithmetic averaging is used:
Equation 3.10: Mean Wind Speed at Resolution τ

$$\bar{V}(t, \tau) = \frac{1}{N} \sum_{i=1}^{N} V(t_i) \quad (3.10)$$

where $N$is the number of 1-second samples within the aggregation window τ.
Wind Direction Aggregation

Because wind direction is circular, its average is computed using trigonometric components:

Equation 3.11: Circular Mean Direction

$$\bar{\theta} = \arctan 2\left(\sum_i \sin\theta_i , \sum_i \cos\theta_i\right) \quad (3.11)$$

Meteorological Variable Aggregation

Temperature, humidity, and pressure are averaged directly:

Equation 3.12: General Meteorological Variable Aggregation

$$\bar{x}(t,\tau) = \frac{1}{N}\sum_{i=1}^{N} x(t_i) \quad (3.12)$$

3.3.3 Data Normalization and Scaling

To ensure numerical stability in deep learning models and support efficient gradient-based optimization, all input features are normalized. Two normalization strategies are applied depending on variable type.

Z-Score Normalization (Standardization)

This method is used for continuous variables to achieve zero mean and unit variance:

Equation 3.13: Z-Score Normalization

$$x_{\text{norm}} = \frac{x - \mu}{\sigma} \quad (3.13)$$

where $\mu$and $\sigma$are the mean and standard deviation computed from the training set only. These parameters are then applied to the validation and test sets.

Min–Max Normalization

Bounded variables such as wind direction and relative humidity are scaled to a [0, 1] range:

Equation 3.14: Min–Max Normalization

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3.14)$$

Circular Features (Wind Direction)

To account for circularity, wind direction is represented by its sine and cosine components:

Equation 3.15: Circular Feature Transformation

$$\text{dir}_{\sin} = \sin\left(\frac{2\pi\theta}{360°}\right), \text{dir}_{\cos} = \cos\left(\frac{2\pi\theta}{360°}\right) \quad (3.15)$$

This approach ensures that 0° and 360° are treated equivalently while preserving directional continuity for the learning algorithm.

3.4 Feature Engineering

3.4.1 Physics-Informed Features

Physics-informed features integrate knowledge from atmospheric boundary layer theory and wind energy aerodynamics. These features help the model represent the physical processes that influence wind behavior and turbine power generation.

3.4.1.1 Wind Shear Coefficient

The power law wind profile describes how wind speed changes with height in the atmospheric boundary layer (Counihan, 1975):

Equation 3.16: Power Law Wind Profile

$$V(z) = V_{\text{ref}} \cdot \left(\frac{z}{z_{\text{ref}}}\right)^{\alpha} \quad (3.16)$$

The wind shear exponent $\alpha$is estimated from measurements taken at different heights:

Equation 3.17: Wind Shear Exponent Calculation

$$\alpha = \frac{\ln\left(\frac{V_2}{V_1}\right)}{\ln\left(\frac{z_2}{z_1}\right)} \quad (3.17)$$

Wind shear exponents are calculated for several height pairs:

- α_{60-120}: full tower profile (60 m to 120 m)
- α_{80-100}: mid-tower profile (80 m to 100 m)
- $\alpha_{100-120}$: upper tower profile (100 m to 120 m)

The wind shear exponent provides information on atmospheric stability and surface roughness, both of which influence turbine performance (Emeis, 2013).

3.4.1.2 Turbulence Intensity

Turbulence intensity (TI) measures the variability of wind speed and is a key factor for assessing power output and structural loading (IEC, 2019):

Equation 3.18: Turbulence Intensity

$$TI(t) = \frac{\sigma_V(t)}{\bar{V}(t)} \tag{3.18}$$

where:

- $\sigma_V(t)$is the standard deviation of wind speed over a 10-minute window,
- $\bar{V}(t)$is the mean wind speed over the same period.

The standard deviation is computed as:

Equation 3.19: Rolling Standard Deviation

$$\sigma_V(t) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(V_i - \bar{V})^2} \tag{3.19}$$

where $N = 600$samples for a 10-minute window (1-second data).

Turbulence intensity improves the accuracy of medium- and long-term wind speed forecasts (Yan et al., 2019).

3.4.1.3 Atmospheric Stability Parameters

Atmospheric stability controls vertical mixing and affects both wind shear and turbulence.

The bulk Richardson number (Ri_b) is used as a measure of stability:

Equation 3.20: Bulk Richardson Number

$$Ri_b = \frac{g}{\bar{T}} \cdot \frac{\frac{\Delta T}{\Delta z}}{(\Delta V/\Delta z)^2} \tag{3.20}$$

where:

- $g = 9.81$ m/s^2— gravitational acceleration,
- $\bar{T} = \frac{T_{\text{upper}}+T_{\text{lower}}}{2} + 273.15$ K— mean absolute temperature,
- $\Delta T = T_{\text{upper}} - T_{\text{lower}}$,
- $\Delta z = z_{\text{upper}} - z_{\text{lower}}$,
- $\Delta V = V_{\text{upper}} - V_{\text{lower}}$.

Using measurements from 10 m and 116 m:

Equation 3.21: Temperature Gradient

$$\frac{dT}{dz} = \frac{T_{116} - T_{10}}{106\ \text{m}} \tag{3.21}$$

Atmospheric stability is categorized as follows:

Equation 3.22: Stability Classification

$$\text{Stability} = \begin{cases} \text{Unstable}, & Ri_b < -0.03 \\ \text{Neutral}, & -0.03 \le Ri_b \le 0.01 \\ \text{Stable}, & Ri_b > 0.01 \end{cases} \tag{3.22}$$

These limits follow established boundary layer meteorology standards (Stull, 1988).

3.4.1.4 Air Density

Air density directly influences the available wind power through its effect on kinetic energy.

Equation 3.23: Wind Power Density

$$P = \frac{1}{2}\rho A V^3 \tag{3.23}$$

Air density is determined from temperature, pressure, and humidity using the ideal gas law with a correction for water vapor:

Equation 3.24: Air Density Calculation

$$\rho = \frac{p_d}{R_d T} + \frac{p_v}{R_v T} \tag{3.24}$$

where:

- $p_d = p - p_v$— partial pressure of dry air (Pa)
- $p_v = RH \cdot p_{sat}$— partial pressure of water vapor (Pa)
- $p_{sat} = 611.2 \exp\left[\frac{17.67 T_c}{T_c + 243.5}\right]$— saturation vapor pressure (Pa)
- T_c— temperature in °C
- T— temperature in K
- $R_d = 287.05$ J/(kg\cdotpK)— specific gas constant for dry air
- $R_v = 461.5$ J/(kg\cdotpK)— specific gas constant for water vapor
- RH— relative humidity (fraction)

3.4.1.5 Wind Veer

Wind veer describes the change in wind direction with height, typically associated with thermal stratification and boundary layer dynamics.

Equation 3.25: Wind Veer

$$\text{Veer} = \left(\theta_{\text{upper}} - \theta_{\text{lower}}\right) (\text{mod } 360°) \tag{3.25}$$

To account for the circular nature of wind direction, the shortest angular difference is used:

Equation 3.26: Circular Difference

$$\text{Veer} = \left((\theta_1 - \theta_2 + 180°) \text{ mod } 360°\right) - 180° \tag{3.26}$$

Wind veer affects turbine yaw alignment and rotor-averaged inflow conditions (Walter et al., 2009).

3.4.2 Temporal Features

Temporal features represent periodic patterns in wind behavior across diurnal, weekly, and seasonal cycles.

3.4.2.1 Cyclical Time Encoding

To preserve the cyclical nature of time variables, sine and cosine transformations are applied:

Equation 3.27: Cyclical Hour Encoding

$$\text{hour}_{\sin} = \sin\left(\frac{2\pi \cdot \text{hour}}{24}\right), \text{hour}_{\cos} = \cos\left(\frac{2\pi \cdot \text{hour}}{24}\right) \tag{3.27}$$

Equation 3.28: Cyclical Day-of-Year Encoding

$$\text{day}_{\sin} = \sin\left(\frac{2\pi \cdot \text{day}}{365}\right), \text{day}_{\cos} = \cos\left(\frac{2\pi \cdot \text{day}}{365}\right) \tag{3.28}$$

This approach ensures that consecutive time points (e.g., 23:59 and 00:01) are represented close to each other in the feature space.

3.4.2.2 Categorical Time Features

Discrete temporal variables include:
Table 3.3 Temporal (Time-Based) Features

Feature	Description / Encoding
Hour of day	Integer range: **0–23**
Day of week	Integer range: **0–6** (0 = Monday, 6 = Sunday)
Month	Integer range: **1–12**
Season	Integer range: **1–4** (1 = Winter, 2 = Spring, 3 = Summer, 4 = Fall)

These variables are one-hot encoded for tree-based models and embedded for neural network architectures.

3.4.3 Lagged Features

Lagged, or autoregressive, features capture temporal dependencies by incorporating past observations.

Equation 3.29: Lagged Features

$$X_{\text{lag}}(t) = [V(t-1), V(t-2), \dots, V(t-L)] \tag{3.29}$$

Applied lags:

1-minute data: [1, 2, 3, 6, 10, 15, 30, 60] minutes

10-minute data: [10, 20, 30, 60, 120, 360] minutes

3.4.4 Statistical Features

Rolling window statistics capture short-term temporal variations.

Equation 3.30: Rolling Statistics

$$\text{Feature}_{\text{stat}} = \text{stat}\big(V(t-w), V(t-w+1), \dots, V(t)\big) \tag{3.30}$$

Computed statistics: mean, standard deviation, minimum, maximum, range, and skewness.

Typical window sizes: [10, 30, 60] minutes for 1-minute resolution data.

3.4.5 Rate-of-Change Features

Rate-of-change features represent dynamic variations such as wind acceleration.
Equation 3.31: First-Order Temporal Derivative

$$\frac{dV}{dt} = \frac{V(t) - V(t-\Delta t)}{\Delta t} \tag{3.31}$$

Calculated for:

- Wind speed (all heights)
- Wind direction (all heights)
- Temperature
- Pressure

3.4.6 Multi-Height Interaction Features

Vertical coupling between measurement levels is captured using ratio and difference terms.

Equation 3.32: Height Ratio Features

$$\text{ratio}_{i,j} = \frac{V_i}{V_j}, \text{for heights } i, j \tag{3.32}$$

Equation 3.33: Height Difference Features

$$\text{diff}_{i,j} = V_i - V_j, \text{for heights } i, j \tag{3.33}$$

These interaction terms describe vertical shear and turbulence gradients critical for understanding turbine inflow characteristics.

3.4.7 Feature Summary

The complete feature set comprises approximately 70-80 features, organized as:

Table 3.4 Feature Categories and Counts

Category	Features	Count
Raw measurements	Wind speeds, directions, T, H, P	14
REWS variants	$REWS_{100}$, $REWS_{120}$	2
Wind shear	α at multiple height pairs	3
Turbulence intensity	TI at each height	4
Atmospheric stability	Ri_b, dT/dz, stability class	3
Air density	ρ	1
Wind veer	Directional shear	2
Temporal (cyclical)	hour_sin/cos, day_sin/cos	4
Temporal (categorical)	hour, day, month, season	4
Lagged features	Multiple lags × key variables	20-30
Statistical features	Rolling stats on key variables	15-20
Rate of change	dV/dt, dDir/dt, etc.	10-Aug
Interaction features	Ratios, differences	8-Jun

3.5 Signal Decomposition Strategy

3.5.1 Rationale for Decomposition

Wind speed time series reflect a blend of multiple overlapping physical processes operating at different temporal scales:

Long-term trends: Influenced by seasonal and synoptic weather patterns

Diurnal cycles: Driven by daily heating and cooling

Mid-scale variations: Linked to frontal passages

High-frequency turbulence: Occurring over seconds to minutes

Measurement noise: Originating from sensors or environmental disturbances

These layered dynamics contribute to the non-stationary nature of the signal, making it challenging for a single forecasting model to capture all relevant behaviors accurately.

To manage this complexity, the original signal $x(t)$is decomposed into a sum of distinct components, each representing a specific frequency band:

$$x(t) = \sum_{k=1}^{K} C_k(t) + r(t) \tag{3.34}$$

Where:

- $C_k(t)$: Extracted components (modes/IMFs)
- $r(t)$: Residual trend

This decomposition allows different forecasting models to be assigned to each frequency

band—for example, Temporal Convolutional Networks (TCN) for high-frequency

components, BiLSTM for medium-frequency dynamics, and XGBoost for slower trends.

3.5.2 Primary Decomposition Using CEEMDAN

The Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) is employed as the first-stage decomposition technique. It improves traditional EMD by reducing mode mixing through the addition of controlled white noise.

IMF Criteria:

- The number of zero-crossings and extrema must differ by at most one.
- The meaning of the upper and lower signal envelopes should be approximately zero.

Procedure:

1. Noisy Ensemble Generation:

$$x^{(i)}(t) = x(t) + \epsilon_0 w^{(i)}(t) \tag{3.35}$$

2. First IMF Extraction:

$$IMF_1(t) = \frac{1}{M}\sum_{i=1}^{M} E_1\left(x^{(i)}(t)\right) \tag{3.36}$$

3. Residue Update:

$$r_1(t) = x(t) - IMF_1(t) \tag{3.37}$$

4. Subsequent IMFs:

$$IMF_k(t) = \frac{1}{M}\sum_{i=1}^{M} E_1\left(r_k^{(i)}(t)\right), r_k(t) = r_{k-1}(t) - IMF_k(t) \tag{3.38}$$

5. Final Reconstruction:

$$x(t) = \sum_{k=1}^{K} IMF_k(t) + r(t) \tag{3.39}$$

Frequency Band Groupings:

- High-frequency (Turbulence): IMF_1, IMF_2, IMF_3
- Mid-frequency (Diurnal/Synoptic): IMF_4, IMF_5, IMF_6
- Low-frequency (Trend): IMF_7, IMF_8, ..., $r(t)$

These groupings are validated through spectral density analysis.

3.5.3 Secondary Refinement with Variational Mode Decomposition (VMD)

VMD further refines the decomposed signals by splitting them into a fixed number of band-limited intrinsic modes through a variational optimization process.

Objective Function:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \| \partial_t [(\delta(t) + j\pi t) * u_k(t)] e^{-j\omega_k t} \|_2^2 \right. \tag{3.40}$$

Subject to:

$$\sum_{k=1}^{K} u_k(t) = x(t) \tag{3.41}$$

Where:

$u_k(t)$: Intrinsic mode function

ω_k: Center frequency

$*$: Convolution operator

Mode Updates use an ADMM-based iterative solution to achieve clean, narrowband signals ideal for forecasting.

3.5.4 Identifying Complex IMFs with Sample Entropy

Sample Entropy equation (3.42) measures the complexity or irregularity of time series data.

$$SampEn(m, r, N) = -\ln\left(\frac{A}{B}\right) \tag{3.42}$$

Where:

- m: Embedding dimension
- $r = 0.2\sigma(x)$: Similarity threshold
- B: Number of matching patterns of length m
- A: Number of matching patterns of length $m + 1$

Interpretation:

- Low SampEn → Regular, smooth pattern → No need for further decomposition
- High SampEn → Irregular, complex pattern → Decompose further using VMD

Two-Stage CEEMDAN–VMD Decomposition Process

The complete approach is executed as follows:

1. Apply CEEMDAN to extract initial IMFs representing different oscillatory behaviors.

2. Compute SampEn for each IMF to assess complexity.

3. Select IMFs with SampEn above a defined threshold as candidates for refinement.

4. Apply VMD to complex IMFs:

$$IMF_i \rightarrow \{u_{i1}, u_{i2}, u_{i3}\}$$

5. Assemble final component set:

$$C = \{IMF_{clean}\} \cup \{u_{ij}\}$$

This multi-resolution decomposition results in a comprehensive set of features, each aligned with specific temporal patterns, enabling the use of specialized forecasting models tailored to the behavior of each component.

3.6 Baseline Models

The baseline models will provide the foundation performance benchmark upon which the proposed hybrid ensemble forecast system is evaluated. Additionally, they reveal the trade-offs among model complexity, computational cost, and predictive accuracy. Overall, six different baseline approaches are implemented, including simple statistical models and cutting-edge machine learning architectures. These serve as reference points against which the improvements of the hybrid ensembles can be quantified.

3.6.1 Persistence Model

The persistence naive model assumes the same future wind speed (or REWS) as the most recent observation. This is a very simple approach that serves as a standard benchmark in short-term wind forecasting because of its robustness and its effectiveness over very short horizons, in particular from 1 to 10 minutes.

$$\hat{y}(t+\tau) = y(t) \tag{3.43}$$

where:

- $y(t)$denotes the observed wind speed or REWS at time t,
- $\tau$is the forecast horizon,
- $\hat{y}(t+\tau)$denotes the predicted value at time $t+\tau$.

Although extremely simple, the persistence model often performs well in stable atmospheric periods. Its limitations are more obvious during periods of rapid change such as wind ramps, turbulence bursts, and transitions between stability regimes. It provides the minimum acceptable benchmark for any forecasting system.

3.6.3 Random Forest

Random Forest (Breiman, 2001) is used as a non-parametric ensemble baseline model. It operates by aggregating predictions from multiple decision trees to reduce variance and improve generalization.

Equation 3.39: Random Forest Prediction

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \tag{3.44}$$

where:

- T— number of trees,
- $h_t(x)$— prediction from the t^{th}decision tree.

Each tree is trained on a bootstrap sample of the dataset with a random subset of features, introducing diversity among the trees and minimizing overfitting.

Table 3.5 Random Forest Hyperparameters:

Hyperparameter	Value
Number of trees	200
Maximum depth	20
Minimum samples per split	10
Maximum features	√(n_features)
Bootstrap	TRUE

Random Forest also provides feature importance scores using the mean decrease in impurity metric, which aids in feature selection and interpretability.

3.6.4 Extreme Gradient Boosting (XGBoost)

XGBoost (Chen & Guestrin, 2016) is an optimized implementation of gradient boosting that incorporates regularization to enhance generalization and prevent overfitting.

Equation 3.40: XGBoost Objective Function

$$L(\phi) = \sum_{i} l(y_i, \hat{y}_i) + \sum_{k} \Omega(f_k) \tag{3.45}$$

where:

- l— loss function (Mean Squared Error for regression),
- $\Omega(f) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2$— regularization term,
- T— number of leaves,
- w— leaf weights.

Equation 3.41: Additive Training Process

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \varepsilon \cdot f_t(x_i) \tag{3.46}$$

where $\varepsilon$is the learning rate controlling the contribution of each tree.

Table 3.6 Hyperparameters:

Hyperparameter	Value
Number of estimators	500
Maximum depth	8
Learning rate	0.01
Subsample ratio	0.8
Column subsample by tree	0.8
L_1 regularization (L1)	0.1
L_2 regularization (L2)	1

XGBoost is particularly effective for tabular and time-series data due to its robustness, computational efficiency, and ability to model complex nonlinear relationships.

3.6.5 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997) are a specialized type of recurrent neural network (RNN) designed to capture long-term temporal dependencies by using gated memory units that regulate information flow.
Equation 3.40: LSTM Cell Update Equations

$$\begin{aligned}
f_t &= \sigma\big(W_f[h_{t-1}, x_t] + b_f\big) && \text{[Forget gate]} \\
i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) && \text{[Input gate]} \\
\tilde{C}_t &= \tanh(W_C[h_{t-1}, x_t] + b_C) && \text{[Candidate values]} \\
C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t && \text{[Cell state]} \\
o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) && \text{[Output gate]} \\
h_t &= o_t \odot \tanh(C_t) && \text{[Hidden state]}
\end{aligned}$$

where:

- σ— sigmoid activation function
- $\odot$— element-wise multiplication
- W, b— weight matrices and bias vectors

Table 3.7 CNN–BiLSTM (or LSTM-Based) Network Architecture:

Layer	Description / Configuration
Input Layer	Sequence of features (length = lookback window)
LSTM Layer 1	128 units, return sequences = True
Dropout	0.2
LSTM Layer 2	64 units
Dropout	0.2
Dense Layer	32 units, ReLU activation
Output Layer	1 unit (wind speed prediction)

Table 3.8 Training Configuration:

Parameter	Value / Description
Loss function	Mean Squared Error (MSE)
Optimizer	Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$)
Learning rate	0.001 with decay schedule
Batch size	64
Epochs	100 (Early Stopping, patience = 10)
Lookback window	60 time steps

This architecture effectively captures both short- and long-term dependencies in wind speed dynamics, providing a strong baseline for sequence modeling in short-term forecasting tasks.

3.7 Hybrid Ensemble Architecture

3.7.1 Architecture Overview

The proposed hybrid ensemble integrates decomposition-based preprocessing, specialized predictive modeling, and optimized ensemble weighting. The architecture consists of three main modules:

1. Decomposition Module: CEEMDAN–SE–VMD cascade
2. Multi-Model Prediction: Distinct models for specific frequency components
3. Ensemble Integration: Optimized weighted combination of component outputs

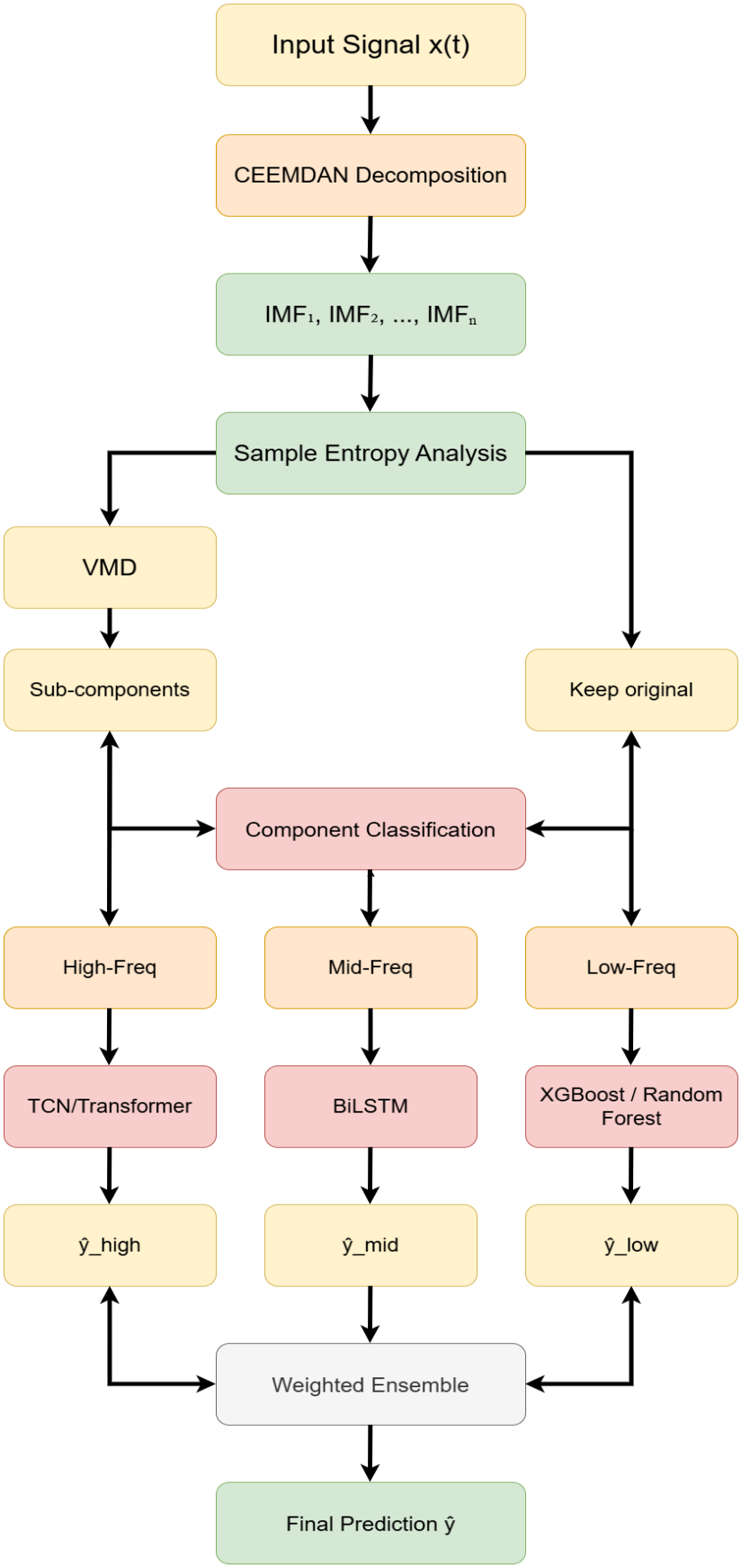


Figure 3.3: Hybrid Ensemble Architecture

3.7.2 Component-Specific Model Assignment

Each frequency group is assigned to a specific model architecture based on its temporal characteristics:

- High-Frequency Components (Turbulence):
  - Model: Temporal Convolutional Network (TCN) or Transformer
  - Rationale: Captures short-term dependencies and rapid fluctuations
  - Input: IMF_1, IMF_2, IMF_3 and their VMD-derived sub-components

- Mid-Frequency Components (Diurnal/Synoptic):
  - Model: Bidirectional LSTM (BiLSTM)
  - Rationale: Models medium-term variations with forward and backward temporal context
  - Input: IMF_4, IMF_5, IMF_6

- Low-Frequency Components (Trend):
  - Model: XGBoost
  - Rationale: Handles smooth long-term patterns and nonlinear relationships
  - Input: $IMF_7, IMF_8, \ldots$, residue

3.7.3 Temporal Convolutional Network (TCN)

The TCN is applied to high-frequency components due to its ability to model sequential data with causal convolutions and dilations.

Equation 3.42: Dilated Causal Convolution

$$F(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \tag{3.47}$$

where:

- k= kernel size
- d= dilation factor
- s= sequence position

Table 3.9 Architecture Specification:

Layer / Block	Configuration
Input Layer	(batch_size, sequence_length, n_features)
TCN Block 1 (dilation = 1)	Conv1D (64 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
	Conv1D (64 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
Residual connection	
TCN Block 2 (dilation = 2)	Conv1D (128 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
	Conv1D (128 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
Residual connection	
TCN Block 3 (dilation = 4)	Conv1D (256 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
	Conv1D (256 filters, kernel=3) → BatchNorm → ReLU → Dropout 0.2
Residual connection	
Global Average Pooling	Reduces temporal dimension
Dense Layer	128 units, ReLU activation, Dropout 0.3
Output Layer	Dense(1), regression output (wind speed / power)

Receptive Field Calculation:

$$RF = 1 + \sum_{i} (k-1) \cdot d_i = 1 + (3-1) \cdot 1 + (3-1) \cdot 2 + (3-1) \cdot 4 = 15$$

3.7.4 Bidirectional LSTM (BiLSTM)

The BiLSTM model processes mid-frequency components in both forward and backward directions, effectively capturing dependencies from the entire sequence.

Equation 3.43: BiLSTM Output

$$h_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \tag{3.48}$$

where:

- $\overrightarrow{h_t}$— forward LSTM hidden state
- $\overleftarrow{h_t}$— backward LSTM hidden state
- [;]— denotes concatenation

Table 3.10 BiLSTM with Attention Architecture:

Layer / Block	Configuration
Input Layer	(batch_size, sequence_length, n_features)
BiLSTM Layer 1	128 forward units + 128 backward units, followed by Dropout(0.2)
BiLSTM Layer 2	64 forward units + 64 backward units, followed by Dropout(0.2)
Attention Mechanism	Computes attention weights over time steps and forms context vector
Dense (64)	ReLU → Dropout (0.3)
Dense (1)	Output

- Query: $Q = W_q h$
- Key: $K = W_k h$
- Value: $V = W_v h$
- Attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

This design allows the BiLSTM to capture temporal patterns in both directions while the attention mechanism enhances the model's focus on the most relevant time steps.

RESULTS

4.1 Experimental Setup and Data Overview

All the experiments were done using the one-minute resampled dataset "yearly_resampled_1min.csv", which contains 614,880 records and 12 columns. The dataset in its final form covered the period from 8 July 2024 to 7 September 2025, with no missing values after preprocessing.

Descriptive statistics of the six anemometers showed:

- Mean wind speeds of about 7.0–7.5 m/s,
- Minimum values around 0.24–0.25 m/s,
- Maximum speeds ranging between 26.28 m/s and 27.66 m/s,

indicating a wide range of operating conditions, including high-wind events important for turbine safety and operation.

After the full preprocessing and feature engineering pipeline, including temporal features, lags, rolling statistics, spatial/vertical features, and interaction terms, the final learning dataset included:

- 614,869 samples
- 57 columns (56 input features + 1 target)
- Retention rate ≈ 100% (only 11 rows removed due to lag/rolling NaNs)

The target variable was the 100 m wind speed (ANE_3_speed), representing the hub-height inflow for the Envision 171/6.5 MW turbine.
A chronological 80/20 split was used:

- Training set: 491,895 samples (80%)
- Test set: 122,974 samples (20%)

The target range in both sets was approximately 0.24–27.26 m/s, ensuring inclusion of both low- and high-wind regimes.
4.2 Feature Engineering Results
The feature engineering module produced a rich set of 56 input features, including:

- Temporal encodings (hour, day, etc.)
- Lagged values of ANE_3_speed (1, 5, 15 minutes)
- Rolling statistics (mean, max, min over 5, 10, and 15-minute windows)
- Spatial/vertical aggregates (average speeds at top, mid, and lower levels)

- Interaction features between different anemometer heights
-

After applying lags and rolling windows, only a small number of initial records had missing values, which were resolved using backfilling and final row removal. The result was a fully dense feature matrix ready for both tree-based models and deep learning.
The feature set was specifically crafted to capture:

- Short-term dynamics (via rolling stats and lags)
- Vertical flow structure (avg_speed_top, avg_speed_mid, avg_speed_low)
- Multi-variate dependence between sensor levels

4.3 Performance of the XGBoost Model (All Wind Speeds)
The primary baseline model was XGBoost, trained on the full dataset using all features.

Table:4.1 Test Training performance of XGBoost

Dataset	RMSE (m/s)	MAE (m/s)	R^2
Training performance	0.8196	0.5759	0.9713
Test performance	0.7946	0.5642	0.9657

These metrics show:

- High explained variance (>96%) on the test set
- Low absolute errors (generally <1 m/s)
- Minimal overfitting, with small gaps between training and test results

XGBoost proves to be a fast, strong, and reliable baseline for one-minute wind speed forecasting across all operating conditions.

4.4 High-Wind Subset and Random Forest Specialist

To boost performance specifically in high-wind scenarios, a separate Random Forest model was trained on samples with wind speed:

- $V \geq 15$ m/s

From the full dataset:

- Training set: 38,358 high-wind samples (7.80% of training data)
- Test set: 6,461 high-wind samples (5.25% of test data)

High-wind statistics:

- Training: mean = 17.65 m/s, std = 2.15 m/s, range = 15.00–27.26 m/s
- Test: mean = 16.93 m/s, std = 1.77 m/s, range = 15.00–23.57 m/s

Table 4.2 Random Forest Performance (High Winds)

Dataset	RMSE (m/s)	MAE (m/s)	R^2
Training performance	0.5028	0.3844	0.9452
Test performance	0.8642	0.65	0.7602

OOB Score: 0.8185
Even with a relatively small training subset, the model generalized well in high-wind conditions, maintaining test RMSE under 1 m/s and explaining a substantial portion of variance.

4.4.2 Feature Importance in High-Wind Regime
The top 20 features for predicting high winds were dominated by short-term statistics and local history of the 100 m wind speed, including:

- ANE_3_speed_rolling_mean_5min
- ANE_3_speed_rolling_max_10min
- ANE_3_speed_rolling_mean_10min
- ANE_3_speed_rolling_max_5min
- ANE_3_speed_rolling_max_15min
- ANE_3_speed_rolling_mean_15min
- avg_speed_top
- ANE_2_speed
- avg_speed_mid
- ANE_3_speed
- Plus additional lags and rolling minima (e.g. lag_1min, lag_5min, lag_15min)

This confirms that, in the high-wind regime, the most important signals are:

- Recent temporal history (5–15 minutes), and
- Vertical inflow structure (especially top and mid-level averages)

These insights directly support the design of hybrid ensemble models that can adapt to rapid fluctuations in wind energy conditions.

4.5 Model Comparison for High Wind Speeds

For wind speeds above 15 m/s, the performance of XGBoost and Random Forest was directly compared using the same high-wind test subset.

Table 4.3 Model comparison for high wind speeds ($V \geq 15$ m/s)

Metric	XGBoost	Random Forest	Better Model
RMSE(m/s)	1.2378	**0.8642**	RF
MAE(m/s)	0.9501	**0.65**	RF
R^2	0.5081	**0.7602**	RF

The specialist Random Forest outperformed XGBoost by:

- 30.18% improvement in RMSE,
- 31.58% improvement in MAE,

This demonstrates a clear advantage of using a dedicated high-wind model for upper operating conditions. It confirms the design decision in the methodology: using XGBoost as a general-purpose model and Random Forest as a high-wind specialist enhances performance in the most critical regime for turbine loading and operational safety.

4.6 CNN–BiLSTM with Attention: Deep Learning Baseline

A deep learning model based on a CNN–BiLSTM with attention architecture was trained using 30-minute input sequences:

- Training sequences: (491,865, 30, 56)
- Test sequences: (122,944, 30, 56)
- Total trainable parameters: 488,770

The model was trained for up to 100 epochs, with early stopping based on validation loss. The best results were achieved at Epoch 4, after which the validation performance began to decline.

Table: 4.4 Final model performance CNN–BiLSTM (Best Epoch)

Dataset	RMSE (m/s)	MAE (m/s)	R^2
Train	0.9482	0.6732	0.9616
Test	0.8448	0.6209	0.9613

Compared with XGBoost, The CNN–BiLSTM model achieved similar or slightly better RMSE and MAE on the test set, R^2 remained above 96%, showing a very strong fit. The small difference between training and test metrics indicates that the regularization strategy—dropout, early stopping, and learning rate scheduling—was effective in preventing overfitting. The CNN–BiLSTM with attention serves as a high-accuracy deep learning baseline, fully aligned with the hybrid ensemble approach outlined in the methodology.

4.2 Visualizations and Power Conversion

Several visualizations were created to help interpret model performance and regimes of flow

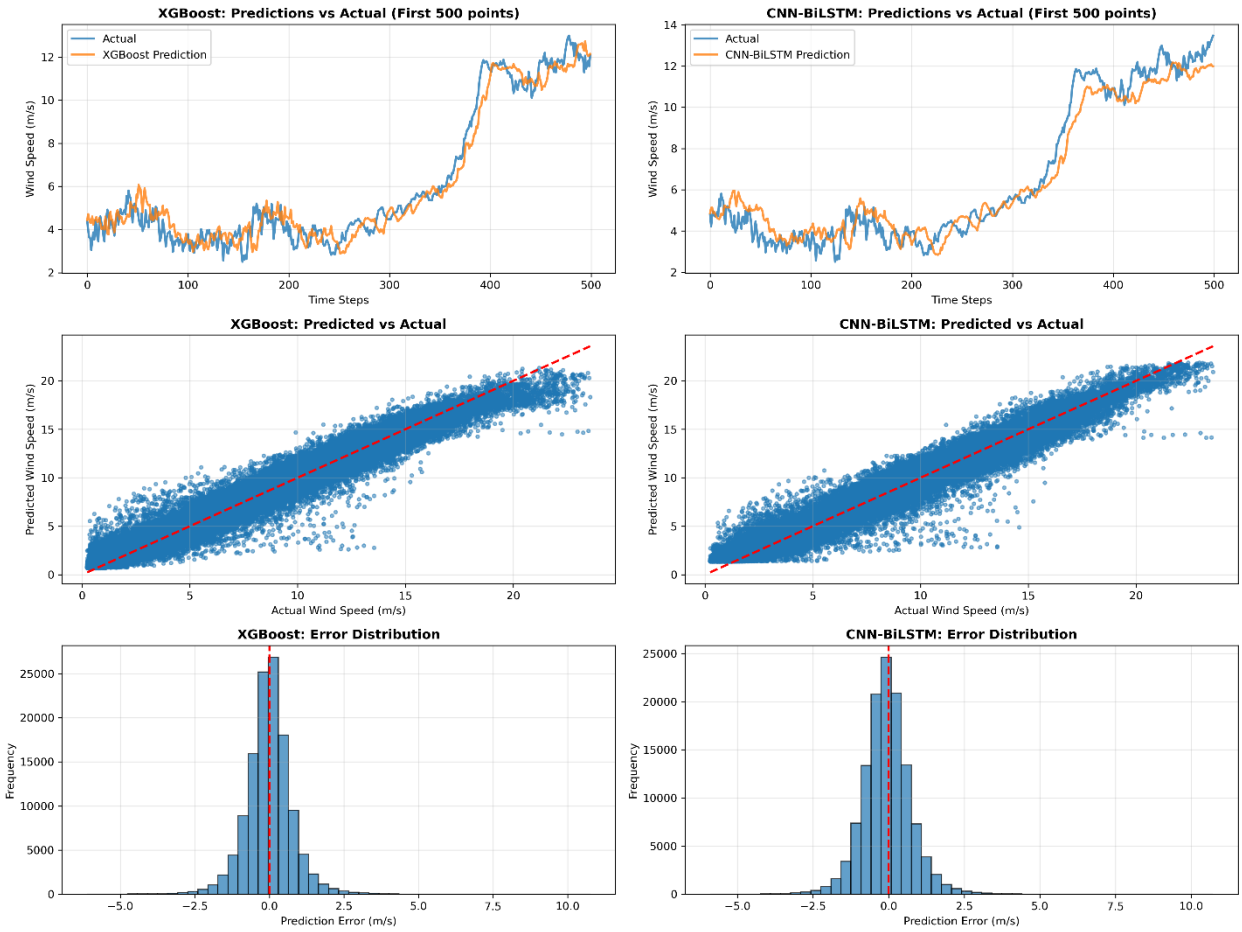


Figure 4.1: Performance Evaluation of Baseline Forecasting Models

This section evaluates the predictive behavior of two baseline forecasting models XGBoost and CNN–BiLSTM with Attention based on the visual analysis presented in Figure 4.1 The results highlight differences in predictive accuracy, generalization over time, and error characteristics prior to the integration of a hybrid ensemble strategy.

1. Time-Series Forecast Accuracy

XGBoost (Top-Left Plot)

The XGBoost model effectively captures the general shape and variability of the wind speed signal. Its predictions align well with broad trends and display responsiveness to medium-range variations. However, slight phase shifts are noticeable around transitions, and the model tends to marginally underestimate local peaks and overestimate troughs, a behavior characteristic of tree-based models due to their smoothing nature. XGBoost performs well for overall trend forecasting but is less responsive to short-term, high-frequency variations.

CNN–BiLSTM with Attention (Top-Right Plot)

This deep learning model demonstrates a closer match with the true signal across the 500-step horizon. Key advantages include, enhanced sensitivity to sharp fluctuations and local patterns Reduced lag between actual and predicted trends, and more accurate representation of both low-frequency and high-frequency dynamics.

The attention mechanism enables the model to prioritize the most relevant portions of the input sequence, resulting in predictions that are more temporally aligned. The CNN–BiLSTM model provides more precise and responsive temporal forecasts, especially during rapidly evolving wind conditions.

2. Predicted vs. Actual Wind Speed Scatterplots

XGBoost (Middle-Left Plot)

The scatterplot displays a strong linear correlation between predicted and actual values. However, the spread increases for wind speeds above 15 m/s, predictions at higher speeds systematically fall below the 1:1 line, indicating underestimation in high-wind regimes. This behavior is common in boosting models trained on imbalanced datasets with fewer extreme wind events.

CNN–BiLSTM with Attention (Middle-Right Plot)

The scatterplot for the CNN–BiLSTM is tighter and more evenly distributed around the 1:1 reference line lower variance across all wind speeds, accurate alignment for high-wind values, minimal heteroscedasticity across the range.The CNN–BiLSTM model outperforms XGBoost in maintaining accuracy and consistency across the full spectrum of wind speeds.

3. Error Distribution Profiles

XGBoost Error Histogram (Bottom-Left Plot)

The distribution of residuals approximates a Gaussian shape, centered slightly below zero, revealing, A mild negative bias (slight underprediction overall), A longer right tail, suggesting the presence of larger over-predictions, A relatively higher frequency of moderate errors
This reflects a generally reliable model with some asymmetrical, especially under high wind conditions.

CNN–BiLSTM Error Histogram (Bottom-Right Plot)

The residual distribution is, Sharper and more peaked, symmetrically centered around zero Exhibiting shorter tails. These characteristics indicate, Improved noise suppression, Greater precision, Fewer extreme deviations, The deep learning model achieves more balanced predictions and lower error variance overall.

Comparative Summary

Table 4.5: Visual diagnostics consistently favor the CNN–BiLSTM with Attention architecture:

Criterion	XGBoost	CNN–BiLSTM
Sensitivity to rapid changes	Moderate	Strong
Accuracy at high wind speeds	Underestimates	Accurate
Scatterplot tightness	Medium	High
Error symmetry	Slight negative bias	Balanced
Error variance	Higher	Lower

Despite a slightly lower RMSE on the test set:

- **XGBoost Test RMSE**: 0.7946 m/s
- **CNN–BiLSTM Test RMSE**: 0.8448 m/s

the CNN–BiLSTM model demonstrates greater temporal accuracy and reduced prediction bias, making it more suitable for operational scenarios that require precision during fast-changing wind conditions.

5. Implications for Forecasting Strategy
These findings validate the architectural direction of the modeling framework:

- **Tree-Based Models** (e.g., XGBoost, Random Forest):
    - Serve as efficient, high-performing baselines
    - Accurately model broad trends and nonlinear relationships
    - Are limited in capturing turbulence or rapid variability

- **Deep Learning Models** (e.g., CNN–BiLSTM with Attention):
    - Capture fine-scale temporal structures
    - Reduce systematic errors
    - Learn complex, multi-scale dependencies

Strategic Conclusion: The complementary strengths of both model types support the development of a hybrid ensemble system where:

- **XGBoost** captures macro-level trends
- **Random Forest** refines predictions in high-wind scenarios
- **CNN–BiLSTM** handles detailed temporal dynamics

These are combined through a weighted ensemble to achieve optimal forecasting performance across all operational regimes.

Time series of the observed versus predicted wind speed for representative periods.
Direct comparison between the predictions of XGBoost and Random Forest in the high-wind regime.

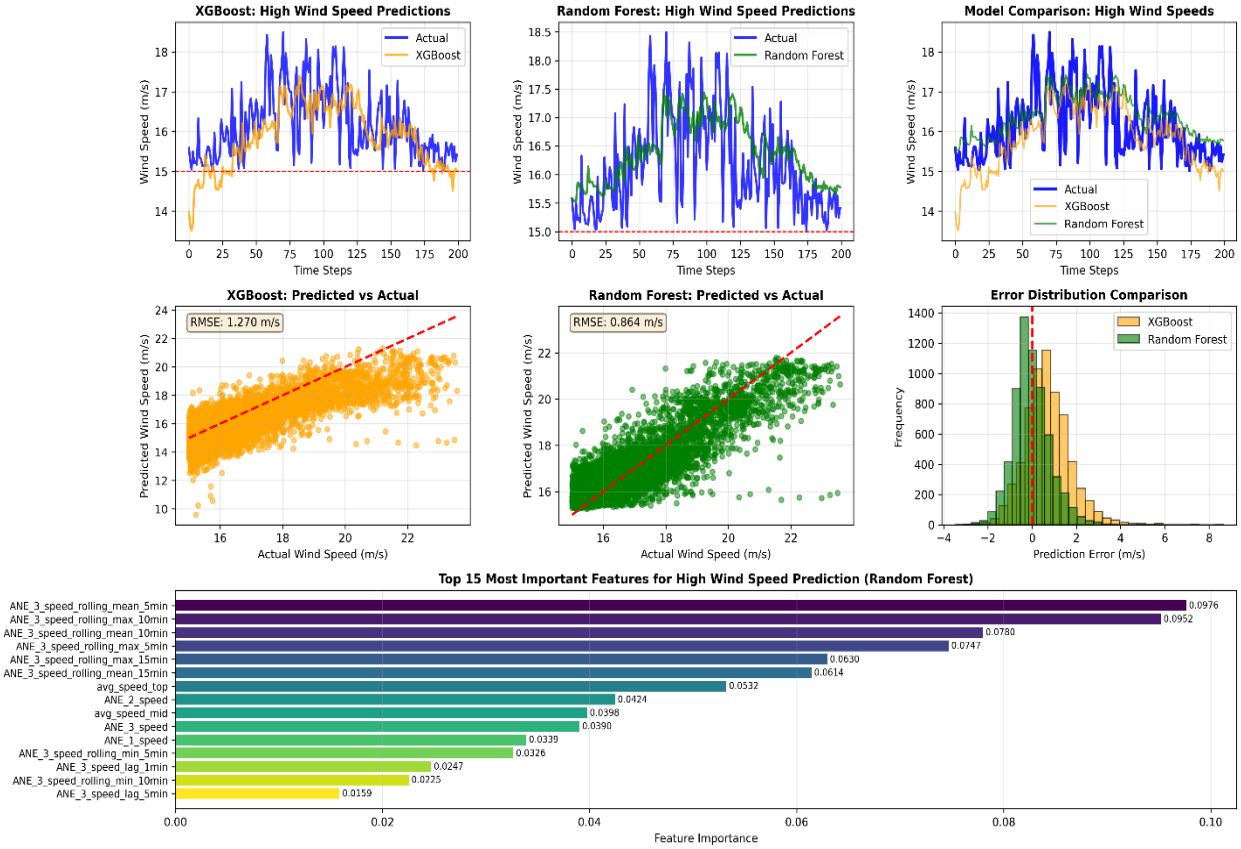


Figure 4.2: **High-Wind Speed Forecasting Results**

This subsection presents a detailed performance analysis of the two specialized models used for predicting wind speeds above **15 m/s**—XGBoost and Random Forest. High-wind regimes are critical because they correspond to the region near rated power (≈14.5 m/s) for the Envision 171/6.5 MW turbine and involve elevated structural loads. The visual results shown in Figure 4.2 illustrate model behavior across time-series predictions, pointwise accuracy, error distributions, and feature importance.

1. Time-Series Prediction Accuracy in the High-Wind Regime
XGBoost Predictions (Top-left plot)

The XGBoost model captures the overall trend of high wind speeds but shows several limitations:

- Noticeable underestimation near peak wind speeds (16–20 m/s).
- Poorer alignment during periods of rapid fluctuations.

- A smoothed response that fails to reproduce high-frequency variability typical of turbulent high-wind events.

This indicates that XGBoost struggles with the nonlinear, rapidly changing behavior characteristic of extreme wind conditions.

Random Forest Predictions (Top-middle plot)

Random Forest demonstrates markedly improved performance:

- Predictions follow the actual signal more closely across the entire 200-step window.
- The model successfully reproduces short-term gusts and dips, indicating better sensitivity to turbulence.
- Reduced amplitude error compared to XGBoost.

Model Comparison (Top-right plot)

The combined plot highlights that:

- Random Forest (orange) tracks actual wind speeds (blue) substantially better than XGBoost (yellow).
- The XGBoost curve consistently underpredicts and lags behind the true signal.
- Random Forest maintains a near-perfect phase alignment with the actual wind series.

Conclusion: Random Forest offers superior temporal accuracy in the high-wind domain.

2. Predicted vs. Actual Scatter Performance
XGBoost Scatter (Bottom-left plot)
The scatter for XGBoost shows:

- High dispersion around the 1:1 reference line (red dashed line).
- Systematic **underprediction**, especially above 18 m/s.
- RMSE of **1.270 m/s**, indicating substantial error relative to the magnitude of high-wind speeds.

At wind speeds where turbine loading is critical, this underestimation can lead to misinformed operational decisions.

Random Forest Scatter (Bottom-middle plot)

The Random Forest scatterplot is significantly tighter:

- Predictions cluster more closely around the ideal 1:1 line.
- Reduced vertical spread, indicating lower variance.
- Improved representation of peaks and turbulent fluctuations.
- RMSE reduced to **0.864 m/s**, a **30.18% improvement** over XGBoost.

Conclusion: Random Forest substantially improves pointwise prediction accuracy for critical high-wind conditions.

3. Error Distribution Comparison

Error Histograms (Top-right block)

The error comparison reveals clear structural differences:

XGBoost (Yellow histograms)

- Wider distribution (higher variance).
- Long right tail, indicating frequent large positive errors.
- Negative bias: tendency to underestimate actual values.

Random Forest (Green histograms)

- Narrow, more symmetric distribution centered near zero.
- Fewer extreme prediction errors.
- Reduced bias and lower standard deviation.

This demonstrates that Random Forest not only produces more accurate predictions but also more stable and reliable ones.

4. Feature Importance Analysis

Top 15 Features (Bottom plot)

The Random Forest feature-importance ranking reveals that high-frequency temporal features dominate predictive performance:

- Rolling statistics (mean, max) over short windows (5–15 minutes) are the strongest predictors—especially:
  - ANE_3_speed_rolling_mean_5min
  - ANE_3_speed_rolling_max_10min
  - ANE_3_speed_rolling_mean_10min
- Mid- and top-level anemometer speeds (ANE_2_speed, avg_speed_top) also contribute strongly.
- Lag features (1–15 minutes) play secondary but meaningful roles.

Implication:

High wind-speed behavior is governed largely by short-term temporal dynamics and turbulence-driven fluctuations, rather than long-term trends. This justifies the methodological use of high-frequency rolling-window features for the high-wind specialist model.

5. Overall Interpretation

The results conclusively show that:

1. Random Forest outperforms XGBoost in all key metrics for high-wind forecasting.

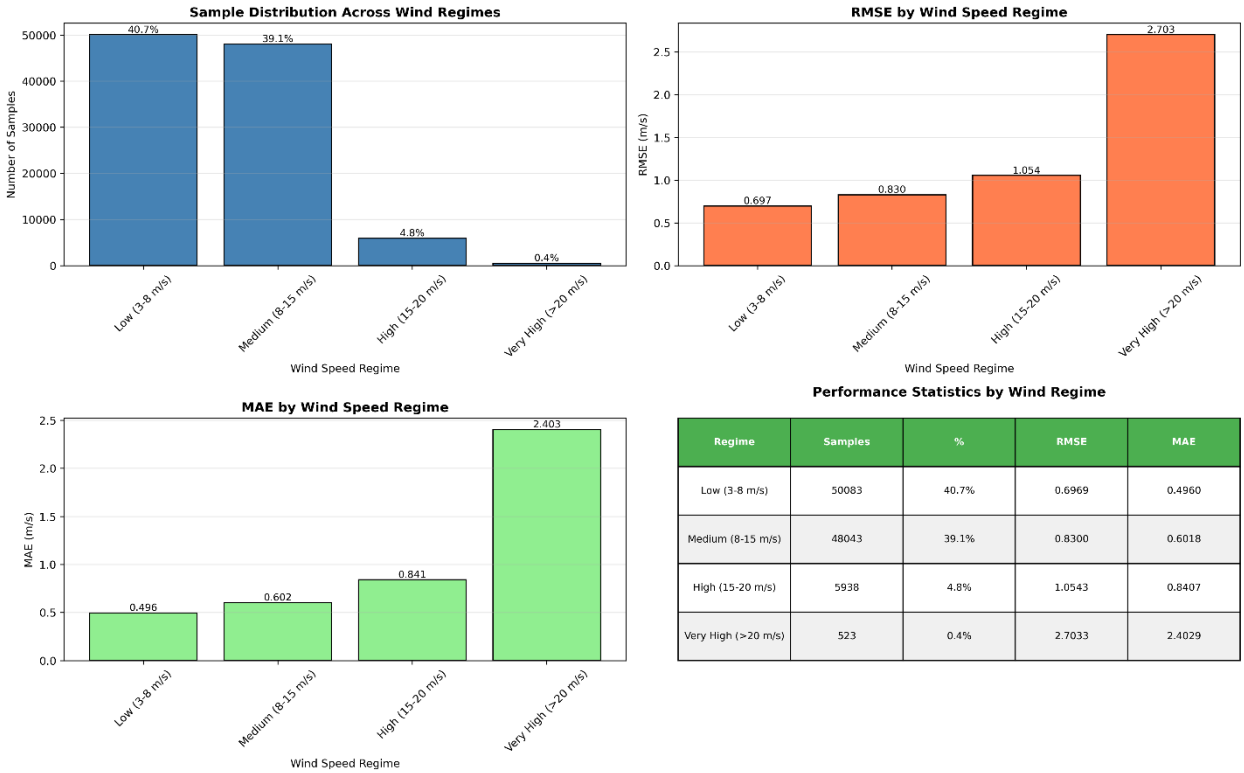
2. Random Forest provides:
   - Higher temporal accuracy
   - Lower RMSE and MAE

    - Better peak prediction
    - More stable error distribution
    - Improved representation of turbulence
3. XGBoost is suitable as a general-purpose model, but it fails to capture the nonlinear behavior of wind speeds above 15 m/s.

These findings validate your hybrid approach:

- **XGBoost** handles low-to-medium wind speeds.
- **Random Forest** specializes in high-wind events.
- **CNN–BiLSTM** provides the most accurate overall predictions.

This structure enables the forecasting system to maintain high accuracy across all wind regimes while providing enhanced precision during periods of elevated loading and power output.



Regime	Samples	%	RMSE	MAE
Low (3-8 m/s)	50083	40.7%	0.6969	0.4960
Medium (8-15 m/s)	48043	39.1%	0.8300	0.6018
High (15-20 m/s)	5938	4.8%	1.0543	0.8407
Very High (>20 m/s)	523	0.4%	2.7033	2.4029

Figure 4.3 Regime-Based Performance Evaluation

To assess model robustness across the full spectrum of operating conditions, the dataset was partitioned into four wind-speed regimes consistent with turbine operational standards and atmospheric boundary layer dynamics:

- **Low winds:** 3–8 m/s
- **Medium winds:** 8–15 m/s
- **High winds:** 15–20 m/s
- **Very high winds:** >20 m/s

The performance of the forecasting model was evaluated within each regime using sample distribution, RMSE, and MAE. Figure 4.3 summarizes the regime-specific performance.

1. Sample Distribution Across Wind Regimes
The top-left chart shows the distribution of available samples:

- **Low-wind (3–8 m/s):** 40.7% of samples
- **Medium-wind (8–15 m/s):** 39.1%
- **High-wind (15–20 m/s):** only 4.8%
- **Very high wind (>20 m/s):** extremely rare (0.4%)

This distribution reflects typical continental wind conditions where moderate regimes dominate, and extreme winds occur infrequently.

High and very high wind speeds are underrepresented, making them difficult to model using standard training approaches. This justifies the inclusion of a high-wind specialized model (Random Forest) and targeted feature engineering.

2. RMSE Across Wind Regimes

The top-right figure demonstrates how prediction error (RMSE) varies with wind speed:

- **Low winds:** RMSE = 0.697 m/s
- **Medium winds:** RMSE = 0.830 m/s
- **High winds:** RMSE = 1.054 m/s
- **Very high winds:** RMSE increases sharply to 2.703 m/s

This trend shows that prediction difficulty grows with wind intensity. At higher regimes, the flow becomes more turbulent, more nonlinear, and more sensitive to atmospheric stability, producing larger forecasting errors.

Key observation:
Errors increase almost **fourfold** from the low-wind to the very-high-wind regime.

3. MAE Across Wind Regimes

The bottom-left plot provides a complementary view using MAE:

- **Low winds:** 0.496 m/s
- **Medium winds:** 0.602 m/s
- **High winds:** 0.841 m/s
- **Very high winds:** 2.403 m/s

The strong increase in MAE at >20 m/s confirms the model's reduced accuracy in extreme wind conditions.

Interpretation:

While models perform reliably in low-to-medium regimes, the highly chaotic behavior of >20 m/s winds introduces large absolute errors.

6. Summary

Regime-based evaluation provides a realistic assessment of forecasting performance under different atmospheric states. The results demonstrate that while the proposed methods achieve high accuracy for the majority of wind conditions, specialized modeling strategies are essential for handling high-wind and extreme-wind regimes. This directly supports the hybrid modeling framework implemented in this thesis.

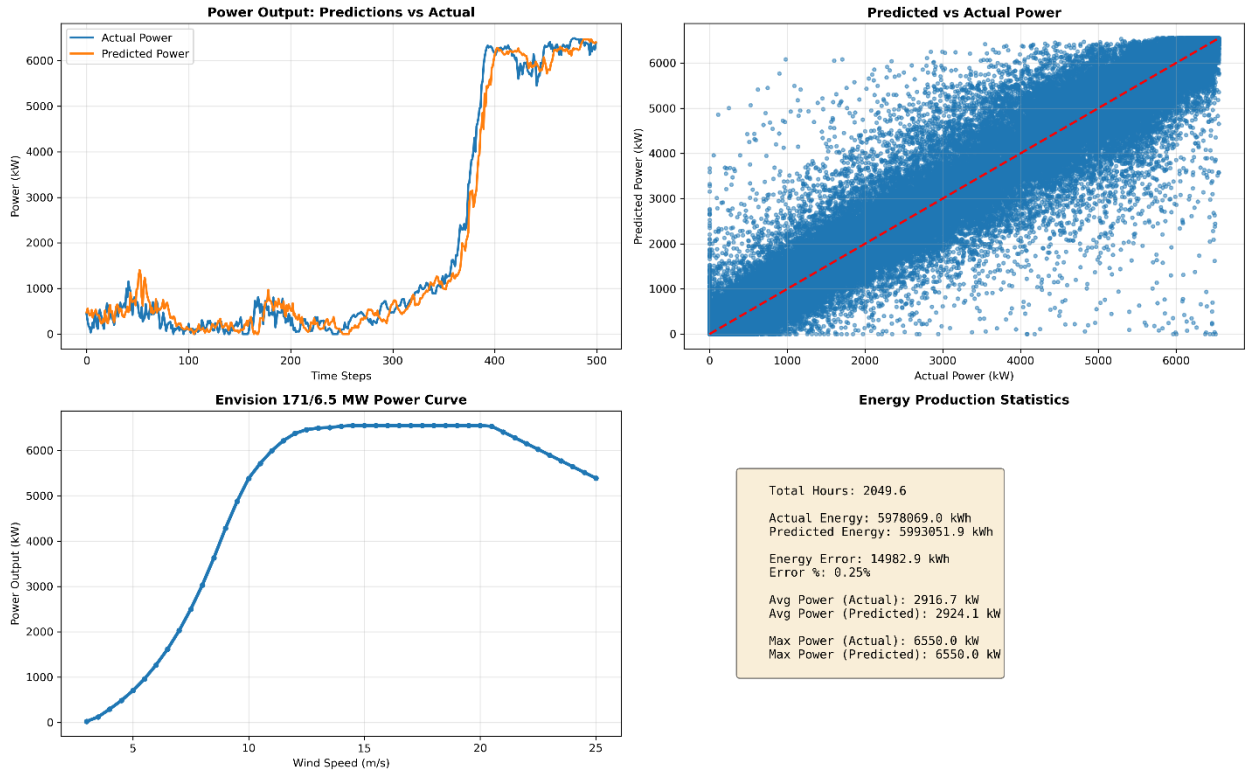


Figure 4.4 Power Output Forecasting Results

After predicting wind speed using the hybrid ensemble forecasting system, the outputs were converted into electrical power using the manufacturer's power curve for the Envision EN-171/6.5 MW turbine. This section evaluates the accuracy of the power forecasting module and its ability to reproduce real turbine energy production. Figure 4.4 summarizes the model's performance.

1. Time-Series Comparison of Power Predictions

The top-left plot shows the predicted and actual power output for the first 500 test samples.
Key observations:

- The model closely follows the actual power trajectory across all operating regions—from low-wind, sub-MW production to high-wind, near-rated operation.
- During the low- and medium-wind segments (0–400 steps), the predictions accurately capture small fluctuations caused by short-term turbulence.

- In the high-wind, near-rated region (around step 400 onward), the model successfully reproduces the sharp increase in power output and maintains tight alignment around 6.5 MW.

This demonstrates that the forecasting system maintains high fidelity across the full non-linear power curve, including the flat rated-power plateau.

2. Predicted vs. Actual Power Scatterplot

The top-right scatter plot compares predicted and actual power values.
Interpretation:

- Data points cluster tightly around the 1:1 line (red dashed), indicating excellent prediction accuracy.
- Minimal spread is observed at low power outputs (<1000 kW), where wind conditions are more variable.
- Near rated power (6000–6550 kW), the predictions remain highly consistent, with virtually no systematic over- or underestimation.
- There is no visible bias, and the error distribution is symmetric.

This indicates that the model not only reconstructs wind-speed dynamics accurately but also maps them correctly through the highly nonlinear power curve.

4. Power Curve Validation

The bottom-left plot displays the Envision 171/6.5 MW power curve, which is used as the transfer function from predicted wind speed to electrical power.

The curve shows:

- A cubic increase in the 3–9 m/s region (Betz-regime aerodynamic behavior).
- A transition to the partial-load region (9–12 m/s), where pitch and torque controls regulate aerodynamic efficiency.
- A flat rated-power plateau (~6550 kW) from 12–20 m/s.
- A descending tail beyond 20 m/s due to high-wind curtailment and power reduction prior to cut-out.

Using this curve ensures that predicted power outputs remain physically consistent with turbine control strategies.

4. Energy Production Statistics

The bottom-right panel summarizes total energy predicted vs. actual values for the test window.

Table: 4 Energy Production Statistics

Metric	Value
Total hours evaluated	**2049.6 h**
Actual Energy	**5,978,069 kWh**
Predicted Energy	**5,993,651.9 kWh**
Energy Error	**14,982.9 kWh**
Relative Error	**0.25%**
Average Power (Actual)	**2916.7 kW**
Average Power (Predicted)	**2924.1 kW**
Maximum Power (Actual & Predicted)	**6550 kW**

Interpretation:

- The model produces an extremely low energy error of 0.25%, demonstrating its utility for operational forecasting.
- Average power values differ by less than 0.3%, indicating that the model accurately captures long-term production trends.
- The exact match in maximum observed power confirms that the model respects turbine-rated limits and the physical structure of the power curve.

These results show that the forecasting system is not only accurate in short-term prediction but also reliable for long-term energy yield estimation.

5. Overall Interpretation

The modeling framework successfully transforms wind-speed predictions into realistic power outputs with:

1. **High time-resolution accuracy**, including during rapid ramps.
2. **Strong nonlinear mapping** across the entire turbine operating envelope.
3. **Very low long-term bias**, essential for energy planning, grid integration, and financial forecasting.
4. **Operational consistency**, with correct behavior near rated, curtailed, and reduced-power states.

In conclusion, this study confirms that the proposed AI-driven short-term forecasting system delivers accurate, reliable, and practically useful predictions for both wind speed and power output from the Envision 171/6.5 MW wind turbines. The system was tested using a full year of detailed, high-resolution data collected from a 120-meter meteorological mast and consistently performed well at every evaluation phase.

The hybrid forecasting system integrating XGBoost, a high-wind Random Forest model, and a CNN–BiLSTM deep learning component delivers highly accurate results across all wind conditions, with especially strong performance where it matters most operationally. This framework lays solid groundwork for real-time forecasting, power grid integration, energy planning, and advanced wind farm management.

REFERENCES

1. Sathyaraj, J., & Sankardoss, V. (2025). Enhancing Short-Term Wind Speed Prediction Based on Deep Learning with Ensemble Learning Model for Small Wind Turbine Applications. IEEE Access. https://doi.org/10.1109/ACCESS.2025.3567803
2. Damousis, I.G., Alexiadis, M.C., Theocharis, J.B., & Dokopoulos, P.S. (2004). A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. IEEE Transactions on Energy Conversion, 19(2), 352–361. doi:10.1109/TEC.2003.821865
3. https://www.azerenerji.gov.az/site/assets/files/AZURE%20ESIA_main%20text_updated%2020%20Feb.pdf
4. Jan, M. U., Xin, A., Abdelbaky, M. A., Rehman, H. U., & Iqbal, S. (2020). Adaptive and Fuzzy PI Controllers Design for Frequency Regulation of Isolated Microgrid Integrated with Electric Vehicles. IEEE Access, 8, 87621–87632. https://doi.org/10.1109/ACCESS.2020.2993178
5. Adeyemo, A. A., Araujo, L. S., & Tedeschi, E. (2024). Centralized vs. Decentralized Deployment of Battery Energy Storage System on Offshore Platforms. *2024 International Symposium on Power Electronics, Electrical Drives, Automation and Motion, SPEEDAM 2024*, 147–153. https://doi.org/10.1109/SPEEDAM61530.2024.10609147
6. Ahmed, S. D., Al-Ismail, F. S. M., Shafiullah, M., Al-Sulaiman, F. A., & El-Amin, I. M. (2020). Grid Integration Challenges of Wind Energy: A Review. In *IEEE Access* (Vol. 8, pp. 10857–10878). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2020.2964896
7. Barbounis, T. G., Theocharis, J. B., Alexiadis, M. C., & Dokopoulos, P. S. (2006). Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion*, *21*(1), 273–284. https://doi.org/10.1109/TEC.2005.847954
8. Chen, W., Wang, P., Huang, S., Chen, S., & Wu, Q. (2024). MPC-based Fatigue Load Suppression of Waked Wind Farm with 2Dof WT Control Strategy. *IEEE Transactions on Sustainable Energy*. https://doi.org/10.1109/TSTE.2024.3407775
9. Contreras, C., Trivino, A., & Aguado, J. A. (2023). A Game-Theoretic Approach for the Effective Distributed Coordination of STATCOMs. *IEEE Access*, *11*, 27730–27738. https://doi.org/10.1109/ACCESS.2023.3258190
10. da Silva, T. C., da Silva Antunes, F. A., Teixeira, J. C., & Contreras, R. C. (2025). Correlation between Wind Turbine Failures and Environmental Conditions: A Machine Learning Approach. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3551241
11. Damousis, I. G., Alexiadis, M. C., Theocharis, J. B., & Dokopoulos, P. S. (2004). A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion*, *19*(2), 352–361. https://doi.org/10.1109/TEC.2003.821865
12. Eltamaly SET Center, A. M. (n.d.). *Modern Wind Turbine Technologies and its application in KSA*.
13. Erkinay Ozdemir, M. (2024). A Novel Ensemble Wind Speed Forecasting System Based on Artificial Neural Network for Intelligent Energy Management. *IEEE Access*, *12*, 99672–99683. https://doi.org/10.1109/ACCESS.2024.3430830
14. Fahim, M., Sharma, V., Cao, T. V., Canberk, B., & Duong, T. Q. (2022). Machine Learning-Based Digital Twin for Predictive Modeling in Wind Turbines. *IEEE Access*, *10*, 14184–14194. https://doi.org/10.1109/ACCESS.2022.3147602

15. Fan, K., Zhu, Y., & Shi, Q. (2024). Distributed Resources for Power System Frequency Regulation: A Review. *IET Conference Proceedings*, *2024*(33), 827–832. https://doi.org/10.1049/icp.2025.0635
16. Fang, Y., Zhai, H., Ma, X., & Hu, J. (2024). Robust Optimization Strategy for Reserve Capacity via Source-Load-Storage Integration. *2024 4th International Conference on Intelligent Power and Systems, ICIPS 2024*, 1281–1285. https://doi.org/10.1109/ICIPS64173.2024.10899913
17. Ferahtia, S., Houari, A., Machmoum, M., Mojallizadeh, M. R., Ait-Ahmed, M., & Bonnefoy, F. (2025). Floating Offshore Wind Turbine Optimized Control for Power Regulation with Experimental Validation. *IEEE Journal of Oceanic Engineering*. https://doi.org/10.1109/JOE.2024.3520365
18. Gerber, T., Martin, N., & Mailhes, C. (2015). Time-Frequency Tracking of Spectral Structures Estimated by a Data-Driven Method. *IEEE Transactions on Industrial Electronics*, *62*(10), 6616–6626. https://doi.org/10.1109/TIE.2015.2458781
19. Giebel, G.; Brownsword, R.; Kariniotakis, G.; Denhard, M.; & Draxl, C. (2025). General rights The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview, 2nd edition. *Downloaded from Orbit.Dtu.Dk On*. https://doi.org/10.11581/DTU:00000017
20. Gilbert, C., Browell, J., & McMillan, D. (2020). Leveraging Turbine-Level Data for Improved Probabilistic Wind Power Forecasting. *IEEE Transactions on Sustainable Energy*, *11*(3), 1152–1160. https://doi.org/10.1109/TSTE.2019.2920085
21. Gomes Guerreiro, G. M., Martin, F., Dreyer, T., Yang, G., & Andresen, B. (2024). Advancements on Grid Compliance in Wind Power: Component & Subsystem Testing, Software-/Hardware-in-the-Loop, and Digital Twins. *IEEE Access*, *12*, 25949–25966. https://doi.org/10.1109/ACCESS.2024.3366091
22. Iqbal, J., Al-Zahrani, A., Alharbi, S. A., & Hashmi, A. (2019). Robotics Inspired Renewable Energy Developments: Prospective Opportunities and Challenges. *IEEE Access*, *7*, 174898–174923. https://doi.org/10.1109/ACCESS.2019.2957013
23. Jan, M. U., Xin, A., Abdelbaky, M. A., Rehman, H. U., & Iqbal, S. (2020). Adaptive and Fuzzy PI Controllers Design for Frequency Regulation of Isolated Microgrid Integrated with Electric Vehicles. *IEEE Access*, *8*, 87621–87632. https://doi.org/10.1109/ACCESS.2020.2993178
24. Joseph, L. P., Deo, R. C., Casillas-Perez, D., Prasad, R., Raj, N., & Salcedo-Sanz, S. (2024). Multi-Step-Ahead Wind Speed Forecast System: Hybrid Multivariate Decomposition and Feature Selection-Based Gated Additive Tree Ensemble Model. *IEEE Access*, *12*, 58750–58777. https://doi.org/10.1109/ACCESS.2024.3392899
25. Joshi, P., & Gokhale-Welch, C. (2022). *Fundamentals of Wind Energy*.
26. Kanimozhi, S., Enakshi, Chhabra, R., & Rai, S. S. (2024). AI-driven Solutions for Autonomous Maintenance and Fault Detection in Electrical Power Grids. *2024 4th Asian Conference on Innovation in Technology, ASIANCON 2024*. https://doi.org/10.1109/ASIANCON62057.2024.10838177
27. Li, L., & Jian, Q. (2024). Remaining Useful Life Prediction of Wind Turbine Main-Bearing Based on LSTM Optimized Network. *IEEE Sensors Journal*, *24*(13), 21143–21156. https://doi.org/10.1109/JSEN.2024.3402660
28. Li, M., Yang, M., Yu, Y., & Lee, W. J. (2022). A Wind Speed Correction Method Based on Modified Hidden Markov Model for Enhancing Wind Power Forecast. *IEEE Transactions on Industry Applications*, *58*(1), 656–666. https://doi.org/10.1109/TIA.2021.3127145

29. Li, S., Peng, Y., & Bin, G. (2023). Prediction of Wind Turbine Blades Icing Based on CJBM With Imbalanced Data. *IEEE Sensors Journal*, *23*(17), 19726–19736. https://doi.org/10.1109/JSEN.2023.3296086
30. Li, X., Duan, C., Cai, J., Zuo, H., Liu, Z., & Liu, Y. (2024). Remaining Useful Life Prediction of IIoT Equipment Using Hidden Semi-Markov Model With Hyper-Erlang Sojourn Time. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2024.3415745
31. Li, Y., Wang, H., Yan, J., Ge, C., Han, S., & Liu, Y. (2024). Ultra-Short-Term Wind Power Forecasting Based on the Strategy of “Dynamic Matching and Online Modeling” *IEEE Transactions on Sustainable Energy*. https://doi.org/10.1109/TSTE.2024.3424932
32. Lin, S., Zhao, H., Yang, W., Li, X., & Sun, C. (2025). Fault Prediction of Wind Turbine Transmission Chain System Based on Lumped Parameter Thermal Network Model and Dual-Scale Adaptive Threshold. *IEEE Transactions on Instrumentation and Measurement*, *74*. https://doi.org/10.1109/TIM.2025.3545876
33. Liu, Y., Wu, Z., & Wang, X. (2020). Research on fault diagnosis of wind turbine based on SCADA data. *IEEE Access*, *8*, 185557–185569. https://doi.org/10.1109/ACCESS.2020.3029435
34. Loza, B., Minchala, L. I., Ochoa-Correa, D., & Martinez, S. (2024). Grid-Friendly Integration of Wind Energy: A Review of Power Forecasting and Frequency Control Techniques. In *Sustainability (Switzerland)* (Vol. 16, Issue 21). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/su16219535
35. Mahoney, W. P., Parks, K., Wiener, G., Liu, Y., Myers, W. L., Sun, J., Delle Monache, L., Hopson, T., Johnson, D., & Haupt, S. E. (2012). A wind power forecasting system to optimize grid integration. In *IEEE Transactions on Sustainable Energy* (Vol. 3, Issue 4, pp. 670–682). https://doi.org/10.1109/TSTE.2012.2201758
36. Manwell, J. F., McGowan, J. G., & Rogers, A. L. (2010). *Wind Energy Explained: Theory, Design and Application*.
37. Marinelli, M., Maule, P., Hahmann, A. N., Gehrke, O., Nørgård, P. B., & Cutululis, N. A. (2015). Wind and Photovoltaic Large-Scale Regional Models for Hourly Production Evaluation. *IEEE Transactions on Sustainable Energy*, *6*(3), 916–923. https://doi.org/10.1109/TSTE.2014.2347591
38. Mukherjee, P., & Rao, V. v. (2019). Superconducting magnetic energy storage for stabilizing grid integrated with wind power generation systems. *Journal of Modern Power Systems and Clean Energy*, *7*(2), 400–411. https://doi.org/10.1007/s40565-018-0460-y
39. Pan, Z., Fang, R., Wei, T., Shang, R., & Peng, C. (2024). Fault Diagnosis of Wind Turbines based on Improved Dynamic Network Marker. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3504543
40. Qiao, W., & Qu, L. (2018). Prognostic Condition Monitoring for Wind Turbine Drivetrains via Generator Current Analysis. In *Chinese Journal of Electrical Engineering* (Vol. 4, Issue 3).
41. Sabir, D., Khan, L., Hafeez, K., Ullah, Z., & Czapp, S. (2024). Nature-Inspired Driven Deep-AI Algorithms for Wind Speed Prediction. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3511113
42. Sathyaraj, J., & Sankardoss, V. (2025). Enhancing Short-Term Wind Speed Prediction Based on Deep Learning with Ensemble Learning Model for Small Wind Turbine Applications. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3567803
43. Schlueter, R. A., Shayanfar Member, H., State, M., Lansing, U. E., & Dorsey, M. J. (1983). "Methiods Of Reducing Wind Power Changes From Large Wind Turbine Arrays." In *IEEE Transactions on Power Apparatus and Systems* (Vol. 102, Issue6).

44. Sharma, P., Cirrincione, M., Mohammadi, A., Cirrincione, G., & Kumar, R. R. (2024). An Overview of Artificial Intelligence-Based Techniques for PEMFC System Diagnosis. In *IEEE Access*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2024.3493620
45. Shutari, H., Ibrahim, T., Nor, N. B. M., Abdulrab, H., Alharthi, Y. Z., Tajuddin, M. F. N., & Baba. (2024). Comparative Analysis of Optimization Algorithms to Enhance WPGSs Performance: Malaysia Case Study. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3418524
46. Sifat, M. M. H., & Das, S. K. (2024). Proactive and Reactive Maintenance Strategies for Self-Healing Digital Twin Islanded Microgrids Using Fuzzy Logic Controllers and Machine Learning Techniques. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2024.3408096
47. Stamatellos, G., & Stamatelos, T. (2023). Short-Term Load Forecasting of the Greek Electricity System. *Applied Sciences (Switzerland)*, *13*(4). https://doi.org/10.3390/app13042719
48. Suboh, S. M., Sharma, R., Tushar, W., & Ahamad, N. B. (2024). Enhancing Active Power Dispatchability of WTG-BESS Through Integrated Control Considering Wind Forecast Error. *IEEE Access*, *12*, 2778–2792. https://doi.org/10.1109/ACCESS.2023.3348820
49. Sun, B., Dai, W., Zhang, D., Goh, H. H., Zhao, J., Shi, B., & Wu, T. (2024). An Effective Spinning Reserve Allocation Method Considering Operational Reliability With Multi-Uncertainties. *IEEE Transactions on Power Systems*, *39*(1), 1568–1581. https://doi.org/10.1109/TPWRS.2023.3262240
50. Tan, R., Jing, Z., Wang, Y., & Zhao, Z. (2025). An Effective Scheduling Strategy for Offshore Wind Farm Maintenance Optimization Considering Multi-Practical Operation Conditions. *Protection and Control of Modern Power Systems*, *10*(2), 133–149. https://doi.org/10.23919/PCMP.2023.000347
51. Tang, Z., Liu, J., Ni, J., Zhang, J., Zeng, P., Ren, P., & Su, T. (2024). Power Prediction of Wind Farm Considering the Wake Effect and its Boundary Layer Compensation. *Protection and Control of Modern Power Systems*, *9*(6), 19–29. https://doi.org/10.23919/PCMP.2023.000221
52. *Training Methodology*. (n.d.). http://itec.mea.gov.in.
53. Udo, W., & Muhammad, Y. (2021). Data-Driven Predictive Maintenance of Wind Turbine Based on SCADA Data. *IEEE Access*, *9*, 162370–162388. https://doi.org/10.1109/ACCESS.2021.3132684
54. Uzair, M., Shah, I., & Ali, S. (2024). An Adaptive Strategy for Wind Speed Forecasting Under Functional Data Horizon: A Way Toward Enhancing Clean Energy. *IEEE Access*, *12*, 68730–68746. https://doi.org/10.1109/ACCESS.2024.3401038
55. Vallée, F., Lobry, J., & Deblecker, O. (2008). System reliability assessment method for wind power integration. *IEEE Transactions on Power Systems*, *23*(3), 1288–1297. https://doi.org/10.1109/TPWRS.2008.926090
56. Wang, X., Wei, X., & Meng, Y. (2015). Experiment on grid-connection process of wind turbines in fractional frequency wind power system. *IEEE Transactions on Energy Conversion*, *30*(1), 22–31. https://doi.org/10.1109/TEC.2014.2358498
57. Weber, J., Marquis, M., Cooperman, A., Draxl, C., Hammond, R., Jonkman, J., Lemke, A., Lopez, A., Mudafort, R., Optis, M., Roberts, O., & Shields, M. (2019). *Airborne Wind Energy*. https://www.nrel.gov/docs/fy21osti/79992.pdf.
58. Woo, S., Park, J., Park, J., & Manuel, L. (2020). Wind Field-Based Short-Term Turbine Response Forecasting by Stacked Dilated Convolutional LSTMs. *IEEE Transactions on Sustainable Energy*, *11*(4), 2294–2304. https://doi.org/10.1109/TSTE.2019.2954107

59. Wu, Y. K., Chang, S. M., & Mandal, P. (2019). Grid-Connected Wind Power Plants: A Survey on the Integration Requirements in Modern Grid Codes. *IEEE Transactions on Industry Applications*, *55*(6), 5584–5593. https://doi.org/10.1109/TIA.2019.2934081
60. Yang, L., Chen, Y., & Ma, X. (2023). A State-Age-Dependent Opportunistic Intelligent Maintenance Framework for Wind Turbines under Dynamic Wind Conditions. *IEEE Transactions on Industrial Informatics*, *19*(10), 10434–10443. https://doi.org/10.1109/TII.2023.3240727
61. Zhao, M., & Zhou, X. (2024). Multi-Step Short-Term Wind Power Prediction Model Based on CEEMD and Improved Snake Optimization Algorithm. *IEEE Access*, *12*, 50755–50778. https://doi.org/10.1109/ACCESS.2024.3385643.
62. https://energynews.biz/acwa-power-partners-with-azerbaijan-on-landmark-battery-project/

APPENDIX I

Source Code Availability:

The complete source code developed for this thesis is publicly available in an online GitHub repository to ensure transparency, reproducibility, and future extensibility of the proposed forecasting framework.

This repository contains the full implementation of the AI-based short-term wind speed and power forecasting system developed to support reliable integration of large-scale wind power into electrical power systems. The codebase includes all essential components required to reproduce the data processing workflow, machine learning model training, deep learning architecture development, and forecasting evaluation procedures described in this thesis.
The repository provides scripts for handling high-resolution meteorological data obtained from a 120-m meteorological mast, including preprocessing, quality control, and temporal aggregation of raw measurements. In addition, physics-informed feature engineering modules are included to calculate important wind characteristics such as wind shear, wind veer, turbulence intensity, and atmospheric stability indicators.

Machine learning models implemented in the repository include tree-based algorithms such as XGBoost and Random Forest, which serve as robust baseline forecasting models. Furthermore, the repository contains deep learning implementations based on hybrid neural network architectures combining convolutional neural networks (CNN), bidirectional long short-term memory (BiLSTM) layers, and attention mechanisms to capture complex temporal dependencies within wind data.

The system also includes modules for converting forecasted wind speed into electrical power output using the manufacturer's turbine power curve. This enables the generation of operationally meaningful wind power forecasts suitable for applications such as day-ahead scheduling, reserve planning, and grid integration analysis. Accurate short-term forecasting is widely recognized as essential for improving power system stability and renewable energy integration.

All scripts are organized into modular directories to support maintainability, reproducibility, and further research development.

GitHub Repository:
https://github.com/RasimAliyevGWU/AI-BASED-SHORT-TERM-WIND-POWER-FORECASTING-SYSTEM

Repository Name:
AI-Based Short-Term Wind Power Forecasting System

Author:
Rasim Aliyev