

WIND TURBINE POWER GENERATION FORECASTING FOR INTRA DAY AND DAY-AHEAD

A THESIS

submitted by

**KAIVALYA KULKARNI
(CH22M522)**

for the award of the degree

of

**MASTER OF TECHNOLOGY,
Industrial Artificial Intelligence**

Under the guidance of

Mr. Phanibhargava

Data scientist at
TATA Consultancy Services Ltd.

In collaboration with

INDIAN INSTITUTE OF TECHNOLOGY MADRAS



Nov 2024

THESIS CERTIFICATE

This is to certify that the project report titled ‘Wind Turbine Power Generation Forecasting for Intra Day and Day Ahead,’ submitted by me to ‘Indian Institute of Technology Madras’ for the fulfilment of the award of ‘Master of Technology’ is a bona fide record of the research work, done by him under the supervision of Mr. Phanibhargava.

This Project Report, in its entirety or segments, has not been previously submitted to any other Institute or University to obtain any degree or diploma.

Mr. Phanibhargava

Research Guide

Lead Data Scientist

TATA Consultancy Services

Place: Chennai

Date: 4th Nov 2024

ACKNOWLEDGMENTS

I deeply thank my mentor, Mr. Phanibhargava, for his invaluable guidance, encouragement, and expert advice throughout my MTech project. His insights and expertise were instrumental in completing this research, and his unwavering support has been a constant motivation.

I am immensely thankful to TATA Consultancy Services, my employer, for sponsoring the MTech course. This sponsorship has been a pivotal step in my academic and professional growth, allowing me to advance my education and skills. I am also grateful for the support and understanding of my colleagues at TATA Consultancy Services, which enabled me to effectively balance my work and study commitments. I acknowledge the indispensable contributions of the professors at IIT Madras, without whom this research would not have been possible. I sincerely thank all the professors who were directly or indirectly involved in the course.

Lastly, I extend my heartfelt thanks to my family and friends for their unwavering support and encouragement throughout this journey.

ABSTRACT

KEYWORDS: Renewable energy, time series forecasting, machine learning algorithms

Renewable energy, particularly wind power, is pivotal in achieving sustainable and clean energy sources. Effective wind turbine power generation forecasting is essential for optimizing grid operations, resource management, and energy market participation. This project aims to develop accurate and reliable wind turbine power generation forecasting models, addressing intraday and day-ahead prediction horizons.

Key objectives of this project are:

Data Acquisition: Gather real-world or generate synthetic wind turbine power generation data from reliable sources, considering factors like wind speed, wind direction, temperature, and historical power generation records. The importance of features will be based on correlation among variables. Further training will be considered by analyzing correlation only.

Intraday Forecasting: Implement state-of-the-art time series forecasting techniques, such as LSTM networks and other sequential models like encoder-decoder architectures and denoising techniques like wavelet transform, to provide short-term forecasts at sub-hourly intervals of 10 minutes.

Day-Ahead Forecasting: Develop advanced day-ahead forecasting models that utilize historical data, weather forecasts, and sequential-based deep learning algorithms to predict wind power generation for the next timesteps.

Evaluation and Validation: Employ appropriate evaluation metrics like MSE, RMSE, and MAPE, MDA to rigorously assess the performance of forecasting models, ensuring their reliability and effectiveness.

ABBREVIATIONS

- ❖ **LSTM** - Long Short-Term Memory
- ❖ **ANN** - Artificial Neural Network
- ❖ **RNN** - Recurrent Neural Networks
- ❖ **CNN** - Convolutional Neural Network
- ❖ **ARIMA** - Autoregressive Integrated Moving Average
- ❖ **SARIMA** - Seasonal Autoregressive Integrated Moving Average
- ❖ **EKF**- Extended Kalman Filter
- ❖ **ANFIS**- Adaptive Neuro Fuzzy Inference System
- ❖ **PSO**- Particle Swarm Optimization
- ❖ **MFO**- Moth Flame Optimization
- ❖ **ACF**- Autocorrelation Function
- ❖ **PACF**- Partial Auto Correlation Function
- ❖ **MLP** - Multi-Layer Perceptron
- ❖ **MSE** - Mean Squared Error
- ❖ **RMSE** - Root Mean Squared Error
- ❖ **MAE** - Mean Absolute Error
- ❖ **MAPE**- Mean Absolute Percentage Error
- ❖ **MDA**- Mean Directional Accuracy
- ❖ **ED**- Encoder Decoder
- ❖ **SE**- Simple Exponential
- ❖ **DE**- Double Exponential
- ❖ **FFT**- Fast Fourier Transform
- ❖ **WT**- Wavelet Transform

TABLE OF CONTENTS

ACKNOWLEDGMENT	3
ABSTRACT	4
ABBREVIATION	5
LIST OF FIGURES	8
LIST OF TABLES	10
CHAPTER 1	
INTRODUCTION	13
CHAPTER 2	
LITERATURE REVIEW	26
<u>2.1 Statistical algorithms for time series forecasting</u>	26
<u>2.2 Hybrid models implementation with Statistical methods</u>	27
<u>2.3 Machine Learning Models</u>	28
<u>2.4 Advanced Deep Learning Approaches Towards Forecasting</u>	29
<u>2.5 Sequence to Sequence Architectures</u>	30
<u>2.6 Wind Power/ Speed forecasting in colder regions</u>	32
<u>2.7 Smoothing Techniques to Denoise the data</u>	33
CHAPTER 3	
Problem definition and formulation	35
3.1 Problem Definition and preliminary idea.....	35
3.2 Problem of power loss in colder regions due to snow deposition over turbines.....	35
3.3 Problem of power loss in hot climatic regions due to dust deposition over turbines.....	37
3.4 Mathematical Formulation of SARIMA model.....	44
3.4 Stacked LSTM Algorithm formulation.....	45
3.5 Encoder-Decoder Architecture Formulation.....	47

CHAPTER 4.

Dataset Description And Methodology	49
4.1 Dataset Description	49
4.2 Proposed Approach	63

CHAPTER 5

Implementations And Results.....	69
5.1 Implementations of Proposed Methods	69
5.1.1 Power generation forecasting using SARIMAX.....	69
5.1.2 Power Generation Forecasting using Stacked LSTM Architecture....	72
5.1.3 Power Generation Forecasting using Encoder-Decoder Architecture..	74
5.2 Results	75
5.2.1 Results from Sarimax model.....	76
5.2.2 Results of Stacked LSTM Architecture	78
5.2.3 Results of Encoder Decoder Architecture	81
5.2.4 Implementation of Denoising Techniques	90
5.2.5 Results after introducing Smoothing techniques	97
A] Simple exponential smoothing	98
B] Double Exponential Smoothing.....	105
C] Fast Fourier Transform.....	114
D] Wavelet Transform.....	122
5.2.5 Overall Comparison Among All Smoothening Techniques	125
5.3 Results Performed on Dataset-2 based on Ireland Region.....	133
5.3.1 1 step ahead (10 min in future) Forecasting.....	134
5.3.2 3 step ahead (30 min in future) Forecasting.....	137
5.3.3 6 step ahead (60 min in future) Forecasting.....	140
5.4 Calculation of Potential Wind Power and Average power loss.....	144
5.5 Results Performed on Dataset-3 based on Kuwait (High-Temp Region).....	150
5.5.1 1 step ahead (10 min in future) Forecasting.....	154
5.5.2 3 step ahead (30 min in future) Forecasting.....	147
5.5.3 6 step ahead (60 min in future) Forecasting.....	158
5.6 Calculation of Potential Wind Power and Average power loss.....	162
5.7 Cross Dataset Experiments.....	169

CHAPTER 6

CONCLUSION AND FUTURE WORK 174

REFERENCES 180

LIST OF FIGURES

Fig 1.1 : Different renewable energy sources used to generate energy.....	18
Fig 1.2: Turbine blade sectional structure.....	24
Fig 1.3: Wind turbine blade sections.....	24
Fig 3.1: Snow distribution in Europe.....	40
Fig 3.2: Basic architecture of LSTM unit	47
Fig 3.3 : Simple Representation of encoder decoder architecture.....	50
Fig 4.1- Correlation Heatmap of wind turbine data features.....	52
Fig 4.2- Power Curve Relationship between Active Power and Wind Speed.....	53
Fig 4.3: Key feature definitions and characteristics of Dataset-2.....	54
Fig 4.4- Correlation Heatmap of wind turbine data features for dataset-2.....	55
Fig 4.5- Power Curve Relationship between Power and Wind Speed-dataset2	56
Fig 4.6 Time series plot of wind speed feature.....	57
Fig 4.7 Residual analysis of denoised data.....	58
Fig 4.8 Timeseries plot of Power feature.....	58
Fig 4.9 Timeseries plot of Temperature feature.....	59
Fig 4.10 Timeseries plot of Humidity feature.....	60
Fig 4.11- Correlation Heatmap of wind turbine data features for dataset-3.....	61
Fig 4.12- Power Curve relationship between Power and Wind Speed for dataset-3... .	62
Fig 4.13 Timeseries plot of wind speed feature.....	64
Fig 4.14: Timeseries plot of Power feature.....	64
Fig 4.15 Timeseries plot of Temperature feature.....	65
Fig 4.16 Timeseries plot of Humidity feature.....	66
Fig 4.17- Structure to show next time step prediction.....	67

Fig 4.18- Structure to show next time step prediction for multivariable targets.....	67
Fig 4.19– Data input for next time step prediction of impacting features.....	68
Fig. 4.20- Proposed structure of Methodology	69
Fig 4.21– Wind Gusts analysis flow chart.....	70
Fig 4.22– Workflow to handle Noisy Data.....	70
Fig 5.1- Trend. Seasonality and Residuals of data.....	71
Fig 5.2: ACF and PACF plots for non stationary series.....	72
Fig 5.3: ACF and PACF plots for differenced series.....	73
Fig 5.4: Model Residual Analysis.....	74
Fig 5.5 : Stacked LSTM model Summary.....	75
Fig 5.6 : Stacked LSTM model summary for multivariable	76
Fig 5.7: Formulated Architecture of Encoder Decoder algorithm	76
Fig 5.8: Encoder-Decoder architecture summary.....	77
Fig 5.9- Training and validation loss of stacked LSTM architecture.....	78
Fig 5.10: Forecasting Plots for SARIMAX model.....	79
Fig 5.11 : Week wise Plots showing forecasting of wind power for each month.....	83
Fig 5.12: Performance of stacked LSTM for multivariable 1 step ahead prediction..	85
Fig 5.13 : Week wise Plots showing forecasting of wind power and wind speed for each month using Encoder-Decoder architecture	85
Fig 5.14: Performance of Encoder Decoder for multivariable one step ahead prediction	86
Fig 5.15: Performance comparison of stacked LSTM and Encoder-Decoder architecture for one step ahead forecasting.....	87
Fig 5.16: Training and validation loss of Encoder Decoder architecture for 30min ahead forecasting.....	87

Fig 5.17: Performance Metrics results of wind power and wind speed 30 min ahead forecasting.....	88
Fig 5.18: Training and validation loss of Encoder Decoder architecture for 60min ahead forecasting.....	88
Fig 5.19: Performance Metrics results of wind power and wind speed 60 min ahead forecasting.....	93
Fig 5.20: Smoothing plot for simple exponential with different alpha values.....	93
Fig 5.21: Residual performance check for simple exponential smoothing.....	94
Fig 5.22: Residual performance check for Double exponential smoothing.....	95
Fig 5.23: Residual performance check for Holts' Winter smoothing.....	97
Fig 5.24: Residual performance check for Fast Fourier Transform.....	99
Fig 5.25: Residual performance check for Wavelet Transform.....	101
Fig 5.26: Weekly Pred Plots using SE+ED Model for 10Min Forecasting Window.	101
Fig 5.27: Model Performance for SE+ED Model-10 Min Forecasting Window.....	102
Fig 5.28: Performance comparison of ED and SE+ED architecture.....	103
Fig 5.29: Weekly Pred Plots using SE+ED Model for 30Min Forecasting Window.	104
Fig 5.30: Model Performance for SE+ED Model-30 Min Forecasting Window.....	104
Fig 5.31: Performance comparison of ED and SE+ED architecture 30Min Ahead Forecasting Window.....	106
Fig 5.32: Weekly Pred Plots using SE+ED Model for 60Min Forecasting Window.	106
Fig 5.33: Model Performance for SE+ED Model-60 Min Forecasting Window.....	108
Fig 5.34: Performance comparison of ED and SE+ED architecture 60Min Ahead Forecasting Window.....	108
Fig 5.35: Weekly Pred Plots using DE+ED Model for 10Min Forecasting Window.	109
Fig 5.36: Model Performance for DE+ED Model-10 Min Forecasting Window.....	110

Fig 5.37: Performance comparison of ED and DE+ED architecture 10Min Ahead Forecasting Window.....	111
Fig 5.38: Weekly Pred Plots using DE+ED Model for 30Min Forecasting Window.	112
Fig 5.39: Model Performance for DE+ED Model-30 Min Forecasting Window.....	113
Fig 5.40: Performance comparison of ED and DE+ED architecture 30Min Ahead Forecasting Window.....	114
Fig 5.41: Weekly Pred Plots using DE+ED Model for 60Min Forecasting Window.	115
Fig 5.42: Model Performance for DE+ED Model-60 Min Forecasting Window...	116
Fig 5.43: Performance comparison of ED and DE+ED architecture 60Min Ahead Forecasting Window.....	117
Fig 5.44 Cumulative Energy Plots.....	118
Fig 5.45: Weekly Pred Plots using FFT+ED Model for 10Min Forecast Window...	118
Fig 5.46: Model Performance for FFT+ED Model-10 Min Forecasting Window....	119
Fig 5.47: Performance comparison of ED and FFT+ED architecture 10Min Ahead Forecasting Window.....	120
Fig 5.48: Weekly Pred Plots using FFT+ED Model for 30Min Forecasting Window.....	121
Fig 5.49: Model Performance for FFT+ED Model-30 Min Forecasting Window....	122
Fig 5.50: Performance comparison of ED and FFT+ED architecture 30Min Ahead Forecasting Window.....	123
Fig 5.51: Weekly Pred Plots using FFT+ED Model for 60Min Forecasting Window	124
Fig 5.52: Model Performance for FFT+ED Model-60 Min Forecasting Window....	125
Fig 5.53: Performance comparison of ED and FFT+ED architecture 60Min Ahead Forecasting Window.....	126

Fig 5.54: Time Series Plots for Wind Speed and Active Power After Smoothing...	127
Fig 5.55: Weekly Pred Plots using WT+ED Model for 10Min Forecasting	
Window.....	128
Fig 5.56: Model Performance for WT+ED Model-10 Min Forecasting Window...	129
Fig 5.57: Performance comparison of ED and WT+ED architecture 10Min Ahead	
Forecasting Window.....	130
Fig 5.58: Weekly Pred Plots using WT+ED Model for 30Min Forecasting	
Window.....	131
Fig 5.59: Model Performance for WT+ED Model-30 Min Forecasting Window....	132
Fig 5.60: Performance comparison of ED and WT+ED architecture 30Min Ahead	
Forecasting Window.....	133
Fig 5.61: Weekly Pred Plots using WT+ED Model for 60Min Forecasting	
Window	134
Fig 5.62: Model Performance for WT+ED Model-60 Min Forecasting Window....	135
Fig 5.63: Performance comparison of ED and WT+ED architecture 60Min Ahead	
Forecasting Window.....	138
Fig 5.64: Overall Performance Comparison of All Smoothing Techniques.....	139
Fig 5.65: Weekly Prediction Plots using WT+ED Model for 10Min Forecasting	
Window for Dataest-2	142
Fig 5.66 : Performance metrics of ED + Wavelet Transform model for 1 step ahead	
(10 min ahead) forecasting for Dataset-2	142
Fig 5.67: Weekly Prediction Plots using WT+ED Model for 30Min Forecasting	
Window for Dataest-2	145
Fig 5.68: Performance metrics of ED + Wavelet Transform model for 3 step ahead	
(30 min ahead) forecasting	146

Fig 5.69: Weekly Prediction Plots using WT+ED Model for 60Min Forecasting Window for Dataest-2	148
Fig 5.70: Performance metrics of ED + Wavelet Transform model for 6 step ahead (60 min ahead) forecasting	149
Fig 5.71: Methodology to understand loss in summer and winter region.....	149
Fig 5.72 : Potential Power Vs Forecasted power generated in the summer period for 1 step ahead forecasting.....	150
Fig 5.73: Potential Power Vs Forecasted power generated in the winter period for 1 step ahead forecasting.....	150
Fig 5.74: Potential Power Vs Forecasted power generated in the summer period for 3 step ahead forecasting.....	151
Fig 5.75: Potential Power Vs Forecasted power generated in the winter period for 3 step ahead forecasting.....	151
Fig 5.76: Potential Power Vs Forecasted power generated in the summer period for 6 step ahead forecasting.....	155
Fig 5.77: Potential Power Vs Forecasted power generated in the winter period for 6 step ahead forecasting.....	156
Fig 5.78: Weekly Prediction Plots using ED Model for 10Min Forecasting Window for Dataest-3.....	146
Fig 5.79: Performance metrics of ED model for 1 step ahead (10 min ahead) forecasting.....	159
Fig 5.80: Weekly Prediction Plots using ED Model for 30Min Forecasting Window for Dataest-3.....	160
Fig 5.81: Performance metrics of ED model for 3 step ahead (30 min ahead) forecasting.....	163
Fig 5.82: Weekly Prediction Plots using ED Model for 60Min Forecasting Window for Dataest-3.....	164

Fig 5.83: Performance metrics of ED model for 6 step ahead (60 min ahead) forecasting.....	165
Fig 5.84 : Methodology to understand loss in summer and winter region.....	166
Fig 5.85: Potential Power Vs Forecasted power generated in summer period for 1 step ahead forecasting.....	166
Fig 5.86: Potential Power Vs Forecasted power generated in Winter period for 1 step ahead forecasting.....	167
Fig 5.87: Potential Power Vs Forecasted power generated in Summer period for 3 step ahead forecasting.....	167
Fig 5.88: Potential Power Vs Forecasted power generated in Winter period for 3 step ahead forecasting.....	168
Fig 5.89: Potential Power Vs Forecasted power generated in Summer period for 6 step ahead forecasting.....	168
Fig 5.90: Potential Power Vs Forecasted power generated in Winter period for 6 step ahead forecasting.....	169
Fig 5.91: Cross Dataset Experiment Results.....	171
Fig 5.92: Average accuracy of cross-dataset experiments.....	172
Fig 6.1: Average Improvement Over ED architecture using Smoothing Tech.....	174
Fig 6.2: Average Accuracy using WT+ED model.....	175
Fig 6.3: Average Accuracy using ED model for dataset-3.....	176

LIST OF TABLES

Table 3.1: Month wise loss of energy due ice deposition over turbine.....	40
Table 4.1: Key feature definitions and characteristics for Dataset-1.....	52
Table 4.2 Windspeed feature details	57
Table 4.3 Temperature feature details.....	59
Table 4.4: Key feature definitions and characteristics of Dataset-3.....	61
Table 4.5 Windspeed feature details for dataset-3.....	63
Table 5.1 : Performance metrics of forecasted results for each month.....	80
Table 5.2: Performance metrics of forecasted results for each month.....	82
Table 5.3 Dataset-2 training and testing classification.....	82
Table 5.4 Dataset-2 Characteristics of turbine.....	148
Table 5.5: Average loss calculations for dataset-2.....	152
Table 5.6: Technical characteristics of Wind turbine for dataset-3.....	165
Table 5.7: Average loss calculations for dataset-3.....	169
Table 5.8: Comparative Study of all Models.....	173

CHAPTER 1

INTRODUCTION

Wind power generation forecasting problems aim to predict/forecast an intraday or day ahead generated power value based on many features such as wind speed. Harnessing the power of the wind to produce electricity presents a viable and environmentally friendly solution to the challenges posed by conventional fossil fuels. As a result, wind energy has emerged as a critical player in the transition to a cleaner and more sustainable energy landscape.

1.1 Motivation:

With the increase in demand for energy, energy generation should also amplify. The world demands energy, but unfortunately, we have very few energy resources available. The demand for fossil fuels is on an upward trend, and the availability is very low. So, it is necessary to shift to renewable energy sources.

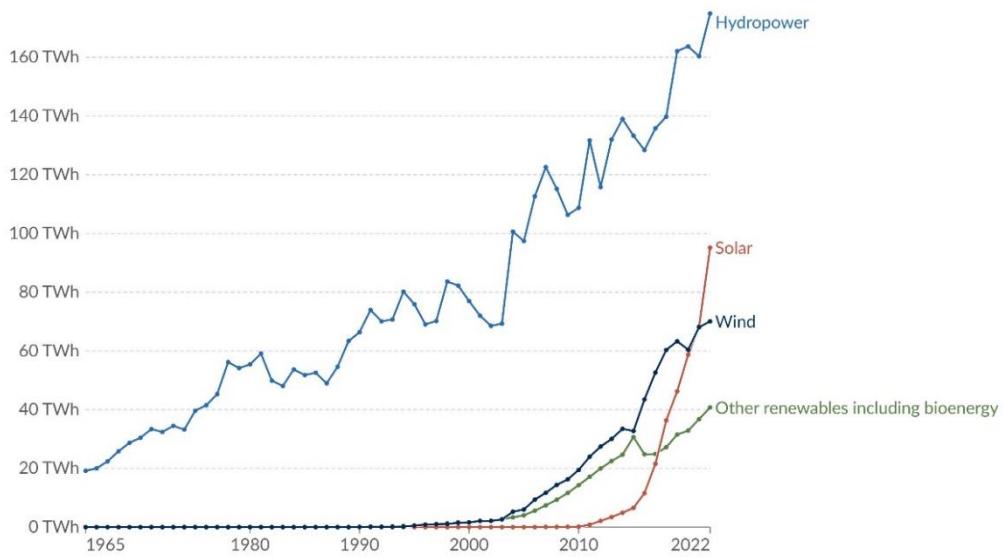
In 2023, the power requirement reached 1600 billion units. As per the Ministry of Renewable Energy Sources [\[1\]](#) , in 2023, around 13 GW of renewable energy was added, and India is 4th largest in Wind power capacity.

The below fig. shows the use of different renewable sources in India to generate energy

Modern renewable energy generation by source, India

OurWorld
In Data

Measured in terawatt-hours¹.



Data source: Ember - Yearly Electricity Data (2023); Ember - European Electricity Review (2022); Energy Institute - Statistical Review of World Energy (2023)
OurWorldInData.org/renewable-energy | CC BY

1. Watt-hour: A watt-hour is the energy delivered by one watt of power for one hour. Since one watt is equivalent to one Joule per second, a watt-hour is equivalent to 3600 Joules of energy. Metric prefixes are used for multiples of the unit, usually: - kilowatt-hours (kWh), or a thousand watt-hours. - Megawatt-hours (MWh), or a million watt-hours. - Gigawatt-hours (GWh), or a billion watt-hours. - Terawatt-hours (TWh), or a trillion watt-hours.

Fig 1.1 : Different renewable energy sources used to generate energy [2]

Denmark country generates 50% of its national electricity demand from wind, and it aims to generate 100% by 2035 [3]. Electricity produced by large central plants located close to available resources or industrial areas is transported to customers through transmission and distribution networks, but electricity is highly unsteady. We can use electricity at the same instant as it is produced and cannot be stored. So, the balance between power generation and consumption at all grid points is necessary.

Wind power generation is one of the finest renewable energy sources, and India is leading in 4th position now. The biggest challenge in wind power generation is the correct weather conditions. Wind power develops only when the wind blows, and electricity is necessary for all time. Hence, some ideas are considered to tackle the increased uncertainties in power systems.

- Energy storage: Until now, no such cheap batteries are available to store electricity, but energy transformation into another form can be done. Research into technologies like compressed air storage, flow batteries, and hydrogen is ongoing.
- Wind power forecasting: Power grid problem operations can be solved through power forecasting. Today, many electrical utilities widely use research on wind power forecasts for better cost reductions.[\[4-5\]](#)

Thus, it becomes necessary to forecast the power generation. Through forecasting, we can get relevant information on wind for particular locations. [\[6\]](#) and thus can motivate operators to optimize energy output.

Also, an accurate wind power forecast helps integrate grid operations, which can replace an expensive power backup resource and thus reduce the overall cost of operations.

1.2 Power Generation in colder climates:

Wind power varies as per geographical region. Wind speed variations in these areas affect the wind power.

Geographical locations are responsible for climatic changes and thus impact the wind energy industry [\[29\]](#). The study has focused on future climate projections to understand the impact of climate change on wind energy.

In colder regions, temperature changes significantly influence wind speed due to the interaction of air masses with different temperatures and densities. In colder regions, especially during winter, there can be sharp temperature gradients between different areas. These gradients create strong pressure differences, leading to higher wind speeds as air moves from high to low-pressure areas. Cold air is denser than warm air, which means it can exert more force as it moves. This can lead to higher wind speeds. During winter, colder regions often experience stronger winds due to the greater temperature contrasts.

But at the same time, snow deposition on turbine blades starts if the temperature falls below freezing point. Icing can lead to increased structural loads on turbines, reduce energy production, and negatively impact the economic viability of wind farm

investments. To address the issue of icing, de-icing systems that use various heating methods to remove ice from turbine blades have been introduced. Additionally, anti-icing systems using hydrophobic coatings to prevent ice accumulation are under development. Wind assessment companies are developing models to predict the severity of icing at potential wind farm sites [30].

The introduction should emphasize the need for further research and the development of standardized methods to accurately predict and mitigate the impact of icing on wind farms, ensuring the long-term sustainability and profitability of wind energy in cold climates.

1.2.1 Snow Formation:

Wind turbines can work in extreme colder conditions down to -4° Fahrenheit using technologies like cold resistant steels and heaters. But, proper maintenance of such equipment is also necessary.

During Texas' record cold weather in 2021, many electric systems failed, leading to the tragic loss of at least 246 lives. While some wind farms halted electricity production due to turbine icing or design limitations for cold temperatures, the issue was rooted in inadequate preparation, not a flaw in wind power itself. In fact, the state's natural gas infrastructure, which provides most of Texas' electricity, suffered an even greater shortfall during the freeze, as it too had not been weatherized for such extreme cold.

For snow to occur, moisture must exist in the atmosphere. While snowstorms significantly depend on temperature, it's not always the ground temperature that matters. Snow forms when the temperature in the atmosphere is at or below freezing (0°C or 32°F). If the ground temperature is also at or below freezing, snow will land intact. However, snow can still reach the ground when it's warmer at the surface under certain conditions. In such cases, as snowflakes enter the warmer air, they begin to melt slightly, triggering evaporative cooling, which cools the surrounding air and allows the snow to persist.

What is a snow?

- Snow is made up of compacted ice crystals. The condition of the snowpack affects qualities like color, Temperature, Water equivalent.
- Changes in weather impact the snowpack, altering its characteristics and the snow's overall properties.
- The character of the snow surface after a snowfall is shaped by both the form of the original snow crystals and the weather conditions at the time of the snowfall.
- For instance, strong winds during snowfall can break the snow crystals into smaller fragments, causing them to pack more densely.
- Once on the ground, snow may melt, evaporate, or persist for extended periods.
- If it remains, the texture, size, and shape of the snow grains will change over time, even if temperatures stay below freezing.
- The snow grains may melt, refreeze, and become compressed by additional snowfall.
- Over the winter season, the snowpack grows and develops a complex, layered structure made of various snow grains.
- These layers reflect both the weather and climate conditions at the deposition time, as well as ongoing changes within the snow cover.

How will it affect wind turbines?

- Icing typically happens when the surrounding temperature is slightly below freezing, and just cold enough for precipitation or fog to freeze on contact with the wind turbine.
- As moisture freezes, thickening layer starts to form, particularly on the leading edge of the blades while the turbine operates.
- When enough ice accumulates, the blade may quickly shed the ice layer, especially when the temperature rises slightly above freezing.

- Ice chunks can vary from several ounces to over 20 pounds.
- Given that blade tips can reach speeds of around 150 mph and are several hundred feet above ground, these ice chunks become potentially dangerous projectiles with a radius of nearly 1,000 feet.
- Usually, the damage is limited to a small area around the turbine base.
- Observing the damage firsthand, such as crushed turbine steps, damaged transformers, junction boxes, and nacelle fiberglass, highlights the destructive power of falling ice and gravity.

Thus proper functioning of ice removal methods is essential.

1.2.2 Ice Accretion and Types of Ice:

In cold climates, **ice accretion** on surfaces is a major challenge, especially for wind turbines. The ISO (International Organization for Standardization) categorizes ice into four primary types, which form under different conditions. These types are influenced by either precipitation or in-cloud icing [35]:

- **Glaze:** This type of ice forms as a smooth, clear layer when rain or drizzle freezes upon hitting a cold surface. Glaze ice is usually dense and heavy, which can lead to structural damage.
- **Rime (Hard/Soft):** Rime ice develops when supercooled water droplets in clouds come into contact with a surface and freeze. Hard rime is dense and durable, while soft rime is lighter and more delicate.
- **Wet Snow:** Wet snow adheres to surfaces and can freeze into a compact ice layer, depending on temperature and moisture conditions.
- **Hoar Frost:** Hoar frost forms directly from water vapor in the air, bypassing the liquid phase and crystallizing into ice. This type of ice is generally light and occurs in very cold environments.

Precipitation Icing vs. In-Cloud Icing

- **Precipitation Icing:** This type occurs when rain or snow freezes on contact with a surface, creating layers of ice. It is common during precipitation events where surface temperatures are at or below freezing.
- **In-Cloud Icing:** In-cloud icing happens when a site's elevation is within the cloud base, causing supercooled droplets and water vapor from the clouds to deposit on surfaces. This form of icing is particularly common in high-altitude locations, where many wind farms are situated.

In regions such as Sweden, **in-cloud icing** is prevalent because wind farms often sit at elevations that intersect with the cloud base [\[33\]](#).

Factors Affecting Ice Accretion:

Several key factors influence the amount and type of ice that accumulates on structures:

- **Air Temperature:** Lower temperatures generally increase the likelihood of ice formation, and the specific type of ice can vary depending on the temperature.
- **Liquid Water Content (LWC) of Clouds:** Higher LWC in clouds means more supercooled droplets are available to freeze upon contact, enhancing the potential for in-cloud icing.
- **Wind Speed:** Faster wind speeds lead to more rapid ice buildup as more droplets are driven onto the surface.
- **Structure Size and Shape:** The size and design of a structure play a significant role in how much ice it accumulates, with larger or more complex shapes typically gathering more ice.

Understanding these factors is essential for developing effective strategies to mitigate ice accretion and enhance the durability of wind turbines in cold climates.

1.3 Turbine blade structure:

As per [\[37\]](#), the main structural features of a typical wind turbine blade are illustrated in Fig 1.2:

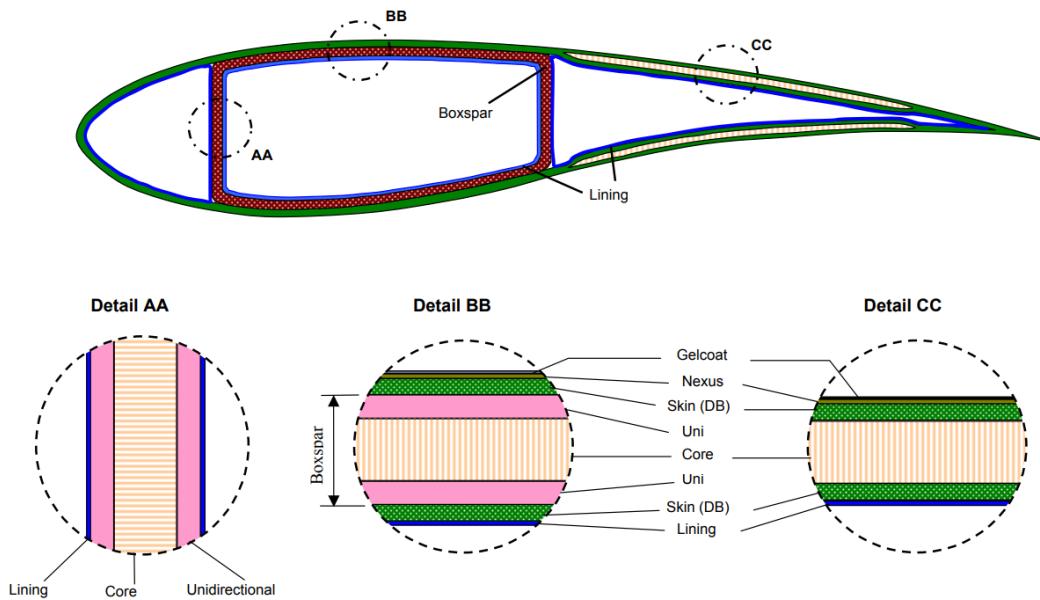


Fig 1.2: Turbine blade sectional structure

The wind turbine blade typically contains 2 shells. As per [37], there are two designs of wind turbine blade- one piece construction cross section and two piece construction cross section. These two pieces were applied by adhesives. Fig 1.3 shows the structure of one piece and two piece cross sectional blade:

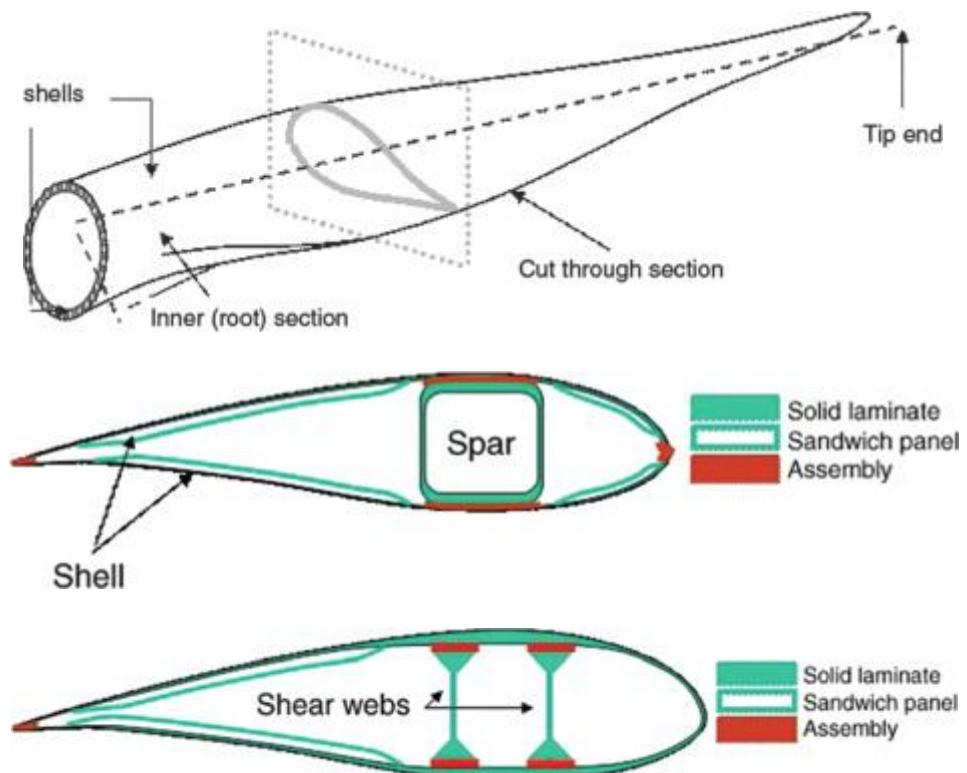


Fig 1.3: Wind turbine blade sections

1.4 Power Generation in Hot Temperature Climates:

Increased Power Loss Due to Heat: In high-temperature environments, wind turbines experience "power derating," where their output is intentionally reduced to prevent overheating. As the ambient temperature rises above a critical threshold (typically 40°C), turbines gradually lower their power output. At extreme temperatures (around 45°C), turbines may shut down entirely to protect internal components.

Impact on Air Density and Power Output: High temperatures reduce air density, which decreases the power that can be generated from a given wind speed. Lower air density directly affects the amount of kinetic energy available to the turbine, thus reducing power output in hot conditions.

Thermal Stress on Components: The heat can cause thermal expansion in turbine components, increasing wear and the likelihood of mechanical issues. Electronics and sensitive equipment inside the nacelle are particularly vulnerable to high temperatures, which can lead to more frequent maintenance and reduced turbine availability.

Cooling System Challenges: High temperatures demand efficient cooling systems to keep the internal components of the turbine within safe operational limits. In desert environments, traditional cooling methods may be less effective, leading to the development of convection-based or sand-resistant ventilation systems to improve reliability and minimize energy loss.

Maintenance and Operational Costs: High temperatures can increase the frequency and cost of maintenance. As turbines operate in high-temperature conditions, components may degrade faster, leading to more frequent replacements or repairs, which increases the operational costs of wind farms in desert regions.

1.4.1 Impact of dusts:

Dust storms, particularly during the summer months, contribute to significant dust accumulation on wind turbine surfaces, including blades and nacelle components. Accumulated dust clogs ventilation systems, further reducing turbine efficiency, especially during high-temperature periods. This effect, referred to as "dust-induced"

derating," is crucial to understanding overall turbine performance in dusty environments.

There are some of the techniques available to tackle the problem of hot climates in dessert conditions [\[45\]](#)

1. Heat and Sand-Proofing with Louvers

The nacelle, which houses critical components like the gearbox, generator, and controller, needs ventilation to prevent overheating, yet sand cannot be allowed to infiltrate. To achieve this balance, the team designed special louvers (angled slats) that act as filters.

Louvers are like plantation shutters that allow air to pass through but physically block sand particles from entering. Four louver sets were installed in strategic locations on the nacelle to enable natural airflow without introducing sand into the turbine. This design ensures air can circulate freely, cooling the interior components while keeping sand out even in strong sandstorms. Louvers are particularly effective because they redirect air without requiring additional power, leveraging the force of convection.

2. Utilizing Convection Cooling

Convection cooling relies on the natural movement of hot air rising and being replaced by cooler air. This is an efficient way to regulate temperature without active cooling systems like fans or air conditioning, which would consume additional energy and could be less reliable in desert conditions.

The turbine modifications utilize convection by allowing hot air to escape through louvered vents, keeping internal temperatures below the critical threshold of 55°C (131°F), which is essential for electronic components to operate. By focusing on natural cooling through convection, engineers eliminated the need for refrigeration or energy-consuming cooling systems, making the turbine better suited for remote and harsh environments where maintenance is challenging.

3. Infiltration Prevention with Sand-Proof Brushes

Sand grains can infiltrate even the smallest openings, damaging machinery over time. Small gaps between the nacelle and tower were sealed using sand-proof brushes.

Engineers added wiry brushes to fill these tiny gaps, creating a physical barrier against sand. These brushes block sand grains blown by strong desert winds from reaching sensitive lower components of the turbine, like the tower's door and connection points. The sand-proof brushes effectively keep the turbine's internal mechanisms sand-free without restricting the structure's flexibility.

4. Temperature-Resistant Lubrication

Turbine gearbox lubrication is vital for smooth operation. However, extreme temperatures can cause lubricants to lose viscosity, leading to friction and potential damage to the gears.

The turbines in Oman are designed to handle higher temperatures than standard turbines. The lubricants used in these turbines are specially formulated to maintain their properties at temperatures beyond 40°C (104°F), reducing the risk of wear and damage. This heat-resistant lubrication plays a key role in ensuring the turbine's mechanical components continue functioning smoothly under high thermal stress.

5. Enhanced Operating Range

Traditional turbines often have to reduce power output or shut down at temperatures above 35°C. The GE-modified turbine, however, is capable of running at full capacity up to 40°C and continues operating between 40°C and 45°C, extending its productive hours during the hottest parts of the day.

By modifying the cooling and sand-proofing aspects, the team allowed the turbine to operate within a broader temperature range. This extended operational threshold ensures the turbines in Oman's Dhofar wind farm can generate maximum power even during peak daytime temperatures, significantly boosting overall energy output in a region where high temperatures are the norm.

How thesis is solving this problem?

The challenge in front of us is in terms of empirical data. Wind power forecasting needs real-world weather and turbine feature data. Most of the windmill data is confidential; everyone cannot access it directly. Thus, this thesis is motivated to access the available public dataset and research this data.

Now the question arises: how much ahead of forecasting is needed and why? As we implement much ahead forecasting, the accuracy of forecasting may reduce. Long-term forecasting is necessary to plan power systems. Medium-term forecasting is required to control fuel supply and maintenance, while short-term forecasting is needed for day-to-day operations and energy management.

This thesis aims to predict/forecast an intraday or day-ahead generated power value based on impacting features such as wind speed and other turbine parameters. Harnessing the power of the wind to produce electricity presents a viable and environmentally friendly solution to the challenges posed by conventional fossil fuels. As a result, wind energy has emerged as a critical player in the transition to a cleaner and more sustainable energy landscape.

The utilization of wind power, however, has its intricacies and challenges. Wind energy generation is highly dependent on many factors, including wind speed. Moreover, the nonlinear nature of wind speed and its inherent variability introduces complexities that necessitate advanced forecasting models. Accurate and reliable wind power forecasts are indispensable for optimizing the integration of wind energy into the electricity grid, ensuring grid stability, and ultimately maximizing the efficiency of wind farms.

Also, in colder regions/hot climatic regions wind speed is highly impacted due to surrounding temperature. The regions where chances of snow fall are on the top, will be the interested problem of this thesis. To fulfil the demand for wind turbine stations, it is necessary to forecast power generation in colder climates as well as in hot weather climates.

From various studies mentioned in the literature survey, deep research has been done on numerous algorithms like statistical models such as ARIMA, ARMA, AR, and MA, as well as advanced models such as LSTM, SVM, RNN, etc. Also, research focuses on hybrid model implementation such as ANFIS, ARIMA+NNs, etc. The study was further extended by developing optimization models with the help of Kalman filters. However, Kalman filters can't perform well on nonlinear data, and thus, it is necessary to implement a method that can study nonlinear patterns in sequential data.

The denoising techniques can be implemented to reduce noise in the data before training the model.

Also, the use of attention models and transformers is well known in the NLP domain, but accurate study needs to develop in the field of time series forecasting.

The main objective of this project is to develop sequential architectures to test time series data to forecast accurate wind power results. Thus, the LSTM model network has been explored and built in this thesis to deal with nonlinearity and sequential data. Further, based on the accuracy of the network, improvement in the structure of the existing algorithm is made. Seq2seq architectures are studied for multistep and multivariable datasets. At last, different algorithm comparisons must be performed to decide the efficient model that gives optimized forecasting easier.

The respective research and research gaps have been mentioned in the upcoming literature survey. The complete problem definition and formulation are explained in Chapter 3, followed by dataset description and detailed methodology in Chapter 4. The results and conclusions are explained in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

Accurate wind power forecasting is essential for better grid planning operations and power plant efficiency. Thus, adequate research has been done in the past years. Research started from the basic use of statistical algorithms to advanced LLM-based architectures. It covers the challenges faced in forecasting, such as nonlinearity and nonstationarity, and how researchers have addressed these challenges through innovative methods. The review starts with traditional statistical algorithms like ARIMA and their limitations in handling nonlinear data. It then progresses to hybrid models that combine statistical methods with neural networks to improve forecasting accuracy. Data cleaning and feature analysis are important before applying any forecasting model.

Hybrid approaches, such as combining ARIMA with neural networks or optimization techniques like wavelet models, are discussed for enhanced forecasting results. Machine learning models like LASSO, KNN, and XGBoost are also explored for daily wind power prediction across different locations. Advanced deep learning approaches, particularly LSTM models, are highlighted for their ability to capture nonlinear and long-term dependencies in time series data.. The review also mentions the use of Kalman filters and marine predator optimizers in conjunction with forecasting models.

The discussion extends to sequence-to-sequence architectures, where LSTM networks are combined with 1D CNN layers to capture both temporal and long-term dependencies in data. Attention mechanisms are introduced to adjust the model's focus on relevant features dynamically, further improving forecasting accuracy.

The detailing of the exploration is mentioned in the subsequent sections below.

2.1 Statistical Algorithms For time series forecasting:

Traditional research aimed to implement statistical algorithms/methods like AR, MA, ARMA, and ARIMA in time series forecasting. Prediction methods from global to local scales and ensemble forecasting were also studied.(Aoife M. Foley et al, 2011) [7]. It was found that statistical methods gave significant results on linear data but

error reduction was not optimum on nonlinear time series data. Implementing Kalman filters was also one of the methods to converge the prediction results. But the use of Kalman filters on nonlinear data proved to be inefficient.

Advanced statistical methods that implement Fuzzy logic and neural networks (George Sideratos, Nikos D. Hatziaargyriou et al., 2020) [8] present a new statistical approach to time series forecasting. When fuzzy logic combines with a neural network model like ANN, weather forecasting results can be optimized. The structure of the algorithm is defined into three parts. Preliminary model, fuzzy logic model, and final wind power forecasting model. These neural networks, also known as radial basis neural networks, are designed to handle complex nonlinear problems.

Before applying any model it was observed that thorough data cleaning and feature analysis is essential (Shui Wang, 2022) [9]. The research emphasizes the significance of data cleaning by detecting and removing outliers. Outliers, including insufficient data and noisy data, can potentially disrupt the accuracy of prediction models. Feature reconfiguration is introduced as a means to enrich the original wind data. By mapping historical wind samples into a multidimensional sample space, the study reconfigures wind features to create a structured information map incorporating time series information. The paper highlights the advantages of deep learning models, particularly convolutional neural networks (CNN), for approximating nonlinear functions in wind power prediction.

The research mentions the concept of the power curve in wind farms, where data points close to the curve are considered ordinary, while deviations are identified as outliers. Outliers are detected using local density calculations. Considering the outcome, it becomes necessary to focus on the data processing part before moving towards the application of any algorithm. In the report, outliers are removed using inter-quartile methods, and missing value imputation is done by KNN imputer, which helped to keep the shape and pattern of the original data.

2.2 Hybrid models implementation with Statistical methods:

Further, to improve forecasting results, hybrid approaches were implemented ('David Barbosa et al., 2017) [10] by introducing neural network models along with the ARIMA model. Introducing multiple neural networks gave good results, but at the

same time, computational cost and training time were more. ARIMA and wavelet models are found to give the same results as that of the ARIMA model.

Another hybrid methodology that came into existence was the combination of ARIMA and ANN models. The use of ACF and PACF using the Box and Jenkins method (G. Peter Zhang, 2001) [\[11\]](#) to identify the order of the ARIMA model was described. ARIMA model deals with linearity, and ANN can treat nonlinear residuals from the ARIMA model. Still, more advanced hybrid techniques were needed to enhance the forecasting results.

To focus on optimization, some of the optimization techniques like wavelet NN model, MOMFO, and a combination of this optimization method with other NN were implemented (Jianzhou Wang, 2019) [\[12\]](#).

Long-term forecasting is necessary to plan power systems. Medium-term forecasting is required to control fuel supply and maintenance, while short-term forecasting is needed for day-to-day operations and energy management.

Sulagna Mahata et al. [\[13\]](#) performed their experiment considering the availability of wind speed in the Kutch region of Gujrat, India, where the average wind speed is 8 m/s. Also, air density in that region helps move rotor blades. Along with the altitude, the density of the air changes, and it was found that dense air exerts more pressure on rotor blades.

The author proposed an ARIMA model considering the AR component, which can capture linear dependencies between current and previous values; the error term can be handled by the MA component, which is a linear combination of all previous error terms, and then the I term indicates order until stationarity is achieved.

Parameters are further optimized using the Nelder-Mead method and Broyden Fletcher Goldfarb Shanno method. The seasonal forecasting is still a gap observed in the solution provided.

2.3 Machine Learning Models :

The research was further reversed to some machine learning techniques. Machine algorithms can observe the change in data trends and could be more efficient than traditional statistical approaches. (Halil Demolli, 2020) [\[14\]](#). Thus, a study on regression algorithms, including LASSO, KNN, XGBOOST, etc., was done to model daily wind

power. This research was done in various locations to observe the effectiveness of models. The study evaluates the performance of machine learning algorithms at different locations. Research could explore more advanced cross-validation techniques, such as spatial cross-validation, to assess how well the models generalize to diverse geographical areas.

2.4 Advanced Deep Learning Approaches Towards Forecasting:

Time series forecasting data is based on time instances and can be treated as sequential data. Based on previous timestamps, the target variable value is decided. Considering the fact, for long-term accuracy, the LSTM model is considered (Farah Shahid et al., 2021) [\[15\]](#). Previously, various machine learning models, such as support vector machines and Gaussian processes, were studied. These models can capture non linear patterns but require high computational resources. Further, to optimize LSTM predictions, a Genetic Algorithm is used. LSTM is used to capture nonlinear and nonstatic sequences. The genetic algorithm's performance is evaluated by comparing its predictions with actual values and considering other performance metrics like MSE, MAE, RMSE, etc. Still, the comparison of the performance of the GLSTM model with other models like ARIMA + GLSTM could be made to analyze comparative results.

The research further approached towards the use of Kalman filters. But as has already been observed, Kalman filters can optimize results onto linear data and can't perform on nonlinear or nonstationary data; time series forecasting results can never be improved with the use of Kalman filters. So, extended Kalman filters can be used along with hybrid models. ANN predictions are further passed through extended Kalman filters for forecasting (Shikha Singh et al , 2007) [\[16\]](#). Wind turbine power plants generate power, and wind speed is a key feature. It is observed that wind power is proportional to the cube of the wind speed. However, wind speed is not always constant, and variations in wind speed with elevation impact wind power generation. Multi-layer perceptron layers, which consist of multiple layers, undergo the training phase where initialization of best weights has been done to find the best network, which minimizes the forecasting error. EKF is used to train the weights of MLP networks in a nonlinear system. EKF is a recursive method in training the parameters. EKF training algorithm is highlighted for its strong learning capability and convergence properties.

Another optimization parameter introduced was the marine predator optimizer. (Mohammed A.A. Al-qaness, 2022) [17]. The marine predator optimizer is used along with the boosted ANFIS model using mutation operators. MPAmu is a new variant of MPA. It is introduced so that the algorithm cannot converge to local optima, which can eliminate the possibility of best optimization. MPAmu optimizes the developed ANFIS parameters to boost the configuration process. A comparison of traditional ANFIS algorithms with developed MPAmu + ANFIS has been done. It was found that the main objective of using these operators is to strengthen the ability of MPA to avoid the attractiveness of the local point. This will increase the convergence rate towards the optimal solution and improve the prediction quality based on the best configuration of ANFIS. In detail, the MPAmu is applied in the learning stage to determine the best weight values.

2.5 Sequence to Sequence Architectures:

Deep learning models are further extended with the use of networks that can take sequential data. Sequential data can be textual or time series data. Using LSTM network architectures can solve many sequential data problems (Liya Sutskever et al., 2014). Two LSTM networks are used in encoder and decoder format. One LSTM network read input sequences and the other read output sequences. As LSTM can capture long-term data, it has been used in sequential architectures so that long-term dependencies of textual data on the next sequential data can be retained. The model works by taking the input of previous instances and forecasting output based on the highest probability of any vector.

Also, LSTM supports multivariate data and can output sequences, which can further be used for multi-step forecasting. (Truc Thi Kim Nguyen et. al) [18] proposes different NN models on multivariate and univariate features. The study is focused on the impact of data fed as input, which uses generating power only for univariate analysis, and for multivariate analysis, wind speed and generating power are used as input to forecast power generation. The author proposed the use of LSTM to tackle the problem of vanishing gradient in the case of RNN.

After a number of iterations and experiments, author stated the use of the combined architecture of 1D CNN and LSTM model. Temporal characteristics of data can be

captured using 1D CNN layer, while to study long term dependencies LSTM can be used.

The limitation of LSTM architecture is that to forecast, architecture needs to be given with all input data which can increase computational cost. Instead, if architecture could focus on selective data with the highest dependency on the next sequential data then it would not be needed to pass every input sequential data.

The multi-step forecasting was the challenge that involved only LSTM layers. To overcome this problem, the study of Seq-seq architectures was done. The study (Avinash Achar et al.,2022) [19] shows that the proposed Encoder-Decoder architecture incorporates stochastic seasonal correlation and allows for multistep forecasting along with exogenous inputs. The limitation was that multiple input data were passed on in sequence, which did not give attention to the required data.

To overcome the requirement of all input data and to give attention to particular types of words/ sentences, a new concept of attention mechanism came into existence (Dzmitry Bahadanau et al., 2015) [20]. The primary focus was on passing the highly weighted vector in the input of the decoder and not all the vectors. At each timestep t , previous instances of inputs and hidden states are given to the decoder. Dimensions of hidden states of the encoder and decoder are the same. The weights given to all of the hidden states are known as attention weights. The sum of products of attention weights and hidden states forms the context vectors, which are then fed as an input to the decoder to give the final forecasting value.

The research was mainly on textual data. But the concept of sequential data can be applied to time series data. In this report the target variable is active power and is dependent on wind speed data. So, here, high weightage can be given to latest time period's wind speed values than the older one. By creating contextual vectors, accurate forecasting results can be generated. So, the attention mechanism helps in forecasting the results of time series data.

In the Luong attention mechanism, attention weights are not dependent on the previous decoder's hidden state; instead, they depend on the current decoder's hidden state (Minh-Thang Luong et al., 2015) [21]. To adjust the output dynamically, the current decoder's hidden state is introduced instead of the previous decoder's hidden state. More updated information is now available as input to the decoder. A simple dot product of hidden

states of the encoder and decoder is performed instead of using FFN to calculate attention weights. This change was to reduce the number of unwanted parameters.

To support the statement, Naixiao Wang et al. [6] propose a method that combines an attention mechanism and bidirectional long short-term memory (Bi-LSTM) to forecast wind power. The attention mechanism helps the model focus on relevant features, while the Bi-LSTM network captures long-term memory and reduces the occurrence of falling into local optimum.

In the field of sequence transduction, attention mechanisms and self-attention have played important roles. In time series data, the attention mechanism can be used to give weightage to different time steps in the input sequence as per the relevance of predicting the next time step. So, the attention mechanism focuses on particular data of time steps, which can be useful for the forecasting of desired data. However, the importance of each time step and how it can play its role in other time steps is found by studying the self-attention mechanism (Ashish Vaswani et al., 2017) [22]. In the context of time series data, self-attention mechanism refers to each time step to another time step to capture dependencies and patterns.

Self attention particularly useful for capturing seasonality, trends and irregular patterns.

2.6 Wind Power/ Speed forecasting in colder regions:

As mentioned earlier, wind speed varies with the change in geographical locations, and thus it affects on wind power generation. (Jon Kasper)[31] made a study report where he examines the effect of cold climates on wind turbine productivity in Sweden through the analysis of 10 wind parks with 45 turbines. It uses a method correlating daily production with daily average wind speed on summer days to predict winter production. Results indicate higher losses in northern sites due to icing, estimating losses of 10-20% at sites with icing. To accurately evaluate production losses, measurements of wind speed, air pressure, turbine production, and air temperature from the sites are required. Higher temporal resolution turbine data can help reduce errors due to wind speed averaging. Additionally, incorporating terrain, roughness, and air stability into the method, using a more accurate representation for high wind speeds, correlating production to the daily energy equivalent in the wind, and considering more wind sectors are some measures that could potentially improve the method.

(Si Han Li) [29] Focussed on impact of climate change on wind energy. This study analyzes the impact of climate change on wind energy using high-resolution Regional Climate Models (RCMs). The study examines the changes in wind speeds and wind energy on a seasonal and yearly basis across North America. It identifies both positive and negative trends in wind energy, with some regions experiencing significant changes. The study also investigates the influence of climate change on design-level wind speeds and highlights the need to consider changing design wind speeds for maintaining structural reliability. The results emphasize the modeling uncertainties and the need for further research in this field.

2.7 Smoothing Techniques to Denoise the data:

Faulty sensors, human errors, and outliers may cause faulty measurements in the data, which can be the reason for noise. The unwanted noise in the data affects forecasting results, and thus, it is needed to denoise the data. A plethora of techniques are involved in denoising the data.

(Irfan Prathama et. al., 2022) [23] have studied linear extrapolation, exponential smoothing, and moving average method. The results generated are compared with proposed Centroid Decomposition algorithm. The RMSE was significantly reduced than other techniques. Linear extrapolation uses the known data using tangent line, exponential smoothing here uses Holt's winter method, Moving Average is used to smooth out the fluctuations. CD method uses Savitzky-Golay method for approximation. The proposed method may not be efficient on short term predictions. To overcome this issue another denoising technique named wavelet transform has been studied.

(Liwen Quin et. al, 2022) [24] proposed wavelet transform as a denoising technique along with Prophet for multi-time scale data. The method extracts the effective signal by eliminating the noise.

Why to use the wavelet transform?

As per [25], wavelet transform is suitable for multi-step and nonstationary signals. The data used in our analysis is nonstationary, and thus, wavelet transform can be the effective denoising technique, if needed.

2.8 Key findings from literature survey:

From the overall literature survey,

- ✓ It can be studied that, effective time series forecasting mainly deals with the memory stored and handled. Neural networks with LSTM units effectively overcome the problem of vanishing gradient and long-term memory storage.
- ✓ The thesis mainly deals with multi-step forecasting, and hence it is found from the literature survey that seq-seq architectures are efficient in handling complex temporal data and, thus, multi-step forecasting.
- ✓ Variation of the forecasting window length is needed apart from the fixed input data length. In such a scenario, seq-seq NN is efficient.
- ✓ The data obtained is nonstationary and contains lots of noise (after evaluating the SNR value). Thus, data is denoised using wavelet transform, and further, Encoder-Decoder architecture consisting of LSTM units is proposed.

CHAPTER 3

Problem definition and formulation

3.1 Problem Definition and preliminary idea :

We know wind power forecasting is very important for effective energy resource management. However, wind power is greatly affected by wind speed, geographical areas, wind direction, and wind turbine parameters such as rotor RPM, generator windings temperature, gearbox bearing temperature, etc. Thus, due to its unpredictable and dynamic nature, wind power generation poses a big challenge.

So, considering these facts, it is essential to implement an accurate forecasting model using suitable models. The primary focus of our problem is to forecast intra-day and day-ahead power generation by implementing advanced sequential NN architectures like variants of LSTM along with denoising techniques.

Based on the literature review, LSTM can be used for long-term dependencies, but when dealing with very large data, it may degrade the performance of LSTM. To tackle this problem, Encoder-Decoder architecture is implemented, and the use of the attention mechanism to find temporal dependencies and extract important temporal weights is highlighted. [22] LSTM and related sequential networks solve the exploding or vanishing gradient problem of simple RNN and thus can help determine wind power generation forecasting.

The accuracy of each model can be found using metrics such as MSE, RMSE, MAPE, and MDA.

3.2 Problem of power loss in colder regions due to snow deposition over turbines:

Snow distribution in Europe

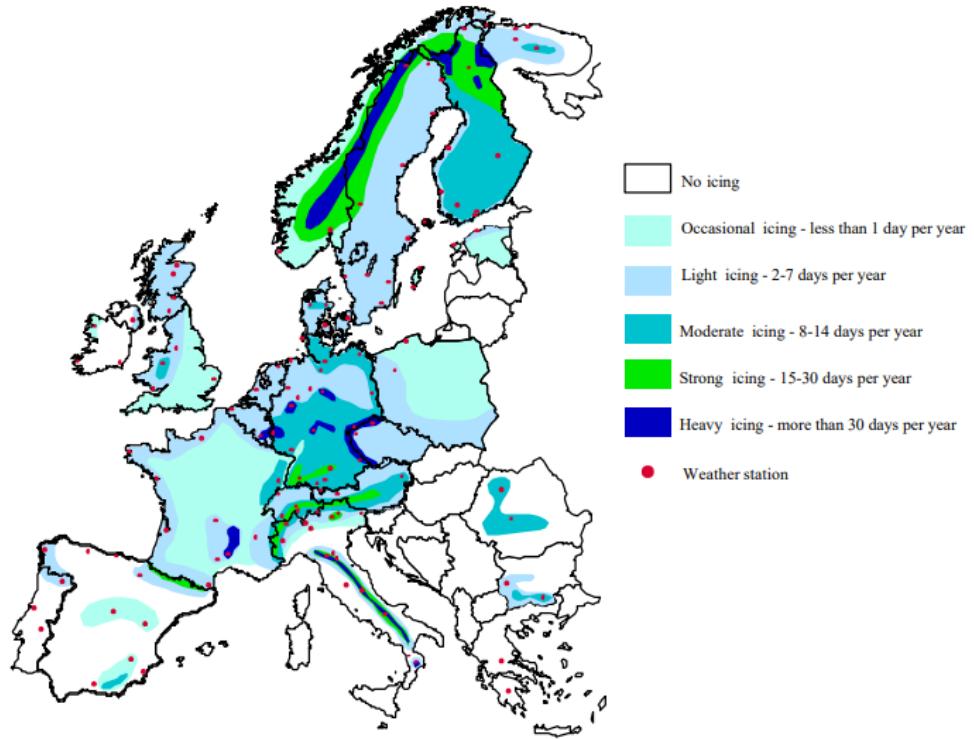


Fig 3.1: Snow distribution in Europe [5]

As per [39], Ice accretion does not occur uniformly along the length of the wind turbine blade. Ice thickness increases significantly near the blade tip, while less ice is observed near the root section. Two main factors contribute to this pattern ,the angular movement of the wind turbine blades and the tapered shape of the blade structure.

The leading edges of wind turbine blades collect more ice around stagnation points. At stagnation points, the local velocity of air is found to be zero [40-41]. The weight of the blade after the accretion of ice may be as high as 50% of the blade's structural weight, and this will be a potential risk for a turbine.

In 2011, at university of Troms, [41] due to ice accretion load, loss of energy is tabulated:

Month	Loss of Energy(%)	Ice Load(Kg/m)
Oct	9.9	245.7
Nov	37.8	424
Dec	12.23	216
Jan	52.5	686.7
Feb	18.5	174

Table 3.1: Month wise loss of energy due ice deposition over turbine

To optimize the performance and reliability of wind farms in colder regions, it is crucial to accurately forecast power generation while simultaneously predicting the occurrence of ice and snowfall, which can significantly impact turbine efficiency and structural integrity. Utilizing climatic data—specifically humidity, wind speed, and temperature—we aim to develop a predictive model that can:

1. **Forecast Wind Power Generation:** Predict the amount of power that will be generated based on current and projected weather conditions, considering the influence of temperature and wind speed on turbine efficiency.
2. **Predict Ice/Snow Accretion:** Anticipate the likelihood and severity of ice and snow accumulation on turbine blades and other critical components based on temperature fluctuations, humidity levels, and wind speed.

De-icing Techniques:

There are already de-icing techniques available. De-icing techniques are advantageous in colder-regions where snow accretion over tip of wind turbine blades held.

What are these techniques?

As per [43], following techniques have already been installed:

Siemens De-icing System:

- Developed an **electrical de-icing system** using carbon heating mats on blade edges.
- De-icing is achieved within an hour, and it involves an ice detection system and a control strategy.
- Installed in Sweden, more than 50% of Siemens turbines in the region are equipped with this system

Enercon De-icing Technology:

- Utilizes **hot air circulation** inside the blade after ice detection via power curve analysis.
- Ice is melted by circulating hot air up to the blade tip, which allows early-stage ice detection without halting turbine operation

Nordex De-icing Technology:

- Features heating elements on the rotor blades' leading edge, automatically activated by an ice sensor.
- The system boosts energy yield by 8-25% during severe icing conditions with minimal power consumption

Vestas De-icing System:

- Uses **hot air circulation** to target critical areas of the blade, focusing on de-icing the outer sections.
- Minimizes power loss and de-ices the blades within two hours

VTT Finland's Electro-thermal Heating:

- Developed an **electro-thermal system** with conductive fibre mats integrated into blades.
- Reduces downtime by allowing de-icing during normal operation and consumes less than 2% of the turbine's annual energy output

Gamesa's Anti-icing Paint:

- Developed **Blade shield anti-icing paint**, preventing ice formation and increasing resistance to corrosion.
- Developed with nanomaterials, its real-world effectiveness is still under evaluation.

But the listed techniques require proper maintenance failing of which may impact on power generation performance.

What kind of maintenance is required?

- The hot air circulation system involves fans and heating units inside the blades, which may require occasional servicing to maintain airflow efficiency and heating effectiveness. Sensors used for ice detection and the fan's heating elements are also likely to need regular checks
- While Nordex's system is designed for low power consumption and minimal interference, the heating elements on the blade's leading edge and the ice sensor will need routine inspections to prevent damage or degradation over time, especially in harsh climates.
- Although the system is designed for low energy consumption, the electrically conductive mats and smart control systems embedded in the blades will require occasional servicing to ensure that they continue functioning efficiently under varying weather conditions
- The anti-icing paint itself is designed to minimize maintenance by preventing ice build-up, but the surface of the blades should be periodically checked for wear or corrosion. Reapplication may be necessary after a certain period due to environmental stress.

Ambient temperature as an input will help in determining the snowfall conditions, and accordingly, in the specific regions of the turbine, routine maintenance can be scheduled. If at all, any precautionary steps are needed, the turbine plant can arrange the safety protocols in advance.

Thus, the goal is to enhance decision-making processes for wind farm operation and maintenance, ultimately improving energy output and reducing the risks associated with ice and snow accretion on turbines in cold climates.

3.3 Problem of power loss in hot climatic regions due to dust deposition over turbines:

As per the study [\[44\]](#) following problems are highlighted

1. Effect of High Temperatures on Wind Turbines

Temperature Constraints: Wind turbines in desert environments like Kuwait face extreme temperatures, which can exceed 45°C (113°F). At these temperatures, the turbines' electronic and mechanical components experience “derating,” where power

output is reduced to prevent overheating. The turbine's control system will initiate a derating process as temperatures rise above 40°C, continuing to limit power output until temperatures return to a safer range.

Power Derating Process: The report highlights that when temperatures exceed 40°C, turbines cannot operate at full capacity and are progressively derated to avoid damage. At 45°C, turbines shut down entirely to protect sensitive components. In Kuwait's Shagaya wind farm, this derating process is significant since temperatures between 40-44°C account for a large portion of the energy losses.

2. Dust Accumulation and its Impact on Performance

Accumulation on Blades and Nacelle: Dust commonly collects on turbine blades and within the nacelle, reducing efficiency. Dust on blades increases drag and decreases lift, directly lowering power output. Dust on nacelle components blocks ventilation, hindering the internal cooling system and exacerbating temperature-related derating.

Frequency of Dust Cleaning: In the Shagaya wind farm study, it was found that regular dust cleaning can mitigate these losses significantly. After a cleaning cycle, high temperatures initially account for most energy losses. However, as dust reaccumulates, both dust and heat contribute to reduced power output. The energy loss from dust increased from 0.31% in the first year to 1.45% in the second year, underlining the importance of frequent cleaning in dusty environments.

3. The “Dust Package” and Sand-Proofing Adaptations

Dust-Resistant Nacelle Design: To address dust-related performance issues, manufacturers have introduced a “dust package” for turbines operating in deserts. This package includes modifications such as filters or louvers on the nacelle’s ventilation system to prevent sand infiltration while maintaining airflow. Louvers are angled to block dust but allow hot air to escape, using natural convection for cooling without added energy consumption.

Sand-Proof Brushes: Engineers have added small, wiry brushes in gaps around the nacelle to prevent sand intrusion. These brushes block sand grains from reaching sensitive areas within the turbine, minimizing the risk of mechanical wear and extending component longevity.

4. Cooling System Adaptations for Desert Environments

Convection-Based Cooling: Standard cooling systems in wind turbines use air circulation to dissipate heat, but this is less effective in hot deserts. Turbine manufacturers for desert operations have enhanced nacelle designs with louvers and venting systems that rely on convection. These systems use the natural upward movement of hot air to cool components without active refrigeration, thus reducing power consumption and the need for maintenance.

Temperature-Resistant Lubricants: In high temperatures, conventional lubricants can thin out, causing gear wear. Special lubricants are used in desert turbines that maintain viscosity at elevated temperatures, ensuring smooth operation and protecting the gearbox.

5. Operational and Strategic Implications

Impact on Energy Production: In hot deserts, optimizing maintenance, particularly cleaning schedules, is crucial for maximizing energy output. For instance, after cleaning, turbines at Shagaya operated at near-full capacity until dust accumulation began impacting performance again. By increasing the frequency of dust removal, operators can reduce dust-induced losses, which account for a significant portion of the turbines' energy losses in the desert.

Problem Formulation:

Power forecasting has been done by implementing various sequential and statistical models. The parameters used and equations related to the wind turbine dataset have been explained in detail in each of the subsequent sections.

Variables:

The variables used in the LSTM algorithm in forecast experiments are as follows:

X_t = input at time t,

h_t = output,

C_t = cell state,

f_t = forget gate,

I_t = input gate,

C'_t = coordinate cell state,

O_t = output gate.

σ = sigmoid activation function,

W_f = weight matrix for forget gate,

tanh = tanh activation function,

W_i = input gate weight matrix,

W_c = coordinate cell gate

Objective function:

Here, the objective of the method is to minimize MSE, RMSE, and MAPE of predicted and actual results and increase MDA.

Obj= $\min \sum_{t=1}^{T=t} L(y'_p, y_p)$ Here, y'_p is the predicted output, and y_p is the actual output of wind power resp.

The loss function we are using is MSE, RMSE and MAPE.

MDA = if the sign of (actual power[t+1] – actual power[t]) = (predicted power[t+1] – predicted power[t]), then the accuracy of that direction is correct.

MDA is calculated in range of 0 to 1.

0 = Least Accuracy

1 = Best Accuracy

3.4 Mathematical Formulation of SARIMA model:

SARIMA is a seasonal statistical model.

Mathematical representation can be shown like this:

$$(1 - \emptyset_1 B)(1 - \theta_1 B^s)(1 - B)(1 - B^s)y_t = (1 + \varphi_1 B)(1 + \beta_1 B^s)^* \varepsilon_t \dots \dots \dots 3.4.1]$$

Where,

B = Lag operator

Φ = Non seasonal AR component

Θ = Non Seasonal MA component

s = Seasonal Period

ε = White Noise error

3.5 Stacked LSTM Algorithm formulation:

Instead of focusing on a single LSTM layer, making the network deeper for more accurate results and long-term dependencies is always better.

Each LSTM unit in the stacked layer consists of a cell state, forget gate, input gate, and output gate.

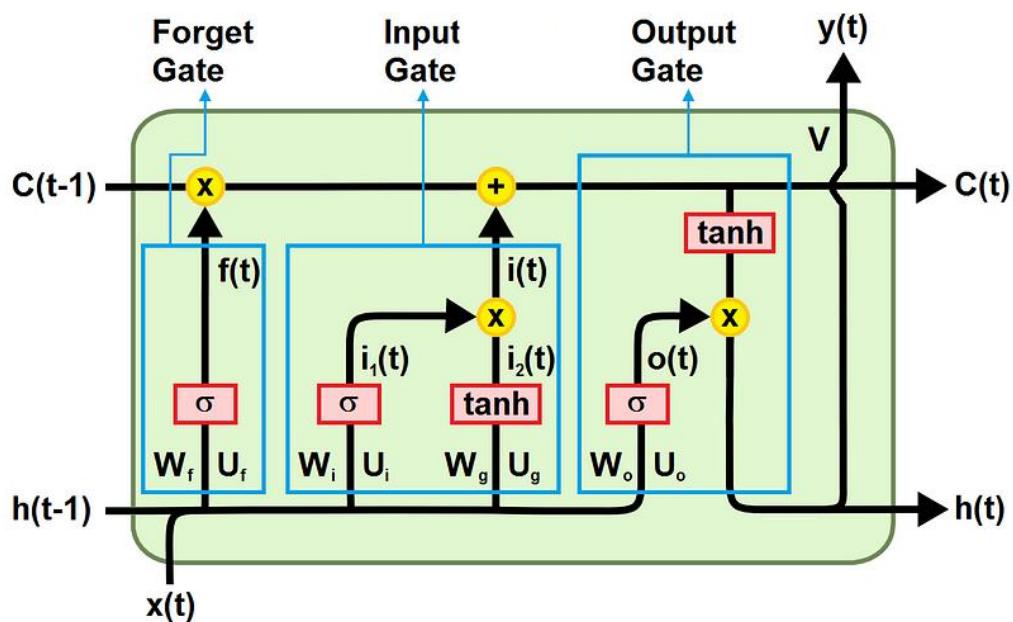


Fig 3.2: Basic architecture of LSTM unit

Forget gate: This gate acts on the information we receive. Important information is stored in the memory, and the rest is forgotten. Information obtained from the previous hidden state and current input information is passed in the sigmoid function. An information output close to 0 means we can forget the information, and an output close to 1 will be stored in the memory.

$X_{i,t}^j$: The value of feature j for the i-th sample at time step t.

Input Gate: This gate is used to update the cell state. As discussed, initially, the previous hidden state and current input are passed through the sigmoid function. Output close to 1 will be considered as important, and further, it is passed from the tanh function to compress the values between -1 and +1. Sigmoid output and tanh output are multiplied point-wise.

$$I_1(t) = \sigma(x(t)) \cdot U_i + h(t-1) \cdot W_i$$

$$I_2(t) = \tanh(x(t) \cdot U_g + h(t-1) \cdot W_g)$$

$$I(t) = I_1(t) * I_2(t) \dots \quad \dots \quad [3.5.2]$$

Cell state: Cell state transfers the required information during processing. The information obtained from the previous step helps reduce short-term memory. So, during training, considering all time steps, the cell state gate learns which information is required to keep and which one to forget.

$$C(t) = \sigma(f(t) * C(t - 1) + i(t)) \dots \quad \dots \quad 3.5.3$$

Output gate: This final gate will evaluate the value of the next hidden state. The previous hidden and current input states are added and then transferred to a sigmoid function. The new state obtained is then passed to the tanh function. In the end, the output obtained from the sigmoid function is multiplied with tanh output to decide the next hidden state information.

Equation 3.5.5 gives you the correct forecast result of power generation.

3.6 Encoder-Decoder Architecture Formulation:

We are already using stacked LSTM to overcome the problem of vanishing gradient caused by standard RNN by introducing the gradient flow in the network.

Considering these facts, some points are required to answer.

- Why is it required to go for encoder-decoder architecture?

In RNN, after using stacked LSTM we get output corresponding with the input for each time step. But in many real cases, we may need to forecast output sequences of different lengths provided input sequences of different lengths. Multi-step forecasting is one of the real-based cases. Thus, the encoder-decoder structure is introduced to address the sequence of the sequence mapping model. As the name suggests, the encoder takes a sequence as input and generates a sequence as output.

Encoder: It is responsible for going through the input time steps and encoding the sequence into a fixed-length vector known as a context vector.

Here, the encoder is a stack of numerous LSTM cells. Each cell accepts the single element from the input sequence, stores information, and moves forward. The hidden state vector is computed by applying a function on weights to the previously hidden state $h(t-1)$ and input vector $x(t)$.

$h(t)$ has all the encoded information obtained from the previous hidden state. Context vector generated from encoder will be initial hidden state of the decoder. To make accurate forecasts, it encapsulates the information for all input elements.

Decoder:

The decoder accepts the input from the previous hidden state $s(t-1)$ and produces output $y'(t)$ as well as its own hidden state $s(t)$.

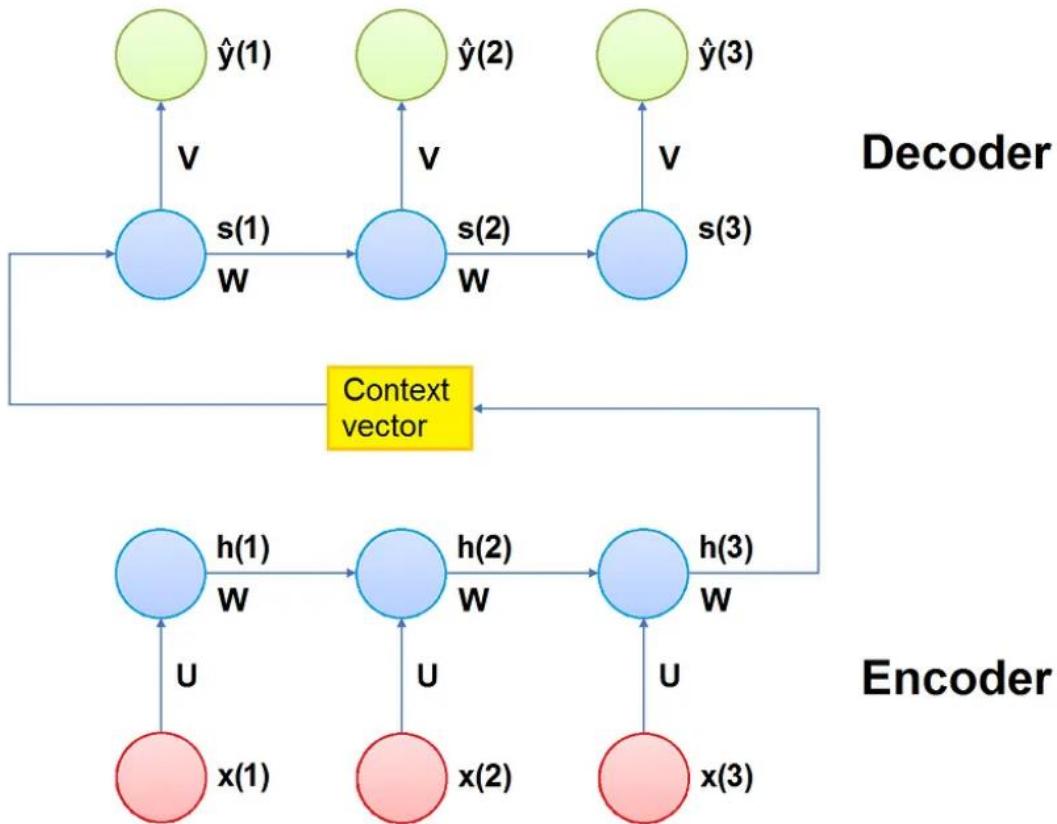


Fig 3.3 : Simple Representation of encoder decoder architecture

$$S(t) = f(Wh(t-1)) \dots \quad [3.6.2]$$

The output is computed using the softmax function applied on current time step $s(t)$ along with the weight V . It creates a probability vector.

$$y'(t) = \text{softmax}(V.S(t)) \dots \quad [3.6.3]$$

This architecture is found to be useful for small sequences. For long sequences, e.g., in seasonal periods, architecture can't store memory to give output in relation to seasonality.

CHAPTER 4

DATASET DESCRIPTION AND METHODOLOGY

This chapter provides details of the dataset used and a high-level understanding of the problem-solving approach. As mentioned, our dataset mainly deals with temporal relationships, which in turn are based on sequence. Thus, for intra-day and day-ahead forecasting sequence to sequence, neural networks play an important role.

4.1 DATASET DESCRIPTION

4.1.1 DATASET-1

Power generation can be done by wind turbines, which is one of the techniques used in renewable power generation sources. To forecast the power generation, we need an ample amount of data and time series observations from a certain windmill. The dataset used in this report is a public dataset, and observations are generated from certain windmills. The data available is from Jan 2018 to March 2020.

Initial data was very clumsy, with 1,18,225 observations and 21 features. Data was with lots of missing values and outliers. It became necessary to impute missing values with appropriate methods to minimize data loss. Imputation by mean was not suggested, as outliers were present in the data. To overcome this, the MICE and KNN imputer have been applied to impute missing values. KNN results helped preserve the data's shape and reduce outliers, so final imputation is done by using KNN imputer. Outlier removal is done using interquartile range.

Now, the updated data has 78460 observations and 19 features. Data is observed at 10 min intervals. Some of the key features in the data are listed below which gives the characteristics of that feature:

Feature	Unit	Characteristic
Active Power	MWhr	It's a target variable and refers to electrical output generated by the turbine
Ambient Temperature	°C	It represents the surrounding Temperature
Bearing Shaft Temperature	°C	Temperature of bearing that supports rotor shaft
Blade pitch angles	radians	The angle at which turbine blades are set to control lift and drag generation while rotating.
Gearbox Bearing Temperature	°C	Bearing temperature supports rotational energy transmission from turbine blades to generators.
Gearbox Oil Temperature	°C	Lubricating oil responsible for dissipating heat from turbine components
Generator RPM	rpm	The speed at which the generator rotor rotates
Generator winding temperatures	°C	Temperature of conductive windings of generator where electricity is generated
Hub temperature	°C	The central part of the turbine where blades are attached.
Reactive Power	KWhr	Maintains voltage levels in the AC system. It does not give useful work as that of active power.
Wind Speed	m/s	The velocity of wind is responsible for rotating turbine blades and converting kinetic energy into electric energy. It is the main component of active power generation.

Table 4.1: Key feature definitions and characteristics for dataset-1

The univariate and multivariate analyses of all features have been done, and the correlation between them is tabulated. Based on the correlation, it is observed that wind speed is highly correlated with the target variable. The correlation observed is:

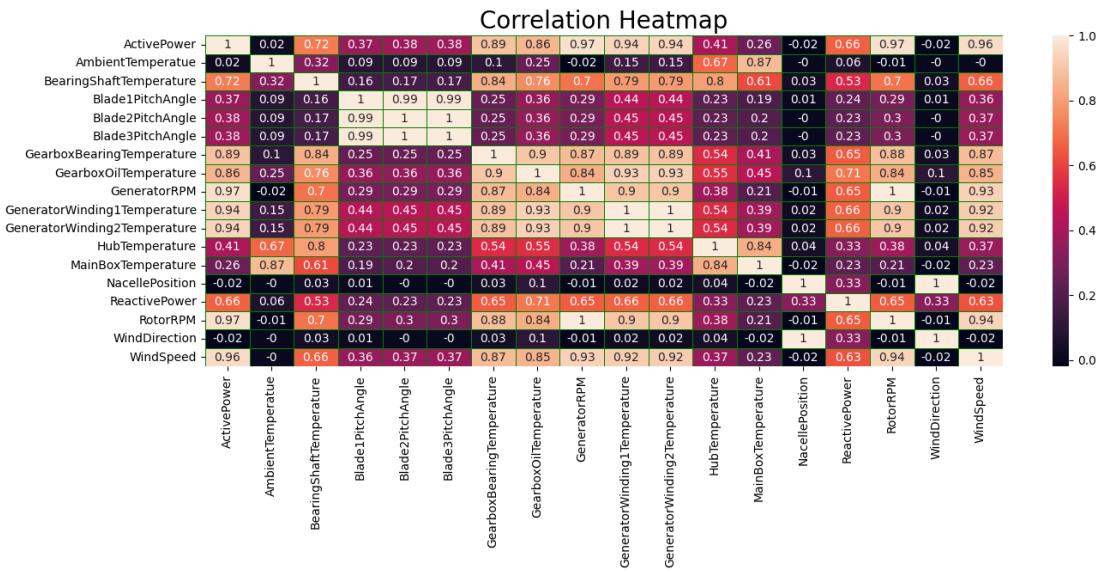


Fig 4.1- Correlation Heatmap of wind turbine data features

The features that correlate highly with the target variable are also highly correlated with wind speed. So, to avoid multicollinearity, wind speed has only been taken for further analysis. As observed, wind speed correlation with power is 0.96.

Considering the importance of wind speed, the power curve is plotted between active power and wind speed.

Power output versus wind speed

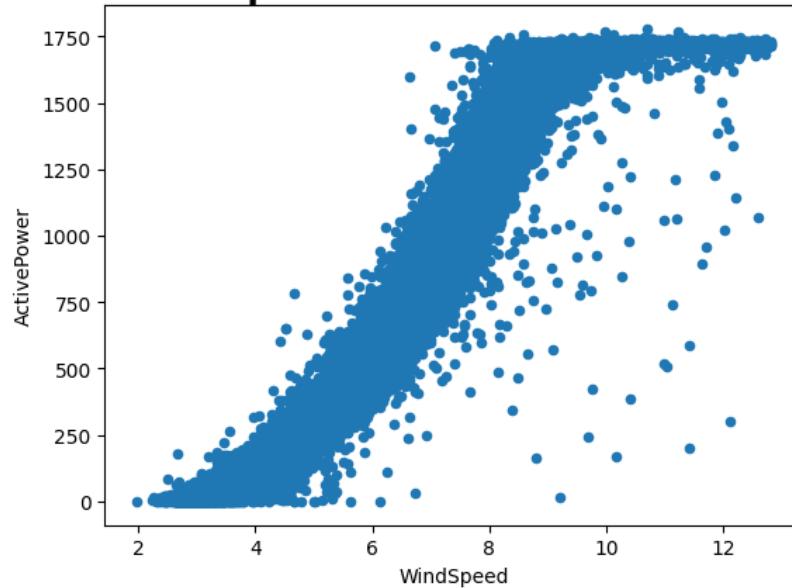


Fig 4.2- Power Curve relationship between Active Power and Wind Speed

The cleaned data has 78459 observations and 2 features. The data is split further into train and test series considering Jan 2018 to July 2019 in train to cover seasonal forecasting and Aug 2019 to March 2020 as test data. 52651 observations are with the train dataset, and 25520 observations are with the test dataset. The observations are not on the same scale, so to bring them on a unique scale, scaling has been done.

The improved scaled data has 52651 observations and two important features: Active power, which is the target variable, and wind speed, which will be input. Further comments on input and output data in respective models is explained in part A and part B resp.

As data obtained is with 10 min of the interval, in 24 hrs, we will have 144 such intervals. The same technique is used to decide the sliding window, and thus, 144×7

(1 week of data) previous time interval data is taken to predict the next data point, and this process continues for each iteration till all the observations are not iterated.

The obtained sequential train and test data is then provided to different Neural Network architectures for training and forecasting purposes.

4.1.2 DATASET-2 :

- **Dataset Overview:** The dataset comprises multi-annual raw SCADA data recorded at 10-minute intervals from a Vestas V52 wind turbine.
- **Location:** The wind turbine is situated at Dundalk Institute of Technology, Ireland (GPS Coordinates: 53.98352, -6.391390), in a peri-urban environment.
- **Operational Period:** The data spans from 30 January 2006 to 12 March 2020.
- **Turbine Specifications:** The turbine features a hub height of 60 meters and a rotor diameter of 52 meters.
- **System Characteristics:** Operating as a behind-the-meter system, this turbine generates power primarily for local use.
- **Maintenance Note:** A significant gearbox replacement occurred between 04 October 2018 and 27 July 2019, during which there was no positive electrical power output.

The features of the dataset are tabulated below:

Characteristic	Description
Target Variable	Refers to electrical power output (the main target of the analysis).
Surrounding Temperature	The ambient temperature around the wind turbine.
Bearing Temperature	Temperature of the bearing that supports the rotor shaft.
Blade Angle	The angle at which turbine blades are set to capture wind energy.
Main Shaft Bearing Temperature	Temperature of the bearing that supports rotational energy transfer to the shaft.
Lubricating Oil Temperature	Temperature of oil responsible for dissipating heat and reducing friction.
Generator Rotor Speed	The speed at which the generator rotor rotates.
Generator Winding Temperature	Temperature of the conductive windings in the generator.
Hub Temperature	Temperature at the central part of the turbine where blades are attached.
Reactive Power	Maintains voltage levels in the AC system (not directly involved in power generation).
Wind Speed	The velocity of the wind responsible for rotating the turbine blades.

Fig 4.3: Key feature definitions and characteristics of Dataset-2

The above features are studied, and their impact on the target variable is measured using a correlation matrix. The correlation matrix is as shown in the figure:

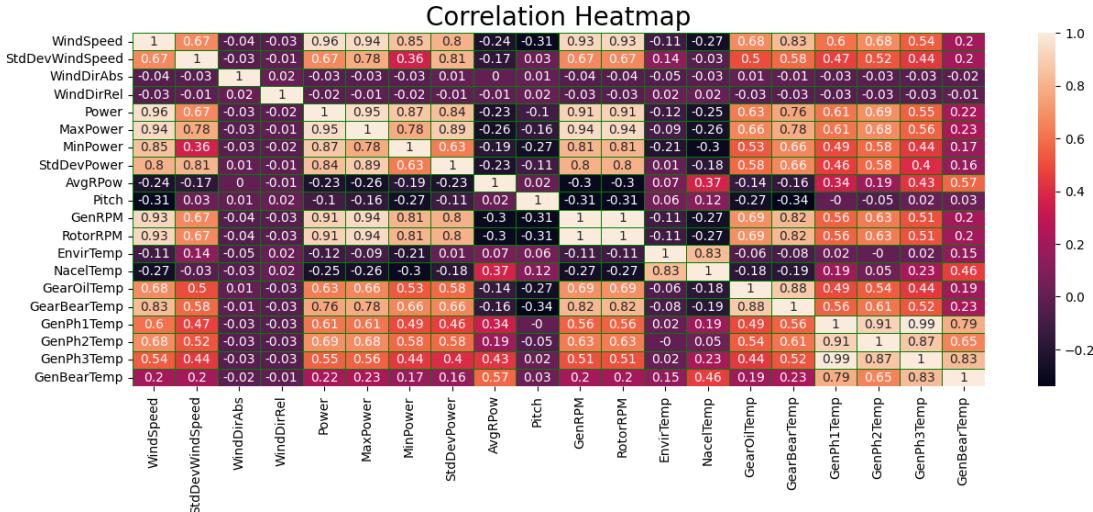


Fig 4.4- Correlation Heatmap of wind turbine data features for dataset-2

Initial dataset obtained from the source was very unstructured with 6,52,997 timestamps and 21 features. The data has missing values and outliers.

It became necessary to impute missing values with appropriate methods to minimize data loss. Imputation by mean was not suggested, as outliers were present in the data. To overcome this, the MICE and KNN imputer have been applied to impute missing values. KNN results helped preserve the data's shape and reduce outliers, so final imputation is done by using KNN imputer.

The features that correlate highly with the target variable are also highly correlated with wind speed. So, to avoid multicollinearity, wind speed has only been taken for further analysis. As observed, wind speed correlation with power is 0.96.

Considering the importance of wind speed, the power curve is plotted between active power and wind speed.

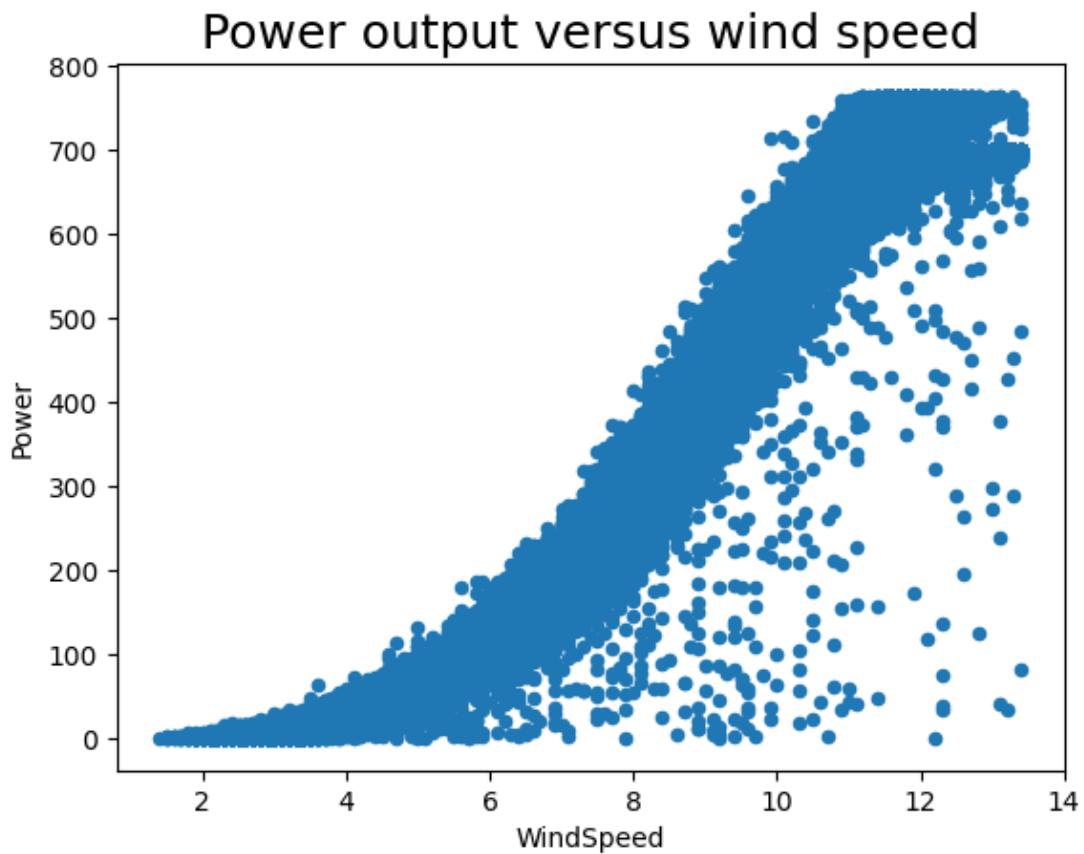


Fig 4.5- Power Curve relationship between Power and Wind Speed for dataset-2

Dataset-2 important features Detailing:

Key Characteristics of

1. Wind Speed:

Wind speed varies season wise. The time series plot of the same is as shown in fig 4.6:

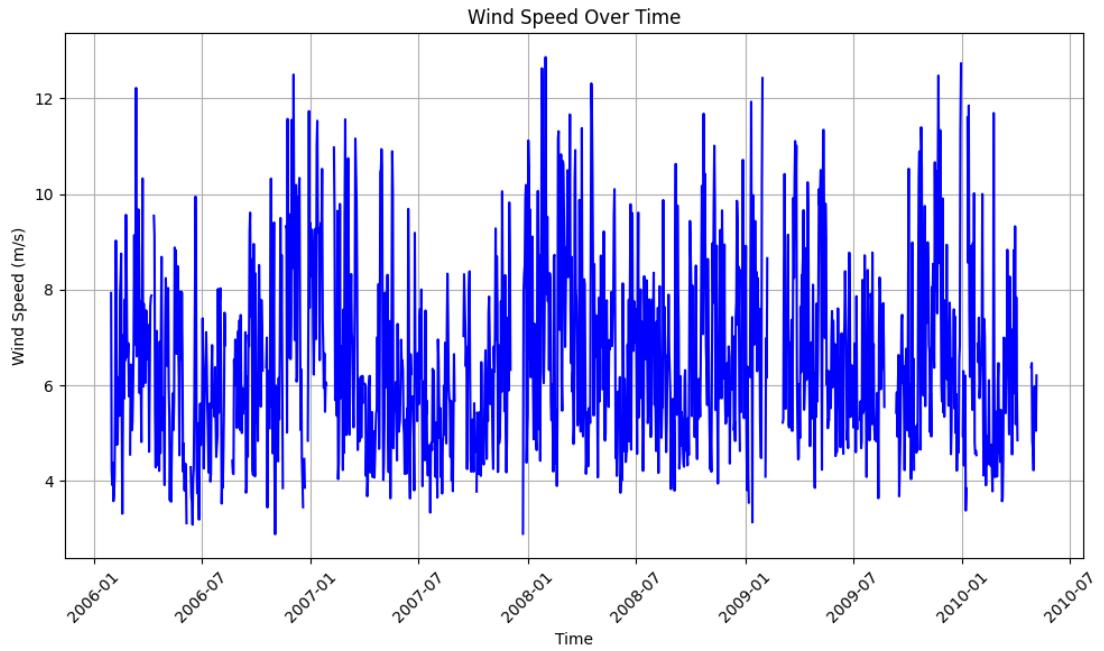


Fig 4.6 Timeseries plot of wind speed feature

Now, it is essential to understand the maximum wind speed recorded so that we can consider wind gust criteria.

As per wind gust analysis, during any gusts, minimum expected speed is above 17 m/s.

Feature Name	Maximum Speed Recorded (m/s)	Minimum Speed Recorded (m/s)	SNR value Before Denoising(dB)	SNR value After Denoising
Wind Speed	13.4	1.4	9.33	14.97 dB

Table 4.2 Windspeed feature details

It is necessary to revalidate the denoised data. It should contain the characteristics of original data. The validation can be done through residual analysis.

Histogram of Residuals and Q-Q plot:

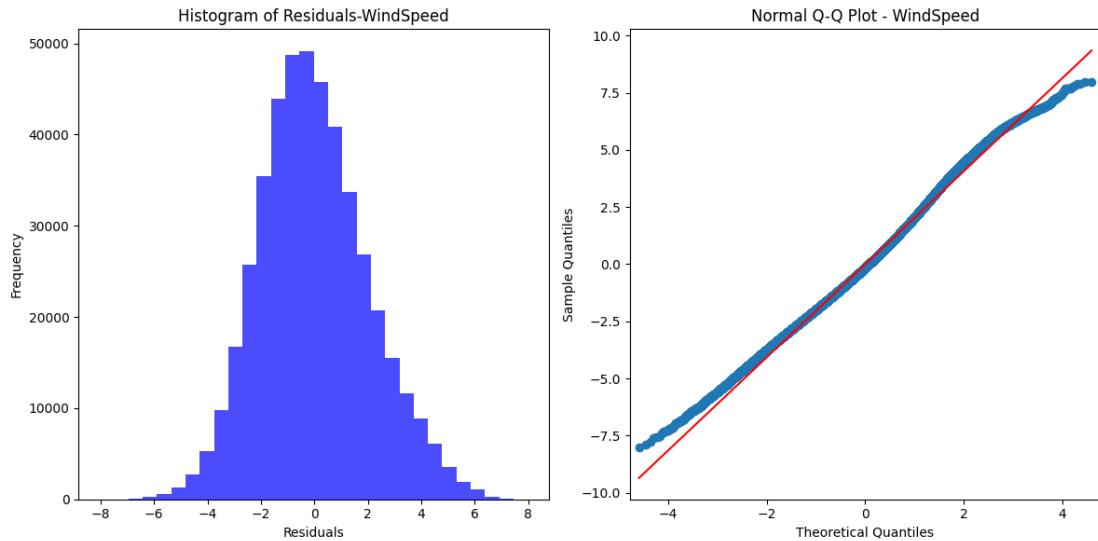


Fig 4.7 Residual analysis of denoised data

Results obtained show that, denoised data has similar characteristics as of original data. The analysis is done on the basis of nature of residuals obtained.

If residuals are normally distributed and if Q-Q plot contains most of data observations on a slant line then the nature of denoised data is similar to original data.

2. Power

Month wise generated turbine power is as shown in fig:

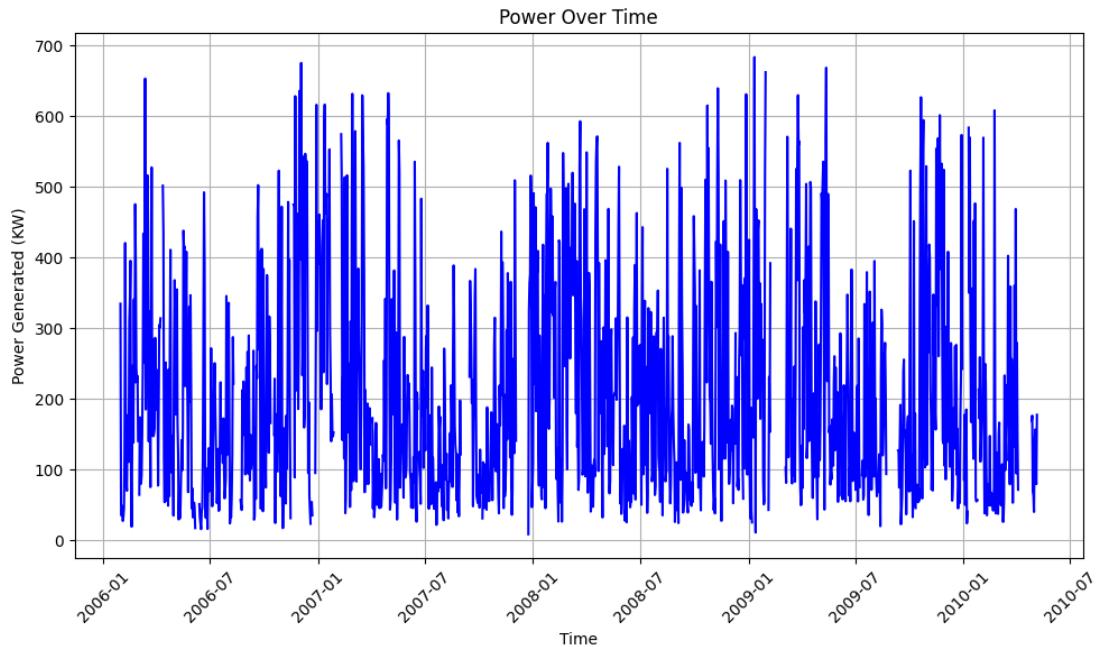


Fig 4.8 Timeseries plot of Power feature

3. Temperature:

As temperature decreases, ideally, wind speed increases. It is because of the pressure difference. Wind moves from a higher pressure area to a lower pressure area. When the temperature decreases, the air becomes denser and creates a high-pressure area, which boosts the speed of the wind.

Turbine is located in Ireland. The minimum and maximum temperatures recorded in that region is as follows:

Feature Name	Maximum Temperature Recorded (°C)	Minimum Temperature Recorded (°C)
Temperature	20	-5

Table 4.3 Temperature feature details

Temperature variation with seasons is as shown:

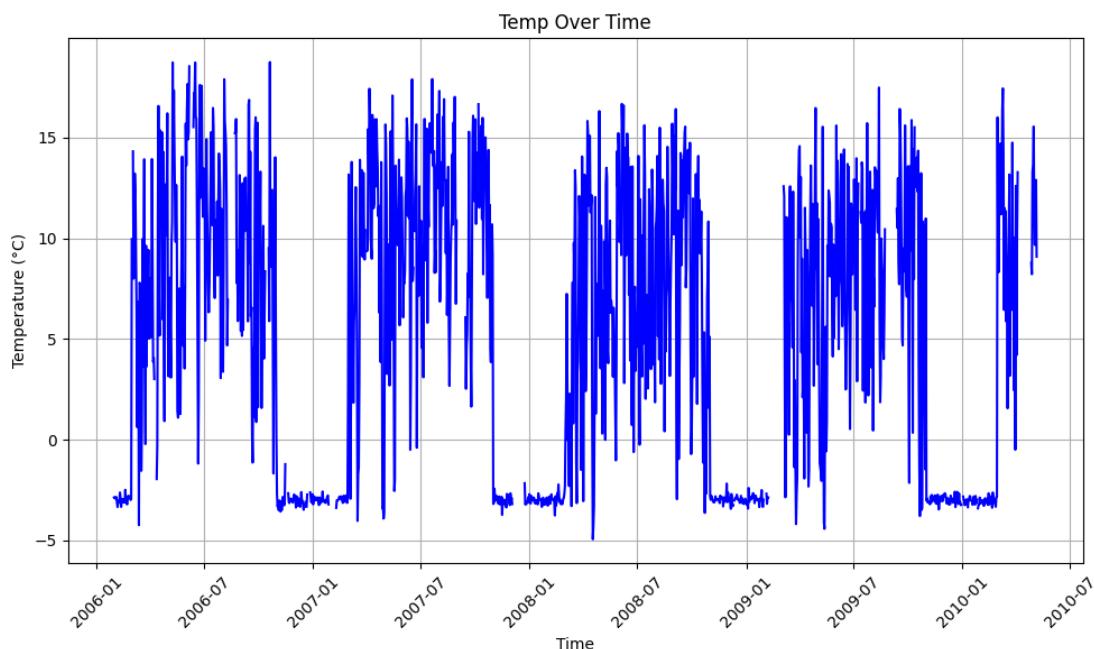


Fig 4.9 Timeseries plot of Temperature feature

4. Humidity:

Relative humidity and temperature has inverse relation.

$$\text{Relative Humidity} = \frac{\text{Amount of water vapour present in water-air mixture}}{\text{Actual amount needed for saturation at the same temperature}}$$

Relative humidity lies between 0 and 1.

Humidity variations according to seasons is shown below:

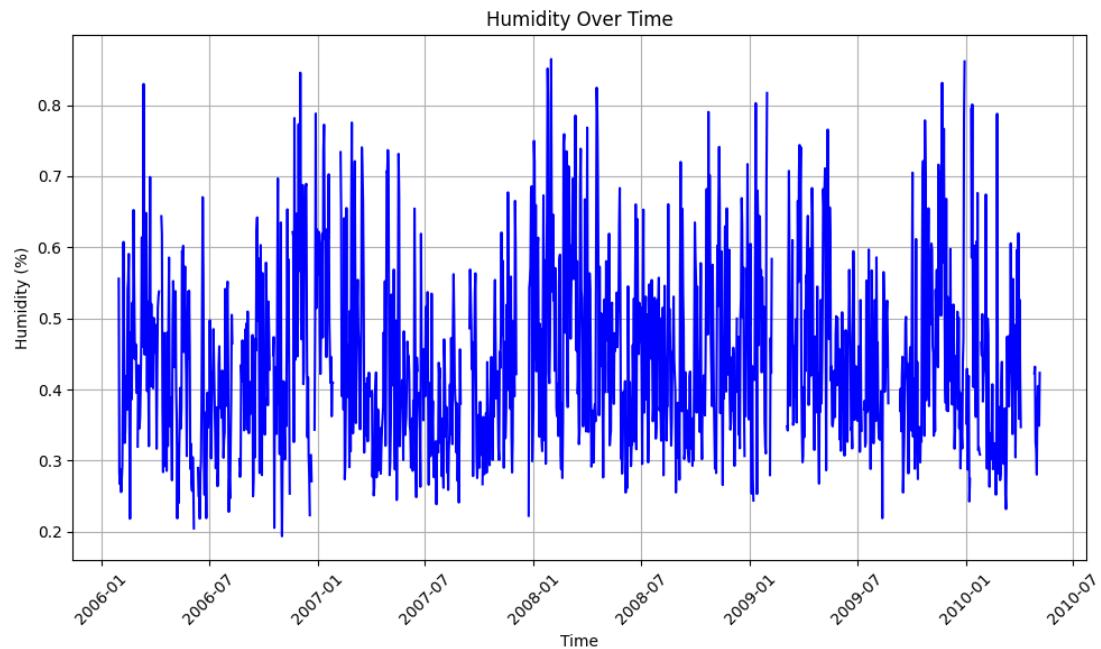


Fig 4.10 Timeseries plot of Humidity feature

As data frequency of this dataset is similar to dataset-1, same methodology can be applied.

4.1.3 DATASET-3 :

- **Dataset Overview:** The dataset comprises multi-annual raw SCADA data recorded at 10-minute intervals from Siemens Gamesa G97 wind turbine.
- **Location:** The wind turbine is situated Al Jahara, Kuwait, (GPS Coordinates: $29^{\circ} 13' 33.1''$, $47^{\circ} 2' 58.4''$).
- **Operational Period:** The data spans from 01st January 2015 to 31st Aug 2017.
- **Turbine Specifications:** The turbine features a hub height of 78.98 meters and a rotor diameter of 97 meters.

The features of the dataset are tabulated below:

Characteristics	Description
Target Variable	Refers to electrical power output (KW)
Ambient Temperature	Ambient temperature around wind turbine (°C)
Wind Speed	Velocity of wind responsible to rotate blades of turbine (m/s)
Relative Humidity	Presence of content of water vapour in atmosphere (%)

Table 4.4: Key feature definitions and characteristics of Dataset-3

The above features are studied, and their impact on the target variable is measured using a correlation matrix. The correlation matrix is as shown in the figure:

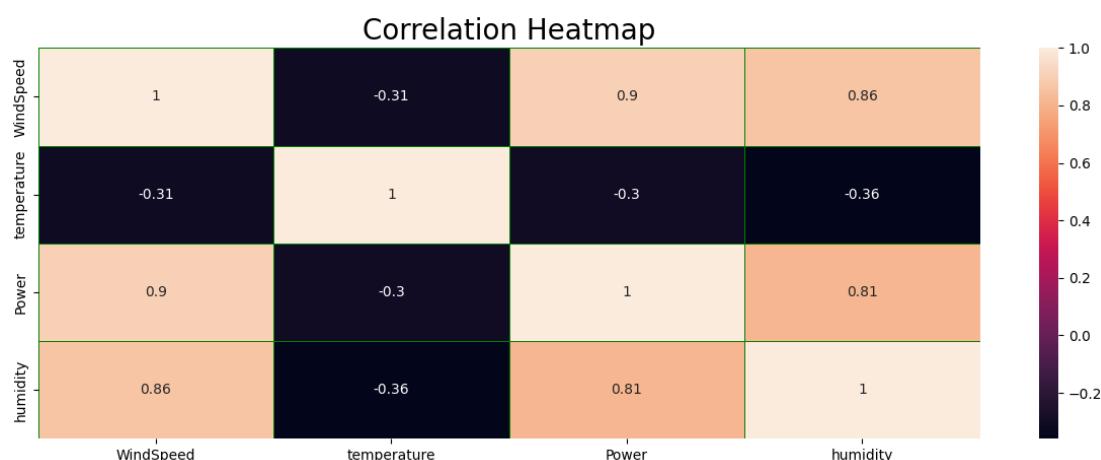


Fig 4.11- Correlation Heatmap of wind turbine data features for dataset-3

Initial dataset obtained from the source was very unstructured with more than 1,00,000 timestamps and 4 features. The data has missing values and outliers.

It became necessary to impute missing values with appropriate methods to minimize data loss. Imputation by mean was not suggested, as outliers were present in the data. To overcome this, the MICE and KNN imputer have been applied to impute missing values. KNN results helped preserve the data's shape and reduce outliers, so final imputation is done by using KNN imputer.

The features that correlate highly with the target variable are also highly correlated with wind speed. So, to avoid multicollinearity, wind speed has only been taken for further analysis. As observed, wind speed correlation with power is 0.90.

Considering the importance of wind speed, the power curve is plotted between active power and wind speed.

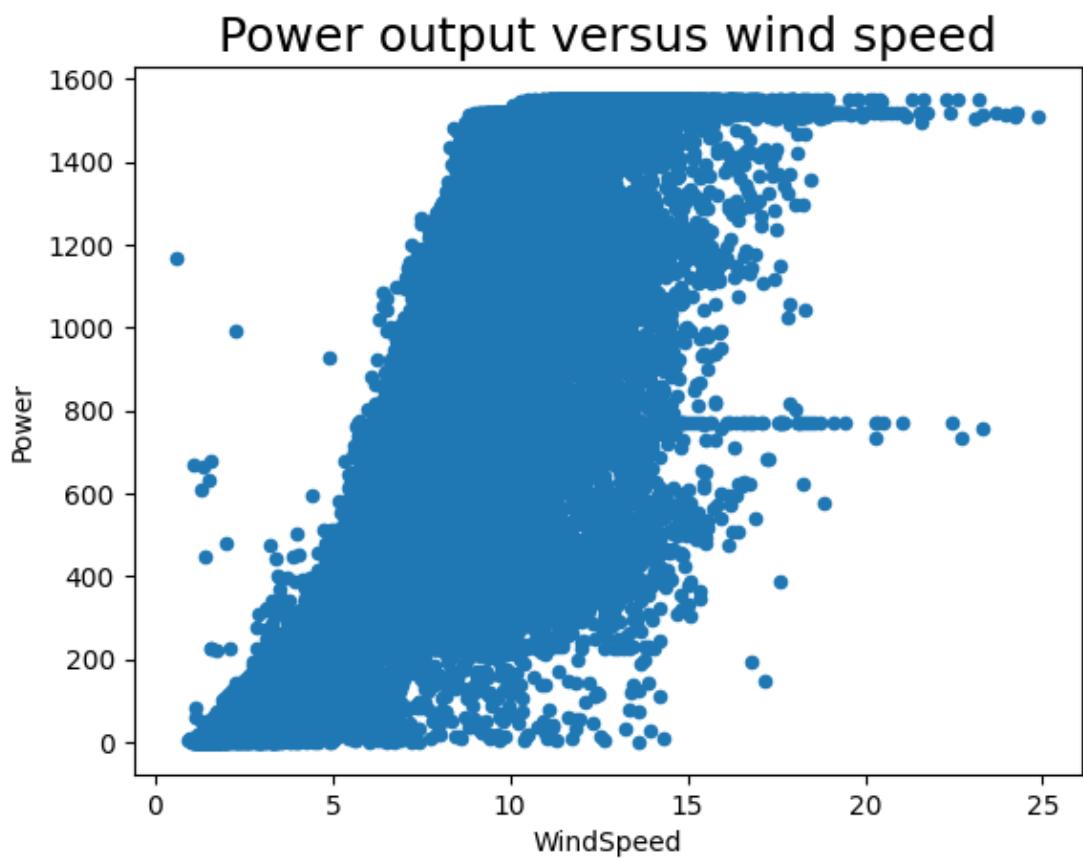


Fig 4.12- Power Curve relationship between Power and Wind Speed for dataset-3

Dataset-3 important features Detailing:

Key Characteristics of

1. Wind Speed:

Wind speed varies season wise. The time series plot of the same is as shown in fig 4.13:

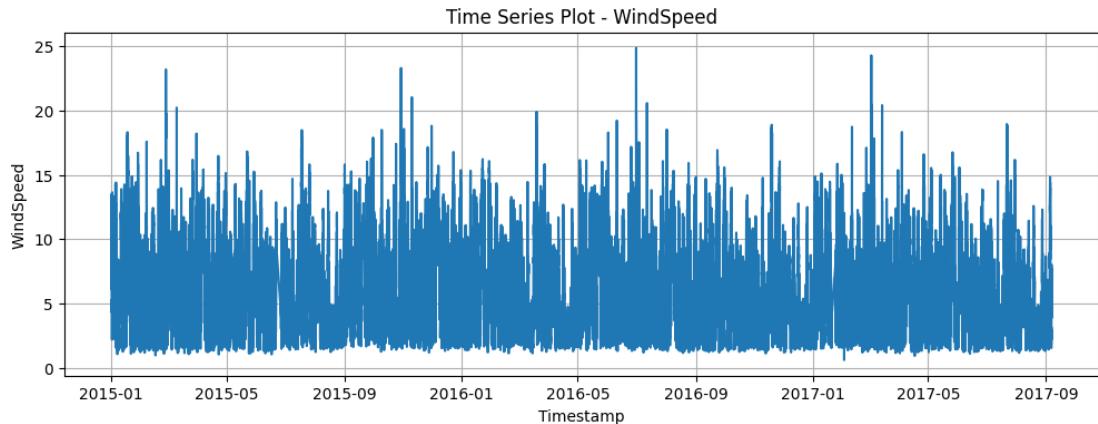


Fig 4.13 Timeseries plot of wind speed feature

Lets calculate maximum and minimum wind speed to consider wind gusts criterion

Feature Name	Maximum Speed Recorded (m/s)	Minimum Speed Recorded (m/s)	SNR value Before Denoising(dB)
Wind Speed	24.87	0.6	6.22

Table 4.5 Windspeed feature details for dataset-3

Here, wind speed is crossing threshold of wind gust limit. So, we consider that wind is bounded with gusts and thus we cannot denoise the data because of its inherent property. Instead, we consider the data as it is and apply encoder decoder model directly.

2. Power

Month wise generated turbine power is as shown in fig 4.14:

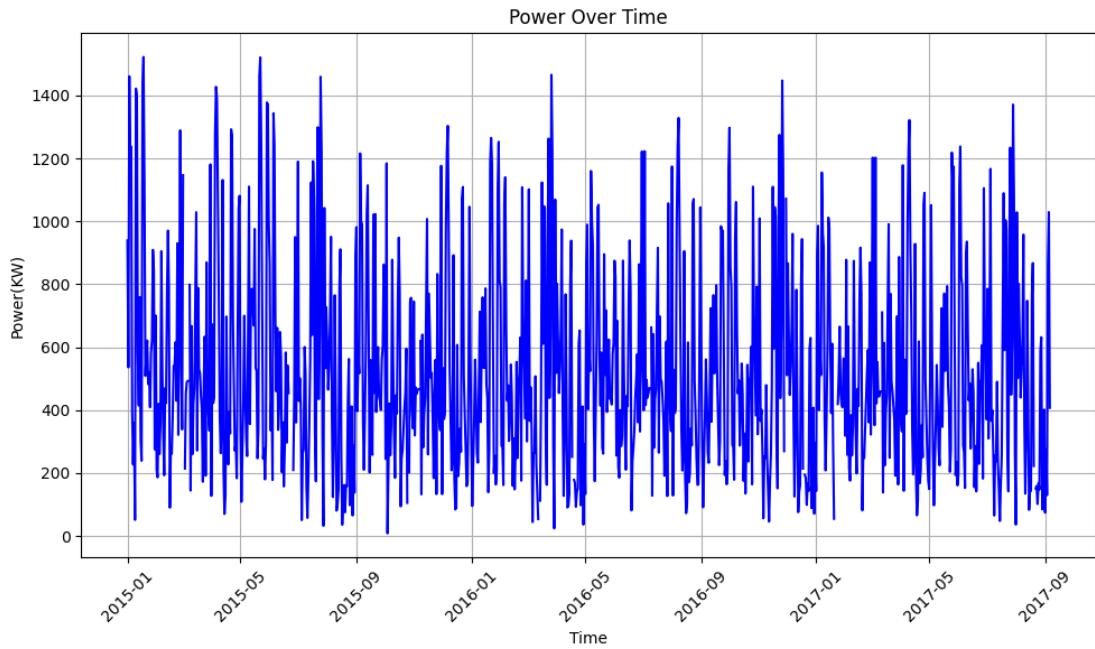


Fig 4.14: Timeseries plot of Power feature

3. Temperature:

Turbine is located in Kuwait which is high temperature region. As temperature increases, pressure difference decreases and thus wind speed decreases due to which power generation decreases.

Temperature profiling is as shown in the fig 4.15:

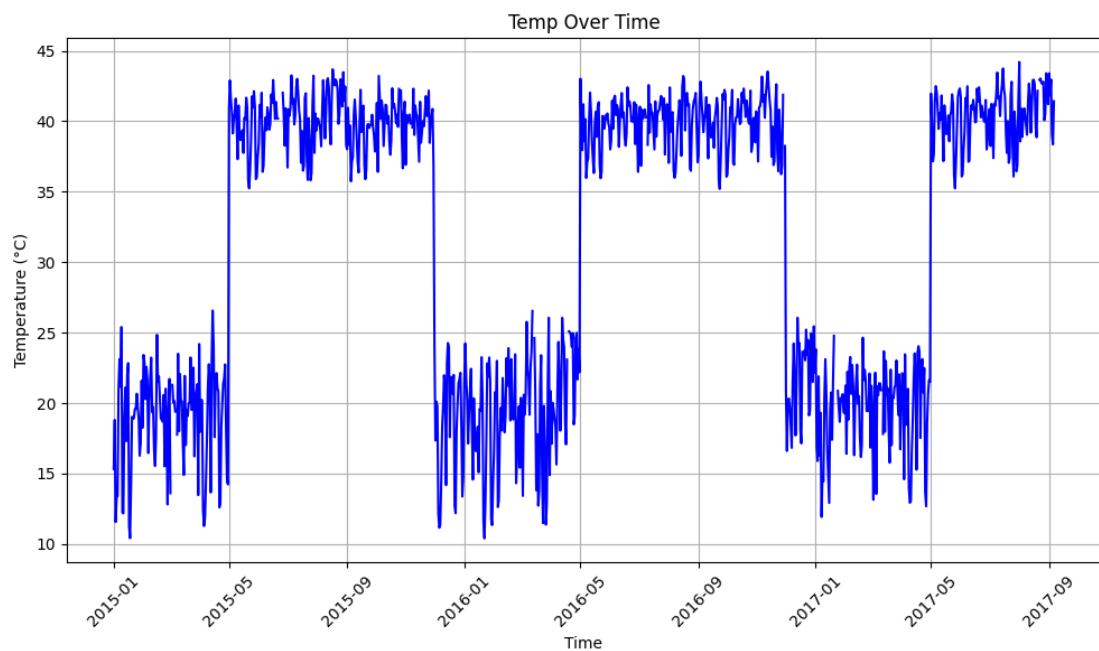


Fig 4.15 Timeseries plot of Temperature feature

4. Humidity

Humidity has an inverse relationship with temperature. Its definition and characteristics has already been discussed. Profiling of humidity along with timestamps is as shown:

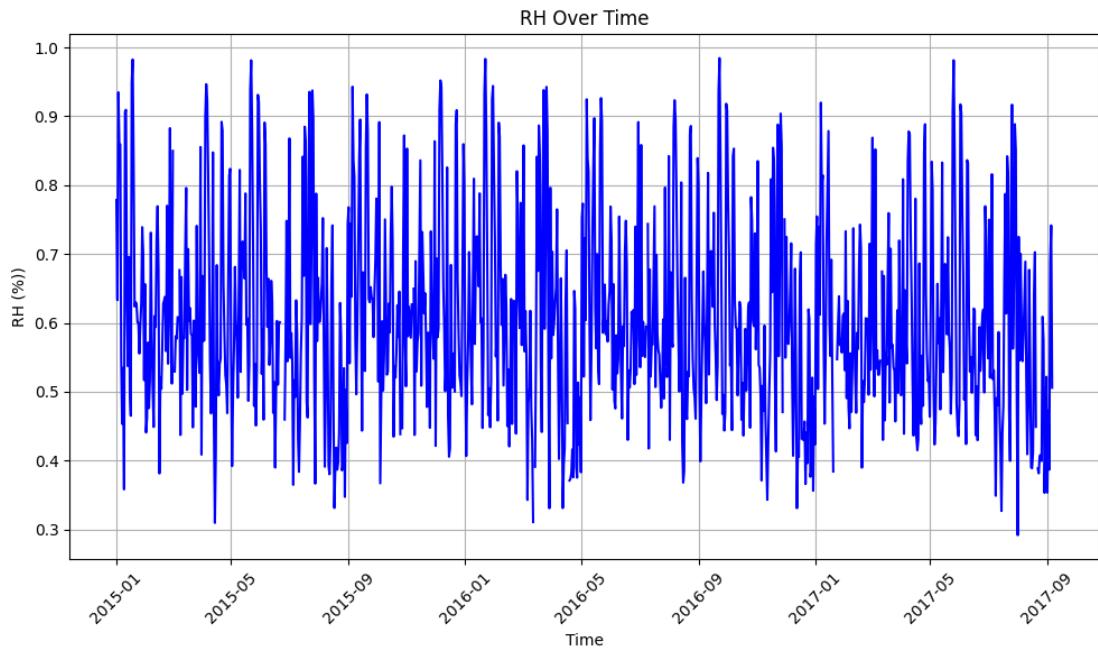


Fig 4.16 Timeseries plot of Humidity feature

4.2 Proposed Approach:

DATASET-1

This thesis is inspired by the experiments done using Neural network models. So, mainly it will focus on the implementation of LSTM and its variants like Stacked LSTM

Now let us go and discuss how these models will be involved in our thesis for power generation forecasting.

Wind turbine power is mainly affected by wind speed. Their high correlation gives an idea of how power generation can be effective in the seasonal periods of wind speed- (April to October). As mentioned, the data collected is in 10 min intervals.

PART-A]

So, our primary focus is 1 step ahead forecasting, which takes one week before (1008 time steps) wind speed and power generation as input and predicts the 1009th time step.

Further, Our sliding window will be 1 step ahead only, and forecasted output will feed back as input to the next observation. For example, our input will be 0th to 1008th time step data to model and 1009th time step power will be forecasted. Now this forecasted output will feed back as input (from 1st to 1009th data). 1009th observation is a forecasted observation of wind power. In this way, the 1010th observation will be forecasted, and this window continues till the last observation.

No. of data points can adjusted based on our requirement of training. Instead of the previous 1 week of data, the previous day of data can also be trained to predict the next time step.

The whole idea is shown in the fig below:

						PREDICTED POWER
INPUT	WIND SPEED	t=0	t=1	t=2t=1008	
	POWER GENERATION	t=0	t=1	t=2t=1008	t=1009
INPUT	WIND SPEED	t=1	t=2	t=3t=1009	
	POWER GENERATION	t=1	t=2	t=3t=1009	t=1010

Fig 4.17: Data input for next time step prediction using forecasted power in subsequent window

PART-B]

What if we want to perform multi-step ahead forecasting? If so, we need to forecast all other impacting features as well. As we pass wind speed and power generation past values as input to our model, it is necessary to forecast wind speed to give justice to multi-step power generation forecasting.

The experiment will start from one step ahead window only; further, we can enhance it for multiple steps.

So, considering this fact, we are providing wind speed and power generation past values as input to the model (1008 time steps data) and forecasting both wind speed and power generation for the 1009th time step. Now, for the next sliding window (1st to 1009th-time steps data will be actual, and 1009th-time step observation will be

forecasted and append to past series for both wind speed and power generation and finally, 1010^{th} observation will be forecasted.

The schematic understanding of the above wordings is shown in fig

							PREDICTED WIND	PREDICTED POWER
INPUT	WIND SPEED	t=0	t=1	t=2t=1008		t=1009	t=1009
	POWER GENERATION	t=0	t=1	t=2t=1008			
INPUT	WIND SPEED	t=1	t=2	t=3t=1009	←	t=1010	t=1010
	POWER GENERATION	t=1	t=2	t=3t=1009	←		

Fig 4.18– Data input for next time step prediction of both wind speed and power using forecasted power in the subsequent window

In this way, data will be forecasted for all observations.

In subsequent experiments, multi-step ahead window will be considered, and the target will be forecasted.

Methodology for Dataset-2 and dataset-3:

Dataset 2 mostly deal with colder climatic locations and hot climatic regions and hence, features like ambient temperature and humidity are essential to consider.

When we go for a multi-step ahead forecasting, all the historical-past observations must be fed to the model, and resulting in all observations until the next step is forecasted.

Here is the required schematic to represent the formulation idea:

							PREDICTED WIND	FUTURE TEMPERATURE READINGS	FUTURE HUMIDITY READINGS	PREDICTED POWER
INPUT	WIND SPEED	t=0	t=1	t=2t=1008		t=1009	t=1009	t=1009	t=1009
	TEMPERATURE	t=0	t=1	t=2t=1008					
	HUMIDITY	t=0	t=1	t=2t=1008					
	POWER GENERATION	t=0	t=1	t=2t=1008					
INPUT	WIND SPEED	t=1	t=2	t=3t=1009	←	t=1010	t=1010	t=1010	t=1010
	TEMPERATURE	t=1	t=2	t=3t=1009	←				
	HUMIDITY	t=1	t=2	t=3t=1009	←				
	POWER GENERATION	t=1	t=2	t=3t=1009	←				

Fig 4.19– Data input for next time step prediction of impacting features

- **Mathematical Formulation of data feeding**

The discussed models take temporal training data as an input in a recursive manner and give output as per the forecasting window.

Let:

- y_t be the observed value at time step t.
- y'_t be the predicted value at time step t.
- X_t be the input features (including lagged values of y, exogenous variables, etc.) at time step t.

Now train the model using first n observations ($t=1,2,\dots,1008$) as discussed.

$$y'_{n+1} = f(X_{1:n})$$

After the first prediction, y'_{n+1} will be taken back as input for the next set of forecasting.

General Recursive Rule:

$$y'_{n+1} = f(X_{t-m+1:t})$$

Where, $X_{t-m+1:t}$ represents the most recent time steps of the features, including both observed and predicted values.

Proposed Structure of Methodology:

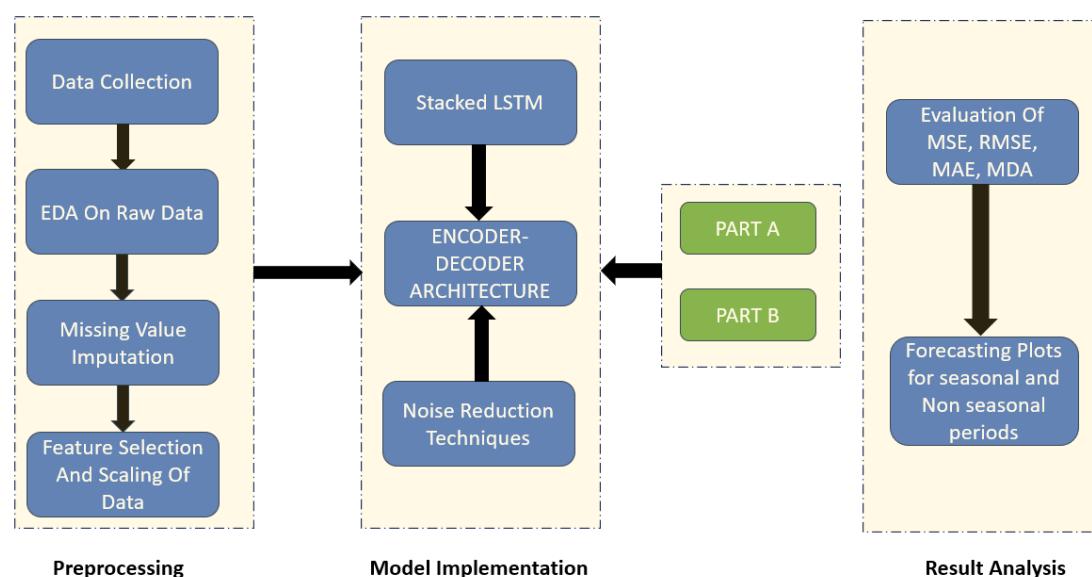


Fig 4.20– Proposed Structure of Methodology

Now lets understand the methodology to handle noise in the data:

- Initially, it is necessary to check whether any kind of wind gusts are present in the data and, if yes, at what intervals. Based on their presence, it will be considered as an inherent property of wind and cannot be considered as noise.
- But if wind gusts are missing from the data, then SNR will be calculated to check the noisiness of the data.

The flow chart below will help in deciding the presence of wind gusts in the data. As per [34], ranges have been decided.

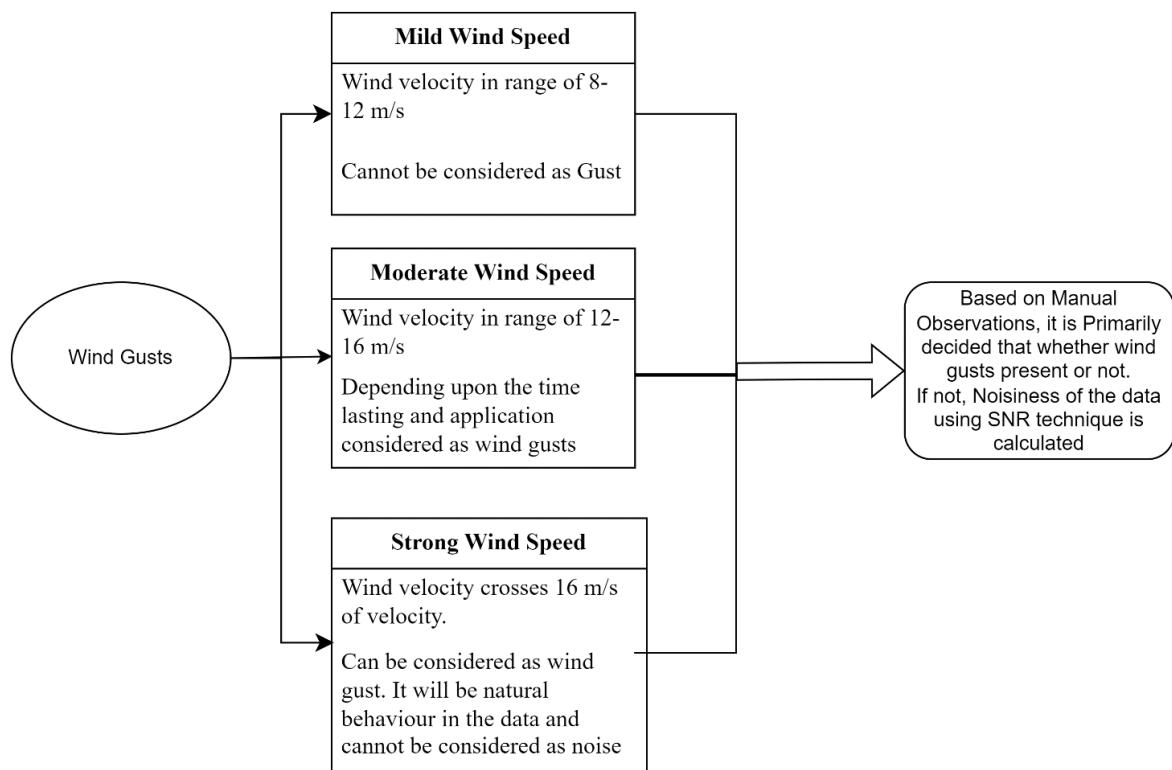


Fig 4.21– Wind Gusts analysis flow chart

Workflow to denoise the data after checking the presence of wind gusts.

Note:

Methodology varies with windmill location and separate training needs to be performed as environmental conditions are not unique.

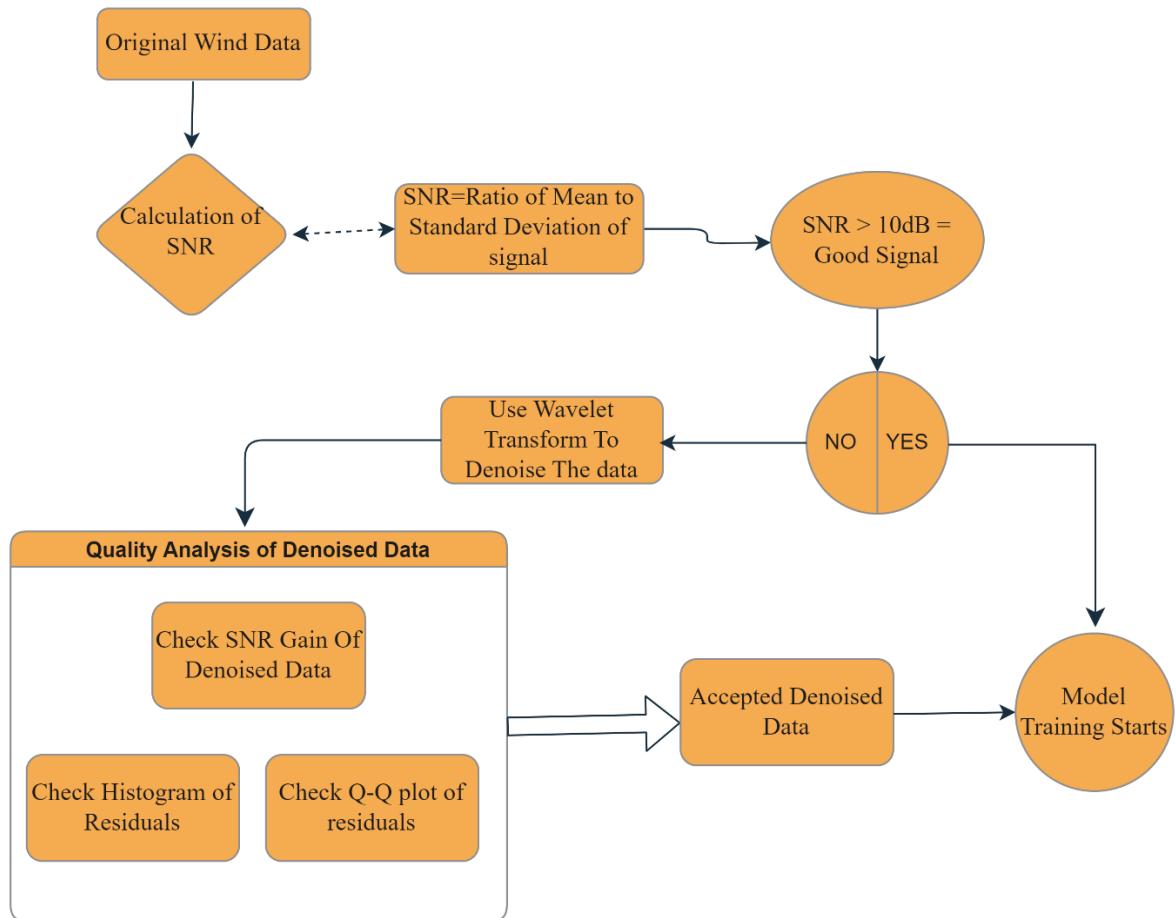


Fig 4.22– Workflow to handle Noisy Data

CHAPTER 5

IMPLEMENTATIONS AND RESULTS

This chapter will provide a detailed description of various methods that have been implemented for forecasting and the results obtained from various formulations.

5.1 IMPLEMENTATIONS OF PROPOSED METHOD:

From correlation charts and multivariate analysis, it is found that for the generation of active power, wind speed plays a vital role, and the contribution of other features can be neglected. So, the target variable ‘Wind power’ will be forecasted based on Wind speed data.

As Power generation data is nonlinear/nonstationary, It became necessary to check for the stationarity of the data involved. ADF and KPSS methods can be used to check stationarity, along with visual inspection.

From visual inspection, it is found that data is nonstationary and involves seasonality.

The figure shows all the components like trend, seasonality, and residuals.

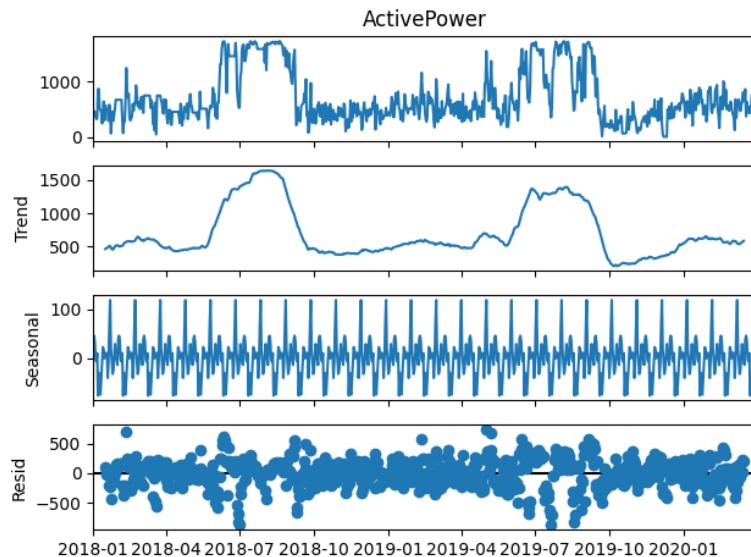


Fig 5.1 : Trend, Seasonality and Residuals of Time series data

ADF and KPSS results are contradictory with each other.

5.1.1 Power generation forecasting using SARIMAX

ADF TEST STATISTICS RESULTS:

Here, ADF test statistics are more negative than the significance level 0.05, i.e., 5%. Also, the p-value is < 0.05 , which shows complete evidence against the null hypothesis.

The null hypothesis in the case of ADF is that the series is nonstationary.

Alternate hypothesis: The Series is stationary

The collected evidence is sufficient to reject the null hypothesis and conclude time series data is stationary.

KPSS TEST STATISTICS RESULTS:

Another test statistic performed is KPSS, whose testing criterion is exactly the opposite of the ADF test.

As per the KPSS test, if the test statistic is more than a significant level (here it is 0.05), and the p-value is < 0.05 , then we reject the null hypothesis.

Null hypothesis: Time series data is stationary

Alternate hypothesis: Time series data is nonstationary.

So, the KPSS test suggests that data is nonstationary.

Conclusion:

ADF results and KPSS results are contradictory.

ACF and PACF plots confirms the non stationarity in the data.

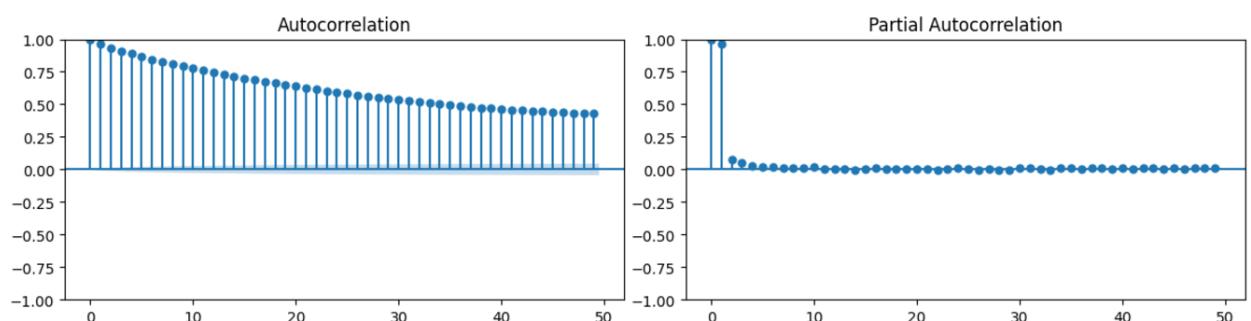


Fig 5.2: ACF and PACF plots for non stationary series

To check for traditional models like SARIMAX, we need to pass the stationary series. So, differencing the series is needed.

Why SARIMAX?

Obtained data has a seasonal component with multiple features as input. To overcome the problem of seasonality, we go with sarimax.

After the first difference, the overall seasonal trend gets removed, and the series becomes stationary.

ACF and PACF plots obtained are shown in fig:

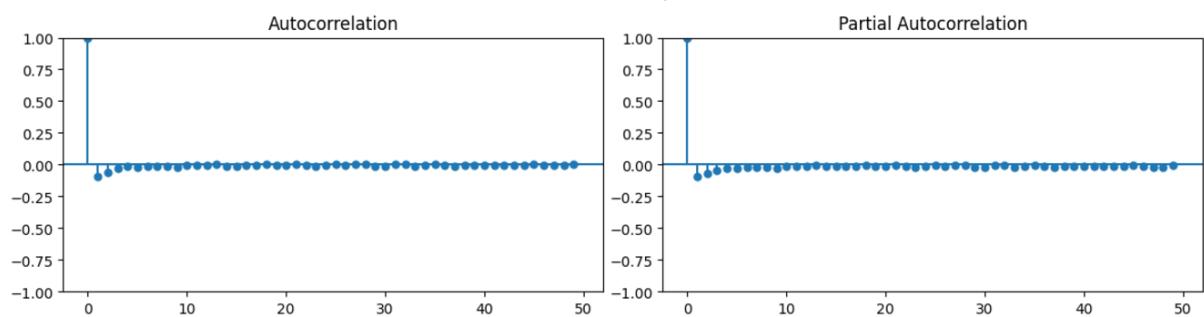


Fig 5.3: ACF and PACF plots for differenced series

From ACF and PACF plots, we can further decide the order of AR and MA

ACF plot helps in deciding the order of AR model(p) and PACF can be used in determining MA (q)

From the plots obtained, order chosen is $(1,1,1)(1,0,1)$ for sarimax model. The model is decided based on the least AIC value obtained.

Also, residual analysis is done for the selected model:

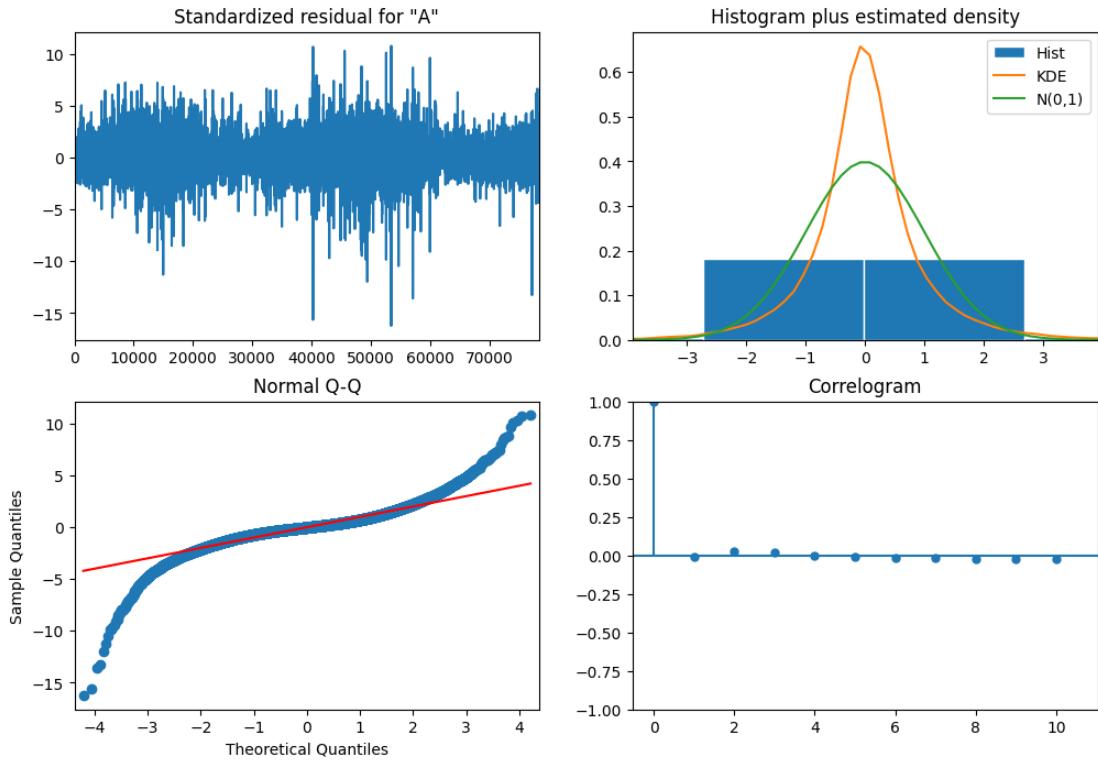


Fig 5.4: Model Residual Analysis

Results are discussed in the 5.2 section.

5.1.2 Power generation forecasting using stacked LSTM architecture

We have many statistical-based methods for forecasting operations like ARIMA, SARIMA, etc. However, these methods will not perform nor give better forecasting for highly noisy, non-linear data. Experiment results with SARIMAX are tabulated above.

LSTM architecture is mainly used to tackle the vanishing gradient problem in sequential data. To learn long-term dependencies of sequential data, i.e. here, time series data, and to capture relevant information, LSTM architecture has been implemented.

Part A: Single feature forecasting

The formulation of LSTM architecture is explained in the 3.2 problem formulation section.

As explained earlier, train and test data samples are generated using the desired sliding window.

In formulated LSTM architecture, one input layer, two hidden layers, and one output layer have been used.

$X_{i,t}^j$ = Input data xtrain with size (51787,1008,2) 3D array at timestamp t

$Y_{i,t}^j$ = Output data ytrain with size (51787,1008,1) 2D array at timestamp t

The summary of LSTM architecture is as shown below:

Parameter Calculation Formula: (Number of neurons * Hidden neurons) + Bias

STACKED LSTM MODEL					
Layer Name	No Of Neurons	No Of I/P Features	No. Of FFN	Parameter Calculation	Total Parameters
Embedding Layer	128	2	4	$((128+2)*128 + 128)*4$	67072
LSTM 1st Layer	128	2	4	$((128*128)+(128*128) + 128)*4$	131584
LSTM 2nd Layer	64	2	4	$((128*64)+(64*64) + 64)*4$	49408
Dense Layer	64	1	-	$64*1 + 1$	65
Total Trainable Parameters					248129

Fig 5.5 : Stacked LSTM model summary

1st and 2nd layer has 128 neurons each and 3rd layer has 64 neurons and h_t the output from each cell will be transferred to the LeakyRELU activation function. LeakyRELU is mainly used to tackle the problem of dying RELU. Dying RELU appears when we keep our learning rate high, and thus, it converts our function negative, which further deactivates that neuron.

To avoid this problem, LeakyRELU brings nonlinearity in the activation, and a small gradient will always flow in each neuron.

Finally, the dense layer gives output, which is the final prediction for that observation. After fitting the model, the loss function is computed between the hidden state output and $Y_{i,t}^j$

Part B: Multivariable forecasting:

As discussed in methodology, it is necessary to consider impacting features on the target variable for multi-head forecasting. Here, wind speed impacts power generation, so wind speed forecasting is considered in this section using stacked LSTM architecture.

The detailing of the architecture is shown below:

STACKED LSTM MODEL FOR MULTIVARIABLES					
Layer Name	No Of Neurons	No Of I/P Features	No. Of FFN	Parameter Calculation	Total Parameters
Embedding Layer	128	2	4	$((128+2)*128 + 128)*4$	67072
LSTM 1st Layer	128	2	4	$((128*128)+(128*128) + 128)*4$	131584
LSTM 2nd Layer	64	2	4	$((128*64)+(64*64) + 64)*4$	49408
Dense Layer	128	2	-	$128*1 + 2$	130
Total Trainable Parameters					248194

Fig 5.6 : Stacked LSTM model summary for multivariable

$X_{i,t}^j$ = Input data xtrain with size (51787,1008,2) 3D array at timestamp t

$Y_{i,t}^j$ = Output data ytrain with size (51787,1008,2) 2D array at timestamp t

The parameter's calculation is same as mentioned in single-feature forecasting.

5.1.3 Power Generation Forecasting using Encoder-Decoder Architecture:

The mathematical formulation of encoder-decoder architecture has already been explained in section 3.3.

The architecture of the encoder-decoder algorithm is as given below:

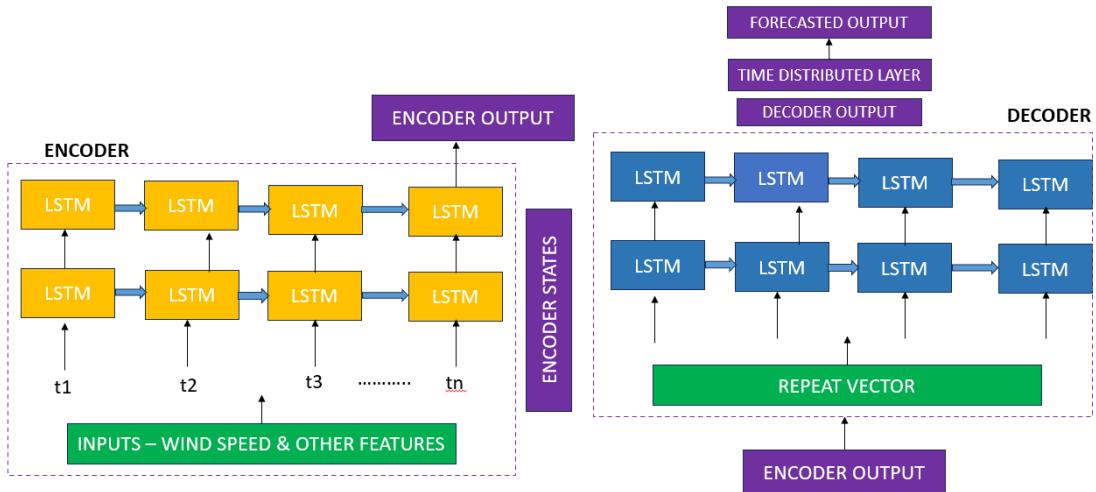


Fig 5.7: Formulated Architecture of Encoder Decoder algorithm

The architecture shown in the figure is used here to forecast wind power and wind speed. The detailing of the architecture is as given in Fig 5.5:

ENCODER DECODER ARCHITECTURE					
Layer Name	No Of Neurons	No Of I/P Features	No. Of FFN	Parameter Calculation	Total Parameters
Embedding Layer	100	2	4	$((100+2)*100 + 100)*4$	41200
LSTM 1st Layer	100	2	4	$((100*100)+(100*100) + 100)*4$	80400
LSTM 2nd Layer	100	2	4	$((100*100)+(100*100) + 100)*4$	80400
LSTM 3rd Layer	100	2	4	$((100*100)+(100*100) + 100)*4$	80400
Dense Layer	100	2	-	$100*2 + 2$	202
Total Trainable Parameters					282602

Fig 5.8: Encoder-Decoder architecture summary

5.2 Results :

Defined LSTM architecture is trained by considering the following parameters:

Optimizer used = Adam

Learning rate= 0.001

Loss metrics = mean squared error

Batch size = 32

No of epochs = 20

Validation split = 0.1

The validation and training loss is observed and displayed in Fig 5.6

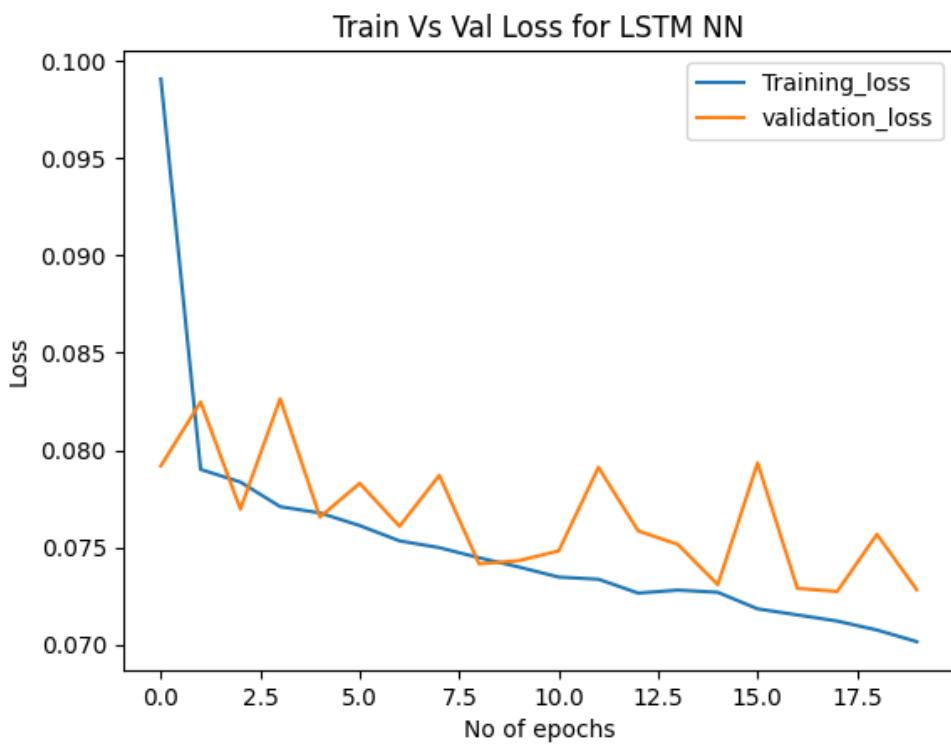


Fig 5.9- Training and validation loss of stacked LSTM architecture

As discussed in methodology, this thesis has focussed on Part A and Part B.

Part A]

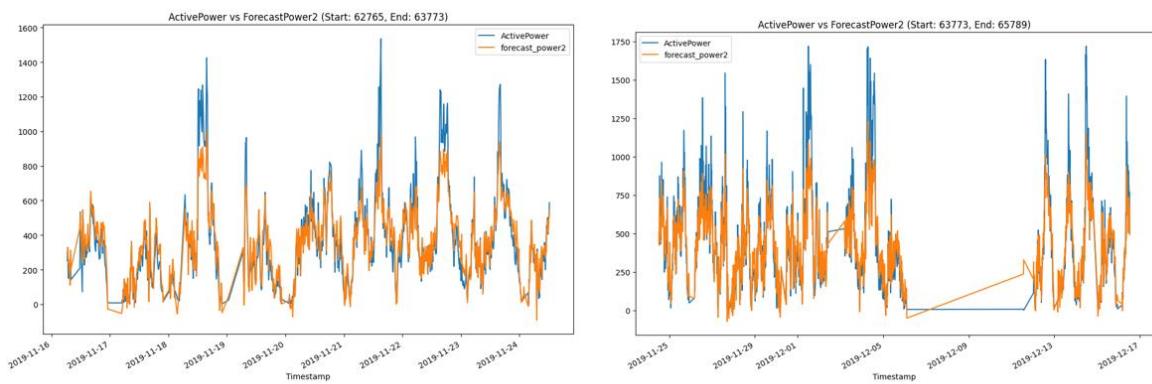
5.2.1 Results from Sarimax model:

To check the efficiency of the traditional model, Sarimax model is trained and results are generated. Here, the sample results are shown and based on obtained metrics the performance of the model is decided.

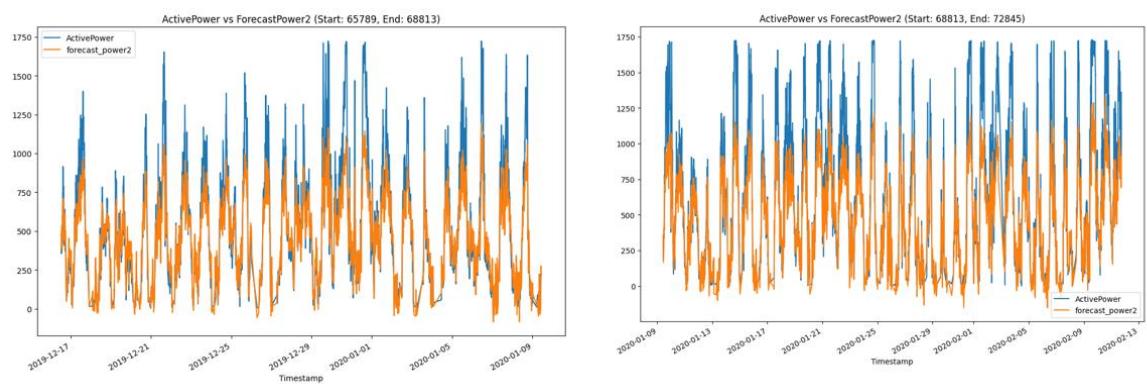
Blue Line= Original Power

Orange Line= Forecasted Power

Nov-Dec:



Jan-Feb:



March:

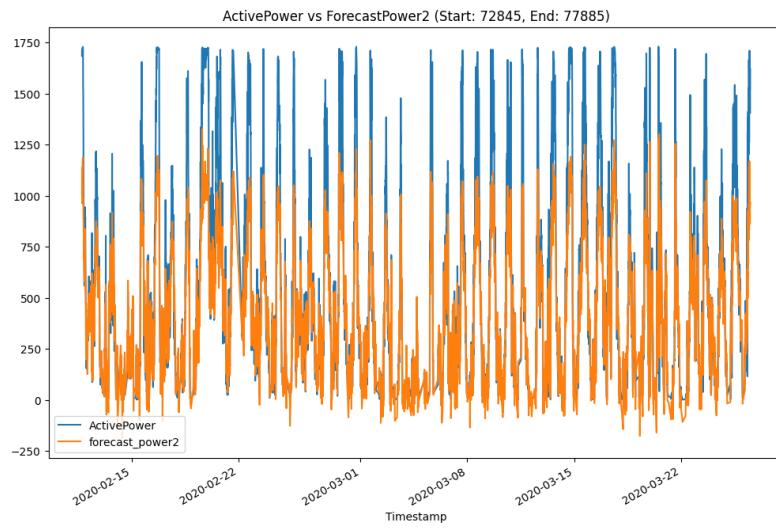


Fig 5.10: Forecasting Plots for SARIMAX model

Performance Metric Results:

Month	RMSE	MSE	MAPE	MDA
Nov	294.33	86628.13	88.67	85.80
Dec	136.93	18750.92	1838.74	84.86
Jan	362.50	131409.38	42.64	80.98
Feb	449.95	202459.36	60.17	82.96
March	236.89	56114.54	81.87	81.88

Table 5.1: Metrics Results for Sarimax model

Sarimax is a statistical method which extends the ARIMA model by incorporating the seasonal components. Using exogeneous variables we can offer an effect of other variables on target one.

Through literature survey it is highlighted that to capture hidden information across multiple time steps statistical models are not effective.[\[47\]](#)

Disadvantages of Sarimax:

1. Sarimax assumes that the data is stationary; if not, we need to make it stationary.
2. It is effective for linear variations in the variable.
3. Sarimax is computationally intensive for large datasets

To handle complex non linear data(Here, wind speed and power generation has cubic relationship), research on NN is done. Results of Deep NN architectures is explained in subsequent sections.

5.2.2 Results obtained from stacked LSTM model.

The below results are sub-part of Part A, where a step window size has been taken for prediction using wind speed and generated power as input, and the target feature will be generated power only.

The results are captured for alternate weeks from August 2019 to January 2020, which also covers the seasonal period.

The displayed plots indicate the behaviour of forecasted wind power with actual wind power.

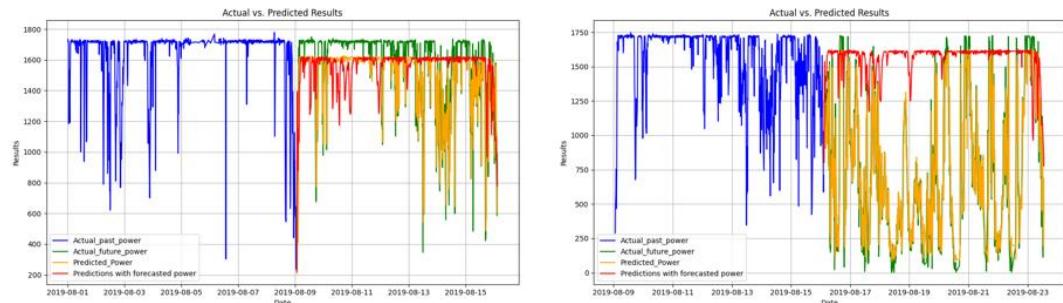
Blue colour: past week observations

Green colour: Actual future observations

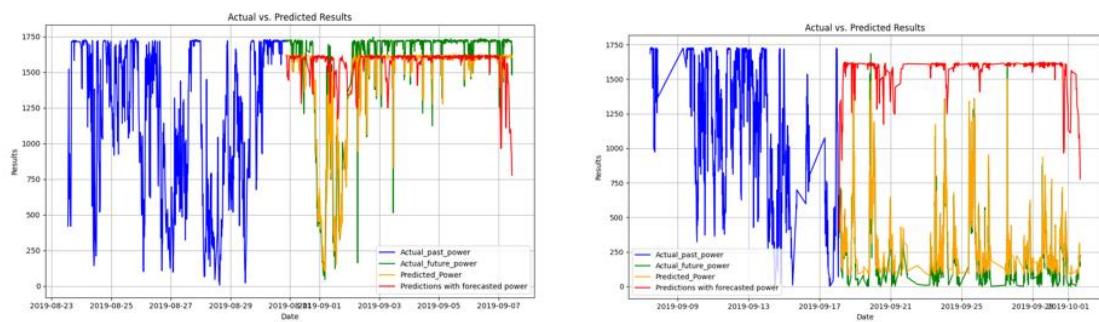
Orange colour: Forecasted observations

Red colour: Forecasted observations with forecasted output as input

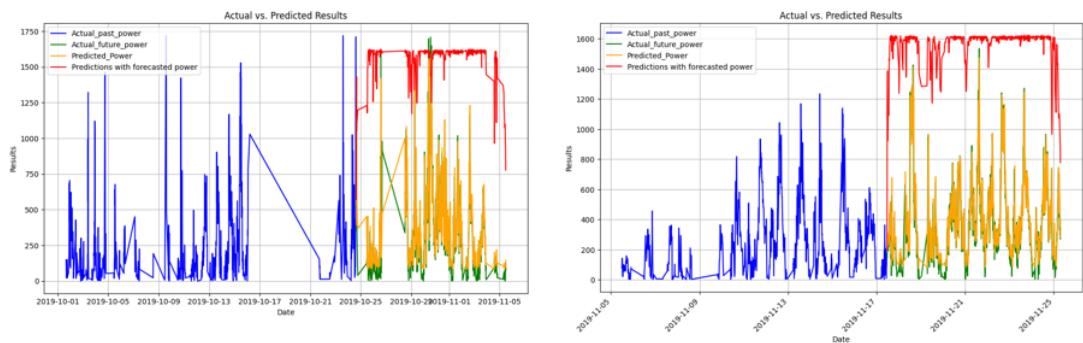
Aug:



Sept:



Oct-Nov:



Dec:

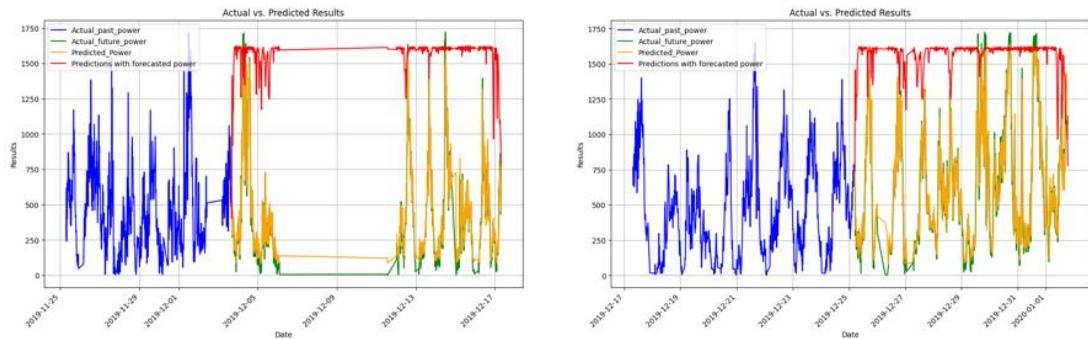


Fig 5.11 : Week wise Plots showing forecasting of wind power for each month

The model's accuracy is analyzed using RMSE, MAPE, and MDA.

Performance metrics of each forecast are tabulated below:

Month	Week	RMSE	MAPE	MDA
Aug	Week 1	192.34	11.22	0.49
	Week 3	287.12	82.21	0.46
Sept	Week 1	166.42	11.97	0.5
	Week 3	250.42	54.32	0.47
Oct	Week 1 & 3	175.73	174.5	0.48
Nov	Week 1 & 3	146.2	78.16	0.48
Dec	Week 1 & 3	148.85	99.55	0.48
Jan	Week 1	182.56	58.4	0.49

Table 5.2: Performance metrics of forecasted results for each month

Part B:

This section, as mentioned in the methodology, focuses on multi-step and multi-variable feature forecasting.

So, the stacked LSTM algorithm is applied considering both wind power and wind speed as input and target features. It means we are forecasting both wind speed and wind power generation to deal with multi-head forecasting.

Initially, 1 step ahead (10 min) forecasting is considered by applying a stacked LSTM algorithm.

Forecasting results are tabulated below:

WIND POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	917.22	55.07	0.48	Aug	Week 2	3.5	32.39	0.5
	Week 3	518.52	257.75	0.43		Week 3	1.67	21.7	0.44
	Week4	730.86	77.12	0.48		Week4	2.4	24.65	0.48
Sept	Week 1	941.81	65.32	0.5	Sept	Week 1	3.62	33.83	0.47
	Week 2	702.67	451.25	0.49		Week 2	2.31	23.65	0.49
	Week 3 & 4	197.39	2584.97	0.44		Week 3 & 4	1.78	38.58	0.45
Oct	Week 1 & 3	468.47	1600.73	0.47	Oct	Week 1 & 3	1.89	42.04	0.47
	Week 4	406.8	881.13	0.47		Week 4	1.5	28.74	0.47
Nov	Week 1 & 2	439.42	1866.29	0.49	Nov	Week 1 & 2	1.91	36.6	0.49
	Week 2 & 3	366.69	381.82	0.48		Week 2 & 3	1.31	2.347	0.5
	Week 4	347.71	266.16	0.48		Week 4	1.21	20.18	0.49
Dec	Week 1 & 2	415.1	730.17	0.45	Dec	Week 1 & 2	1.44	25.79	0.46
	Week 3	361.31	382.37	0.5		Week 3	1.31	22.07	0.5

Fig 5.12: Performance of stacked LSTM for multivariable one step ahead prediction

Forecasting results are not accurate for multivariable forecasting. We can even do better than this. The involvement of more LSTM layers can capture more complex patterns and thus give more accurate results. It is necessary to study the time dependencies of the target variable. Also, it is necessary to study sequential data patterns for better forecasting results. Considering these facts, encoder-decoder architecture is proposed, and multivariable forecasting is performed.

5.2.3 Results of Encoder-Decoder Architecture:

Defined Encoder-Decoder architecture is trained by considering the following parameters:

Optimizer used = Adam

Learning rate= 0.001

Loss metrics = mean squared error

Batch size = 32

No of epochs = 5

Validation split = 0.1

The results obtained are for both wind speed and power generation. Forecasting results displayed are for each month with alternate weeks.

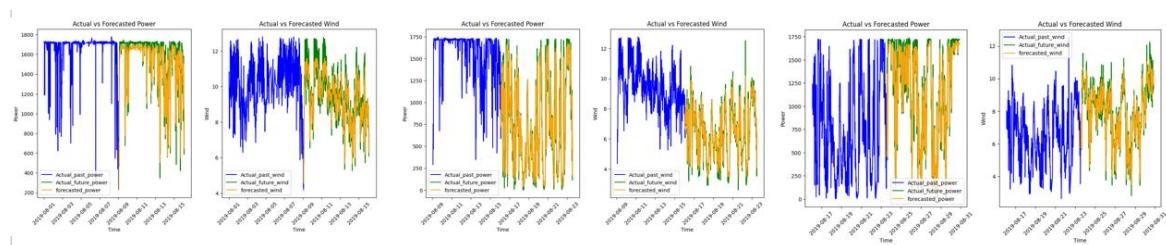
Blue colour: past week observations

Green colour: Actual future observations

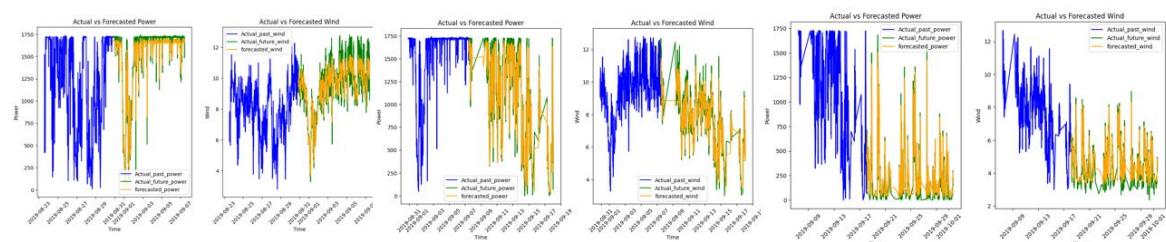
Orange colour: Forecasted observations

A] Results obtained for 1 step ahead (10min) forecasting:

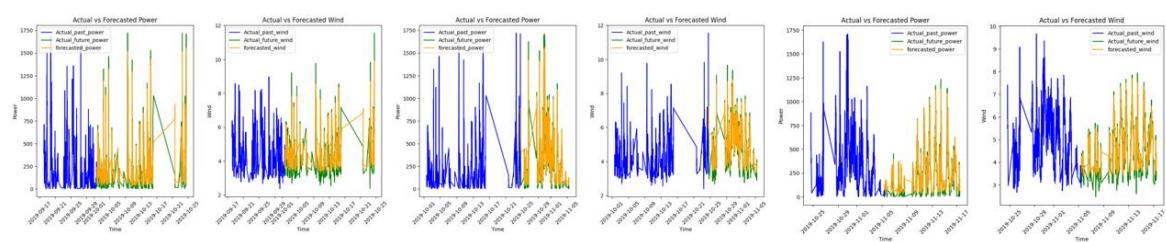
Aug:



Sept:



Oct-Nov:



Nov-Dec:

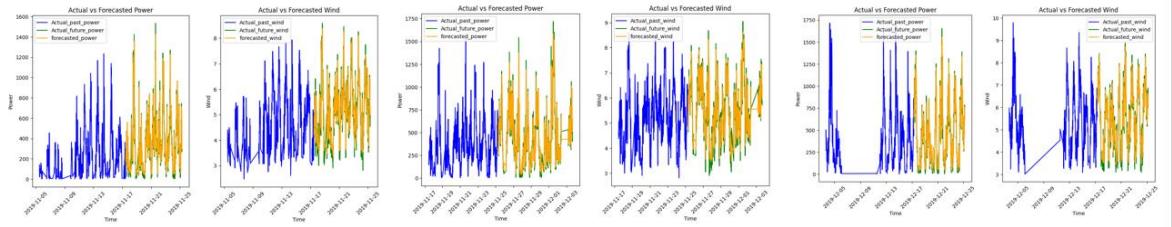


Fig 5.13: Week wise Plots showing forecasting of wind power and wind speed for each month using Encoder-Decoder architecture

Performance metrics of each forecast is tabulated below:

WIND POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	184.22	9.71	0.5	Aug	Week 2	0.89	7.46	0.51
	Week 3	292.54	77.39	0.44		Week 3	0.99	11.82	0.45
	Week4	262.74	27.51	0.48		Week4	0.9	9.38	0.49
Sept	Week 1	151.7	9.79	0.5	Sept	Week 1	0.88	7.36	0.48
	Week 2	244.2	152.55	0.49		Week 2	0.88	9	0.5
	Week 3 & 4	197.39	395.39	0.44		Week 3 & 4	0.93	13.89	0.47
Oct	Week 1 & 3	224.04	289.89	0.47	Oct	Week 1 & 3	0.93	13.89	0.47
	Week 4	177.04	136.63	0.49		Week 4	0.7	10.61	0.48
Nov	Week 1 & 2	117.62	234.31	0.5	Nov	Week 1 & 2	0.56	9.37	0.49
	Week 2 & 3	147.54	66.57	0.47		Week 2 & 3	0.59	9.07	0.48
	Week 4	173.51	58.88	0.49		Week 4	0.64	9.1	0.48
Dec	Week 1 & 2	151.85	58.35	0.46	Dec	Week 1 & 2	0.57	7.97	0.47
	Week 3	142.17	55.08	0.5		Week 3	0.54	7.92	0.5

Fig 5.14: Performance of Encoder Decoder for multivariable one step ahead prediction

Performance comparison w.r.t stacked LSTM architecture:

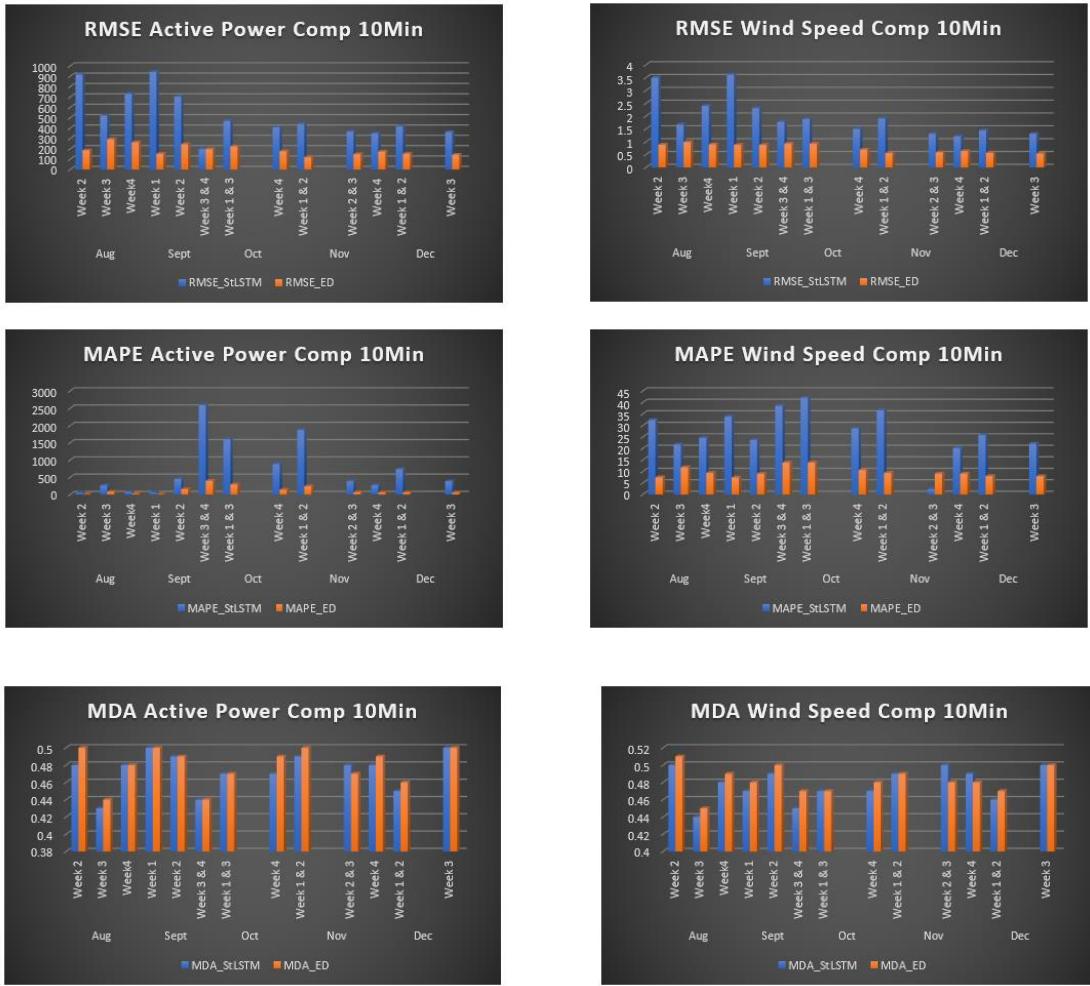


Fig 5.15: Performance comparison of stacked LSTM and Encoder-Decoder architecture for one step ahead forecasting

B] Results obtained for 3 step ahead (30min) forecasting:

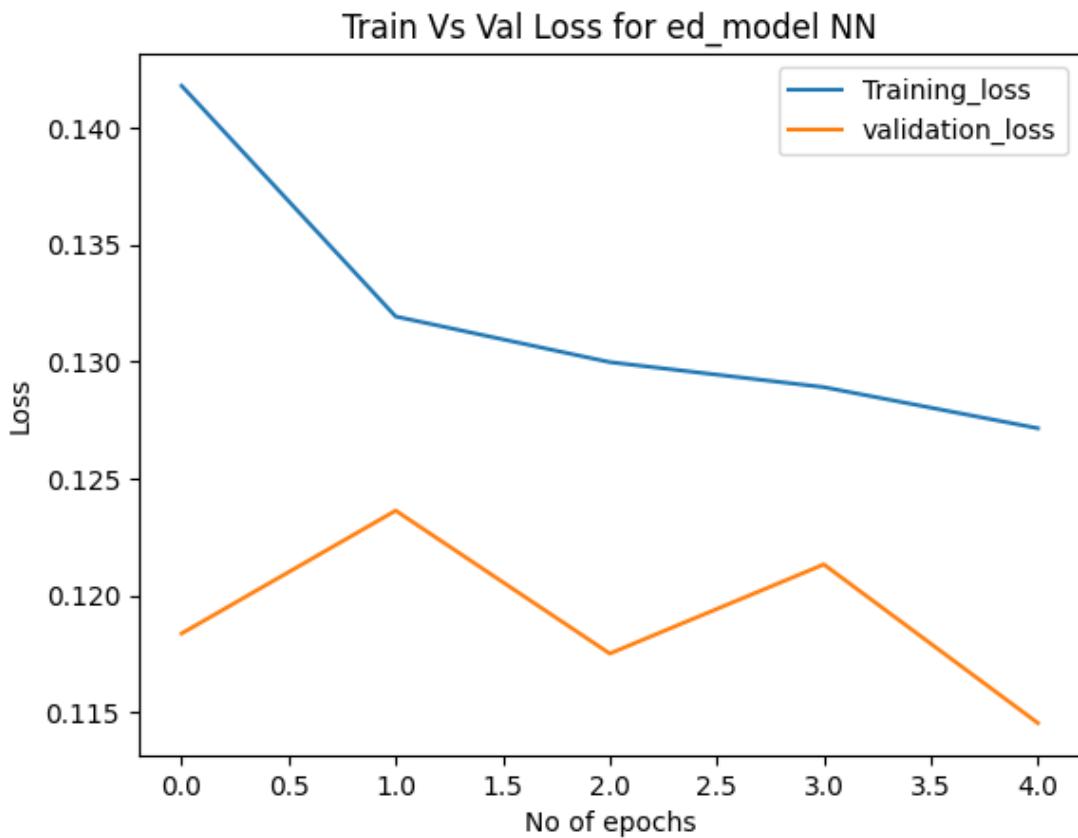


Fig 5.16: Training and validation loss of Encoder Decoder architecture for 30min ahead forecasting

WIND POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	203.63	12.35	0.46	Aug	Week 2	0.96	7.75	0.5
	Week 3	291.29	86.95	0.45		Week 3	0.98	11.81	0.46
	Week4	270.43	31.92	0.49		Week4	0.9	9.27	0.48
Sept	Week 1	185	13.61	0.48	Sept	Week 1	0.94	7.59	0.47
	Week 2	245.32	172.34	0.48		Week 2	0.86	8.59	0.49
	Week 3 & 4	197.7	596.94	0.45		Week 3 & 4	0.84	14.64	0.45
Oct	Week 1 & 3	216.32	453.32	0.47	Oct	Week 1 & 3	0.91	15.63	0.47
	Week 4	174.58	197.97	0.48		Week 4	0.69	11.12	0.47
Nov	Week 1 & 2	128.13	424.41	0.5	Nov	Week 1 & 2	0.59	10.86	0.49
	Week 2 & 3	150.04	81.38	0.48		Week 2 & 3	0.59	9.22	0.48
	Week 4	172.28	66.72	0.5		Week 4	0.63	9.1	0.49
Dec	Week 1 & 2	148.72	147.63	0.47	Dec	Week 1 & 2	0.55	8.13	0.48
	Week 3	143.98	68.62	0.5		Week 3	0.54	8.04	0.49

Fig 5.17: Performance Metrics results of wind power and wind speed 30 min ahead forecasting.

C] Results obtained for 6 step ahead (60min) forecasting:

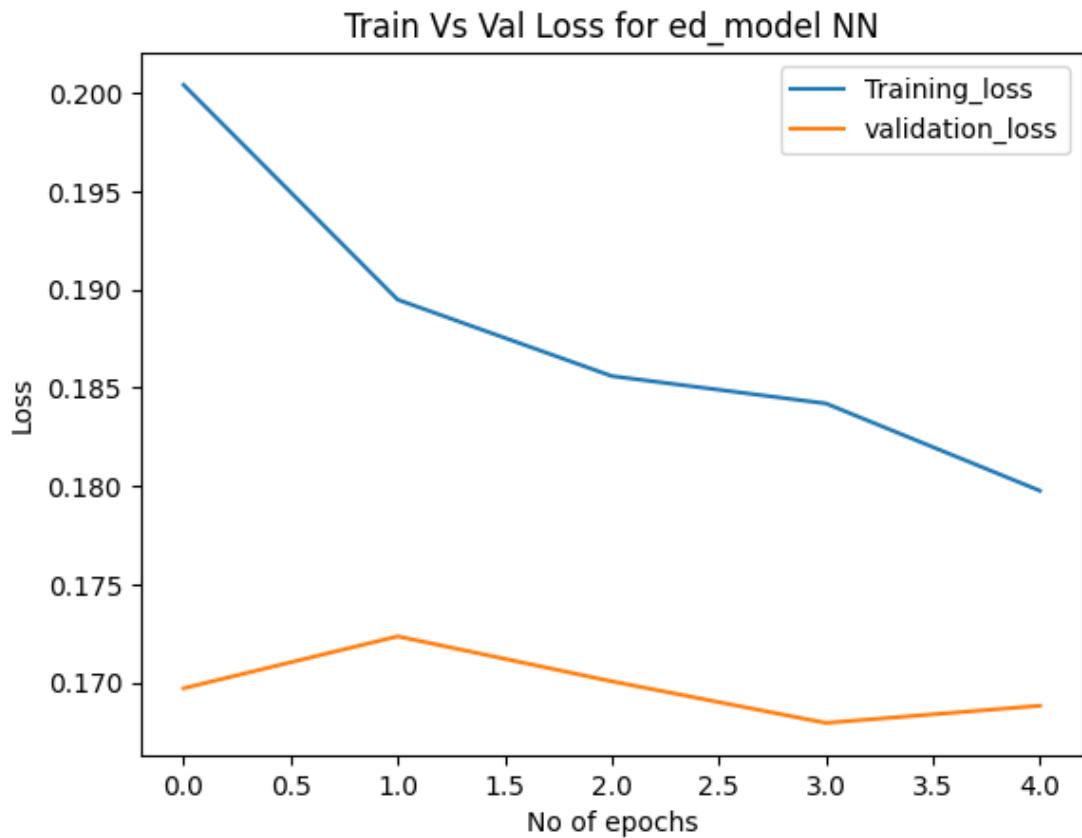


Fig 5.18: Training and validation loss of Encoder Decoder architecture for 60min ahead forecasting

WIND POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	207.94	12.69	0.49	Aug	Week 2	0.99	8	0.52
	Week 3	313.52	92.53	0.45		Week 3	1.04	12.54	0.45
	Week4	302.14	34.87	0.48		Week4	1.01	10.22	0.48
Sept	Week 1	196.79	13.86	0.49	Sept	Week 1	0.99	8.05	0.49
	Week 2	279.21	191.55	0.48		Week 2	0.94	9.44	0.49
	Week 3 & 4	210.67	735.32	0.45		Week 3 & 4	0.88	16.1	0.45
Oct	Week 1 & 3	234.68	618.94	0.49	Oct	Week 1 & 3	1	18.58	0.48
	Week 4	189.02	221.79	0.46		Week 4	0.77	12.5	0.47
Nov	Week 1 & 2	145.15	508.49	0.5	Nov	Week 1 & 2	0.68	12.95	0.5
	Week 2 & 3	163.23	81.32	0.48		Week 2 & 3	0.66	10	0.48
	Week 4	187.72	73.83	0.48		Week 4	0.7	9.92	0.48
Dec	Week 1 & 2	168.12	169.45	0.47	Dec	Week 1 & 2	0.63	9.35	0.48
	Week 3	164.19	71.5	0.49		Week 3	0.64	9.48	0.48

Fig 5.19: Performance Metrics results of wind power and wind speed 60 min ahead forecasting.

Observations:

- The Encoder-Decoder model performs better than stacked LSTM and SARIMAX for multivariable and multi-step ahead forecasting.
- From the result table, it is observed that with an increase in steps, forecasting accuracy is reduced.

Forecasting accuracy can be increased by denoising the data. The dataset has a very low SNR value, which indicates that data cannot give efficient results due to the presence of noise.

We know that time series data is captured with lots of uncertainties known as noise. The noise generation can be due to multiple factors, such as measurement errors (faulty sensors, human mistakes, or rounding errors), seasonal errors due to seasonality, and extreme measurements.

However, some of the noises are essential to highlight the characteristics of the present signal. So, how do we recognize the noise present in the signal as acceptable, or do we need to denoise it?

We calculate the signal-to-noise ratio in decibels, which can guide us in determining the acceptance of the noise.

Signal to Noise Ratio (SNR):

SNR is defined as the ratio of the mean to standard deviation of the signal.

$$SNR = \frac{\mu}{\sigma}$$

As per Rose Criterion [25], an SNR of at least 5-10 DB is needed to accept the signal; otherwise, it is treated as a highly noisy signal. Above 10 dB is considered as an average signal for further processing.

As wind power generation is mainly dependent on wind speed, Accuracy in wind speed forecasting is essential. The standard used for wind energy resource assessment,

named GB/T 18710-2002, highlights that wind farms should pass integrity checks [31]. The study did not consider effects of noise in the data.

The paper [32] briefs about identifying noise in wind speed data and denoising techniques for the better efficiency of wind turbines.

Performance indices are used to evaluate noise in wind speed data. As per [33], the SNR (Signal to Noise Reduction) technique is used to find noise in the wind speed data.

Signal Noise Ratio Gain is another metric to determine the quality of denoised data. The larger the value of GSNR, the more clear will be the wind speed data.

$$GSNR = \frac{SNR_{dn}}{SNR_n}$$

SNR_{dn} = SNR of a denoised signal

SNR_n = SNR of a noised signal

But does the presence of wind gusts be considered noisy data? This will be the inherent property of wind speed.

For that properties of wind gusts need to be understood:

Wind gusts are brief increases in wind speed, typically lasting for a few seconds. They are often associated with changing weather patterns and can be caused by various atmospheric conditions. Below are some key properties and characteristics of wind gusts:

1. **Duration:** Wind gusts usually last from a few seconds to about 20-30 seconds. They are short-lived compared to sustained winds.
2. **Speed:** Wind gusts are typically 20% to 40% stronger than the sustained wind speed. In some cases, gusts can be significantly higher, especially during storms or extreme weather conditions.
3. **Occurrence:** Wind gusts can occur due to turbulence, convection, or rapid changes in weather fronts. They are often experienced during storms, cold fronts, or when wind flows over uneven terrain like mountains.

4. **Measurement:** Wind gusts are measured using anemometers, which capture the peak wind speed within a specific time frame, often 3 to 5 seconds.
5. **Forecasting:** Weather forecasting models predict gusts based on factors like pressure gradients, terrain, and atmospheric stability. However, accurately predicting gusts is challenging due to their localized and transient nature.
6. **Relation to Wind Shear:** Wind gusts are often associated with wind shear, which is the variation of wind speed and direction over a short distance. This is particularly significant in aviation, where gusts can affect aircraft stability.

As per [\[34\]](#), the speed of wind gusts is determined in the below categories:

1. **Mild Conditions:** In mild weather, wind gusts typically range from 20 to 30 Kph (12 to 18 miles per hour).
2. **Moderate Winds:** Under more windy conditions, wind gusts often reach 40 to 60 Kph (25 to 37 miles per hour).
3. **Strong Winds or Storms:** In stormy or severe weather, wind gusts can exceed 70 Kph (43 miles per hour) and can even reach 100 Kph (62 miles per hour) or higher in extreme cases, such as during hurricanes, thunderstorms, or strong cold fronts.

The wind speed data used in this thesis follows the above speed ranges, and it is observed that the wind speed limit does not cross 16 m/s speed. If so, it will be considered as wind gust. Thus the sudden fluctuations in the data can be considered as noise. To confirm the noisiness in the data, the SNR technique is used as described above.

Thus, it is necessary to check the SNR of the features that are impacting target variables. An SNR test for both wind speed and power generation data was performed.

The SNR value **for wind speed is found to be 9.28 dB**, and for power generation, **it is 1.99 dB**. To reduce unwanted noise, data must be denoised. Denoising can be performed by smoothening techniques. Smoothening techniques like:

exponential smoothening can dampen high-frequency signals by revealing hidden trends.

Filtering techniques like Fourier transform, FFT, and wavelet transform can dampen specific frequencies that are responsible for noisy signals.

These techniques were implemented on the original dataset, and the model has been trained further on smoothed data.

The formulation of various smoothing techniques is explained below:

5.2.4 Implementation of Denoising Techniques:

1] Simple exponential smoothing:

This method assumes that the series has no trend or seasonality. The formula for simple exponentials smoothing is expressed as:

$$S_t = \alpha x(t) + (1 - \alpha)S_{t-1}$$

S_t = Smoothed value at time t

$x(t)$ = Observed value at time t

S_{t-1} = Previous Smoothed value at time $t - 1$

α = Smoothing value between 0 and 1

Smoothing value helps control the weight assigned to the current and previous observations.

High value of α = More weightage to the current observation

Lower value of α = More weightage to the past observations.

Smoothing was done for different alpha values for wind speed data and power generation. The respective plots are displayed below:

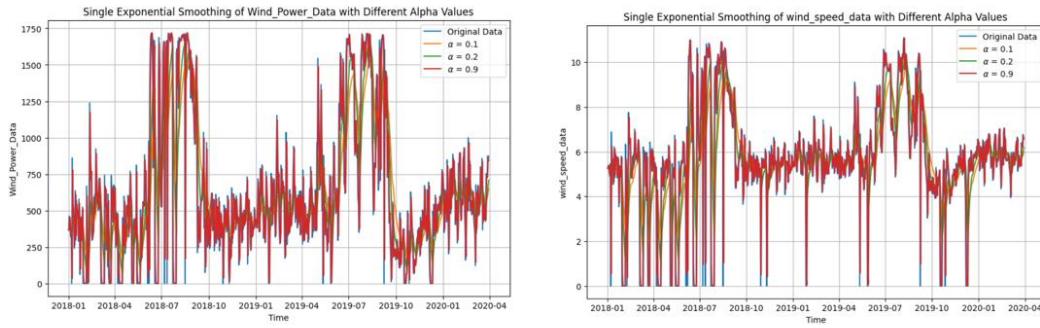


Fig 5.20: Smoothing plot for simple exponential with different alpha values

From various alpha values, the best one needs to be chosen. Forecasting has been done using simple exponential smoothing for various alpha values, and the one with the least MSE is considered the best alpha value.

The best alpha value for both wind speed and power is 0.1.

To check the accuracy of the data after smoothing, we perform residual tests, which focus on the distribution of data and Q-Q plot.

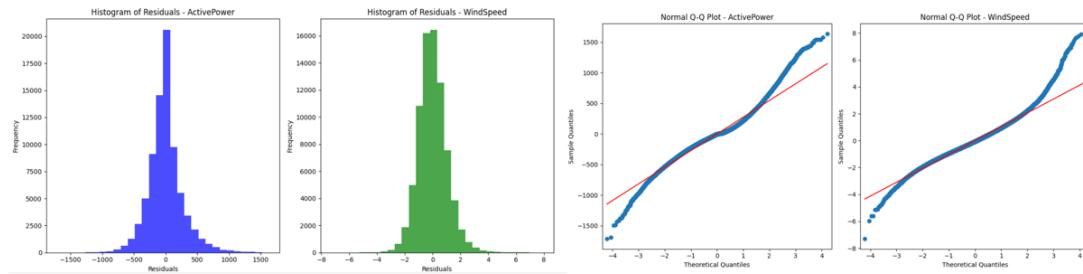


Fig 5.21: Residual performance check for simple exponential smoothing

The data, after smoothing is normally distributed for both the features and Q-Q plot highlights data points at tail and head region are dispersed. Ideally, these points should be on the same line. This indicates that the distribution of original data and smooth data are a bit different from each other.

2] Holt's linear exponential smoothing:

It can be used to find trends in the data but not seasonal factors. In such a scenario, we introduce two smoothing parameters α and β . α for the intercept and β for the trend.

Double exponential smoothing can be expressed as:

$$S_t = \alpha x(t) + (1 - \alpha)(S_{t-1} + b_{t-1})$$

$$\beta_t = (\beta)(S_t - S_{t-1}) + (1 - \beta)b_{t-1}$$

b_t = Slope with best estimate of the trend at time t

β = Smoothing parameter for trend

Here, we deal with the trend. So, for different alpha and beta values, smoothing is verified.

The best alpha and beta values among various ranges are found to be 0.1 and 0.1 resp.

For the best alpha and beta values, smoothing is done, and the accuracy of the data is verified using residual test analysis.

Respective histogram plots and Q-Q plots for both the features are displayed:

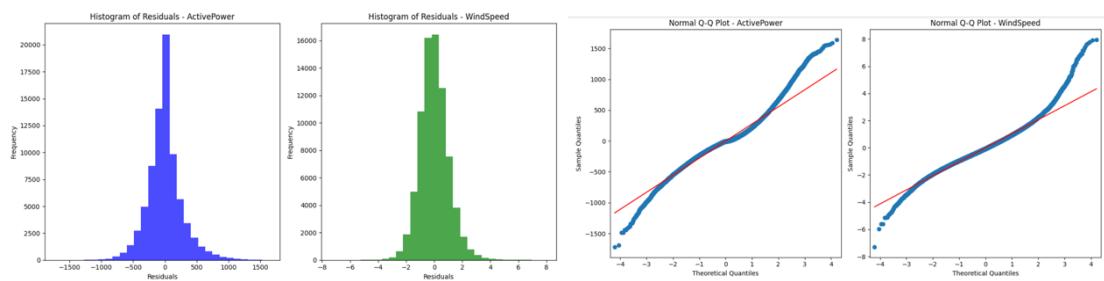


Fig 5.22: Residual performance check for Double exponential smoothing

The residuals are normally distributed, but tail and head region data points are away from the Q-Q plot line. The characteristic of data for some data points is differing.

3] Holt Winters' exponential smoothing:

It is used to smoothen the series, which has both trend and seasonality. It deals with three parameters, α for level, β for trend and γ for seasonality.

Triple exponential smoothing can be formulated as:

$$S_0 = x_0$$

$$S_t = \frac{\alpha * x_t}{c_{t-L}} + (1 - \alpha)(S_{t-1} + b_{t-1})$$

$$b_t = (\beta)(S_t - S_{t-1}) + (1 - \beta)b_{t-1}$$

$$c_t = \frac{\gamma * x_t}{s_t} + (1 - \gamma)(c_{t-L})$$

S_t = Smoothing statistics i.e it is a weighted average of obs. Y_t

c_t = Seasonal component

γ = Seasonal smoothing parameter in the range of 0 and 1.

The given smoothing technique considers both trend and seasonality in the data. For different alpha values, beta and gamma smoothing and exponential forecasting are done. Among numerous combinations, the best parameters are selected.

For power generation data,

Alpha=0.3, beta= 0.1 and gamma=0.7

For Wind Speed data,

Alpha=0.5, beta= 0.1 and gamma=0.7

For the given values, smoothing is performed, and residual analysis has been done.

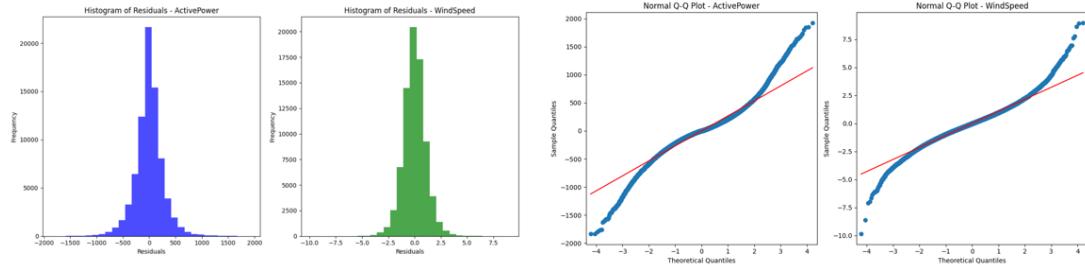


Fig 5.23: Residual performance check for Holt's Winter smoothing

The smoothing is quite complex, and the Q-Q plot highlights that smoothed data is not good for forecasting. The overall forecasting accuracy will be verified based on implemented sequential models.

4] Fast Fourier Transform:

The problem with the Fourier transform lies in its complex computation. It requires $O(n^2)$ to compute all the Fourier coefficients. So, for long-time series data, the use of Fourier transform could be expensive computation.

How does FFT help to solve this problem?

FFT reduces its complexity from $O(n^2)$ to $O(n \log n)$

Mathematically, it is represented as:

$$Z = \sum_{t=0}^{t=n-1} y_t \exp(-2 * pi * i * \frac{t}{n})$$

Where y_t = Time series data

The calculated Z coefficient need not be calculated in each term. It gets stored in the memory of the previous term. So, computation from scratch is not needed.

Fast Fourier transform is applied to data, and residual analysis is done to check the quality of smooth data.

SNR for both the features, Active Power and Wind Speed, is done, and the corresponding values are 6.71dB and 10.51dB resp. The noise is filtered from the original data.

Residual analysis is shown in fig below:

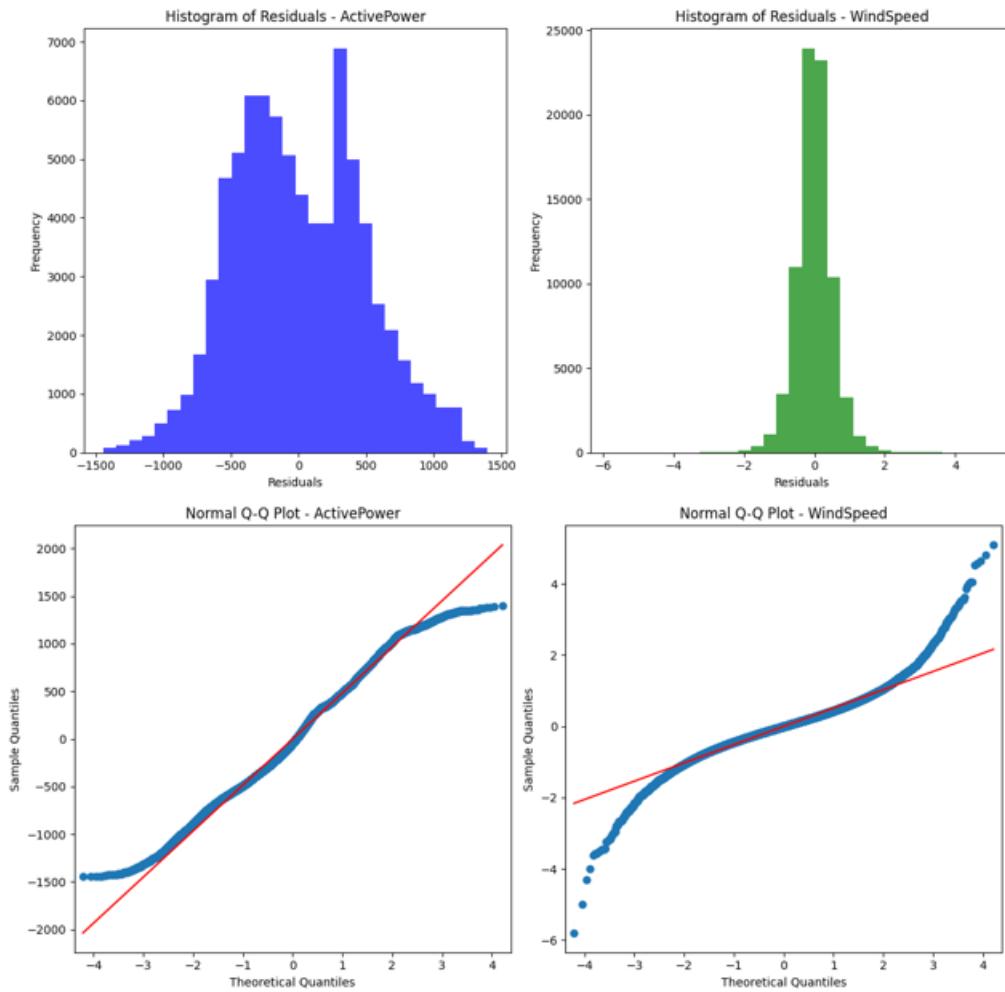


Fig 5.24: Residual performance check for Fast Fourier Transform

The residues of WindSpeed maintain a normally distributed shape, and most of the point lies on the slant line of the Q-Q plot.

5] Wavelet Transforms:

Fourier Transform provides frequency information but fails to achieve time localization. As per wavelet theory, the wavelet can retain both time and frequency information. This helps in analyzing nonstationary signals.

The overall transformation is done in three parts- wavelet decomposition, thresholding, and reconstruction.

Wavelet Decomposition: The original signal decomposed into wavelet coefficients at different scales. Wavelet function, when applied to the data, separates the signal into

low-frequency and high-frequency components. The wavelet function known as the mother wavelet is expressed as:

$$W_\varphi f(a, b) = 1/\sqrt{|a|} \int_{-\infty}^{+\infty} f(t)\varphi\left(\frac{t-b}{a}\right) dt$$

Where,

a = scale parameter. It controls the dilation or compression of the wavelet.

b = Translation parameter. It controls the shifting of wavelets along the time axis.

$f(t)$ = Original signal that we want to analyze.

Ψ_t = Mother wavelet. It is a small wave that is localized in both time and frequency.

Discrete Wavelet Transform (DWT):

It is a mathematical technique to transform time series signals into different components based on frequency. It breaks down the original signal into high-level and low-level frequency components.

Mathematically it is expressed as:

$$x[n] = \sum_k C_a[k]j_{0,k}[n] + \sum_{j=j_0}^{j=j-1} \sum_k C_d[j, k]\varphi_{j,k}[n]$$

Where Φ and Ψ are scaling and wavelet functions.

C_a = Approximation coefficients which captures low frequency components

C_d = Detailed Coefficients which captures high frequency components

$$C_A = \sum_n x[n]h[2k - n]$$

$$C_D = \sum_n x[n]g[2k - n]$$

Here.

$x[n]$ = Input time series signal

$h[n]$ = Low pass filter

$g[n]$ = High Pass filter

$2k$ = downsampling by a factor of 2.

Now, after denoising the data, the SNR value for Active Power and Wind Speed is approx. 5 dB and 12dB resp. which has effectively increased.

The residual analysis is as shown below

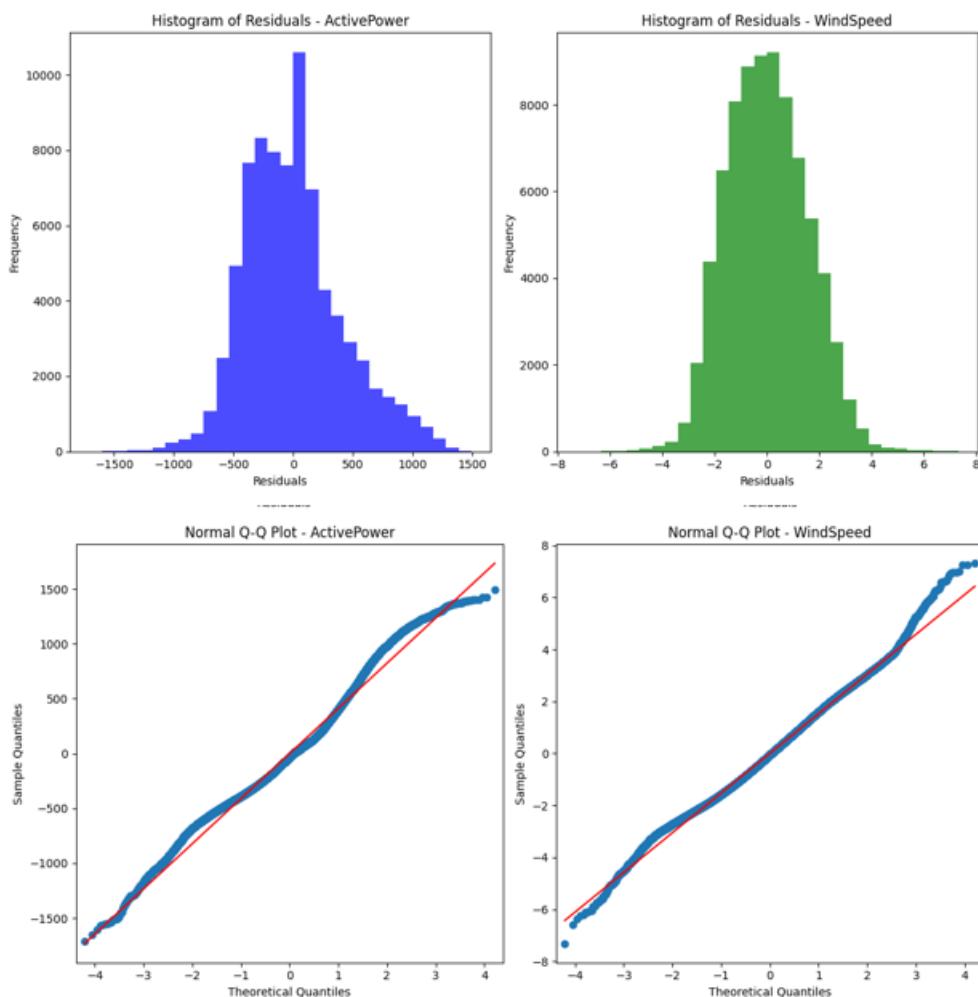


Fig 5.25: Residual performance check for Wavelet Transform

Here, from the distribution curve and Q-Q plot, it is clear that denoised data closely follows the original data distribution.

5.2.5 Results after introducing Smoothing techniques:

The formulation of every smoothing technique is explained in detail in the above sections. The discussed encoder-decoder architecture is applied to smooth data for 10 min, 30 min, and 60 min.

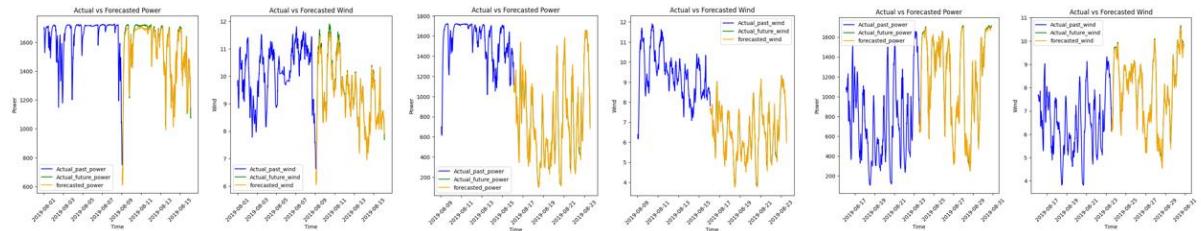
A] Simple exponential smoothing:

Before smoothing the signal, the SNR value of the wind speed signal was 9.28, which needs to be improved. As per [26], it should be at least 5 dB.

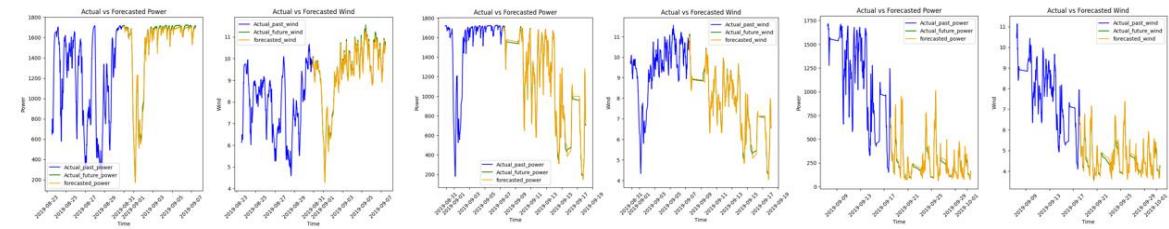
After simple exponential smoothing, the SNR value becomes 10.28 dB. Thus, the signal can be accepted.

- **10 min ahead forecasting results:**

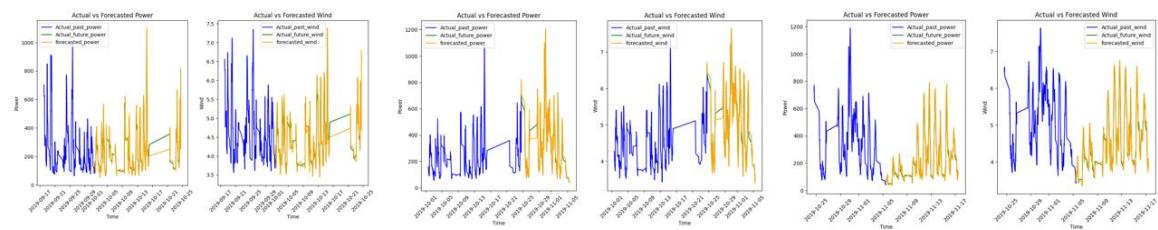
Aug:



Sept:



Oct-Nov:



Nov-Dec:

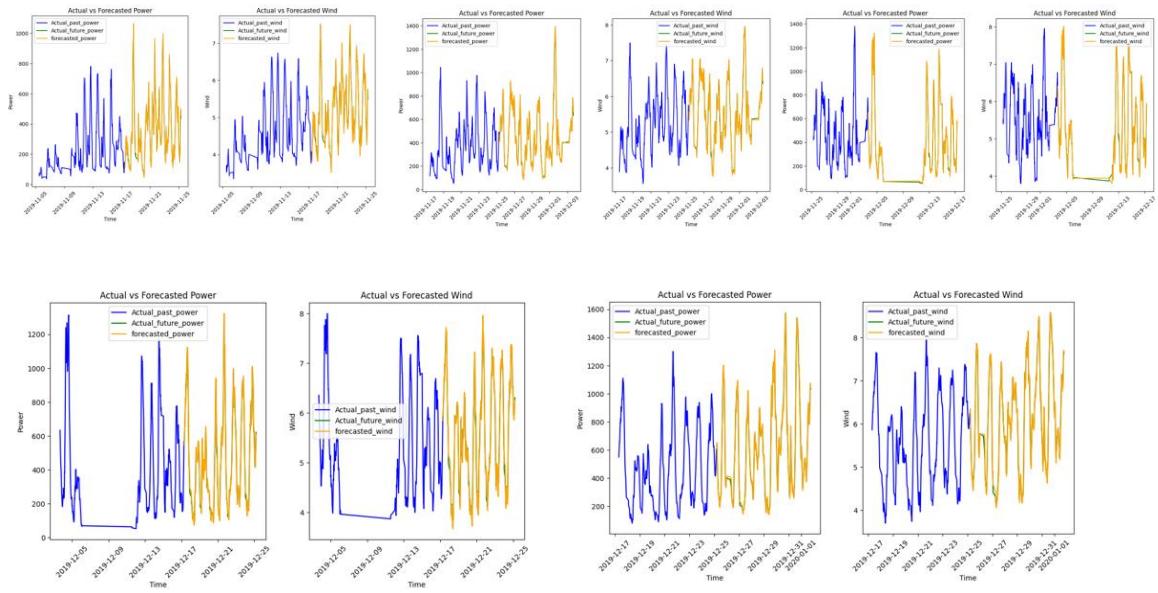


Fig 5.26: Weekly Prediction Plots using SE+ED Model for 10Min Forecasting Window

Model Performance:

SIMPLE EXPONENTIAL SMOOTHENING									
ENCODER DECODER ARCHITECTURE RESULTS for 10 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	32.82	1.41	0.66	Aug	Week 2	0.17	1.38	0.64
	Week 3	51.86	5.91	0.73		Week 3	0.17	2.02	0.71
	Week4	45.25	3.63	0.73		Week4	0.15	1.56	0.7
Sept	Week 1	28.37	1.85	0.71	Sept	Week 1	0.17	1.38	0.67
	Week 2	44.88	3.62	0.73		Week 2	0.15	1.54	0.69
	Week 3 & 4	32.33	8.26	0.69		Week 3 & 4	0.14	2.23	0.67
Oct	Week 1 & 3	35.53	8.6	0.68	Oct	Week 1 & 3	0.15	2.27	0.68
	Week 4	29.86	6.55	0.76		Week 4	0.12	1.81	0.71
Nov	Week 1 & 2	21.48	6.91	0.75	Nov	Week 1 & 2	0.1	1.7	0.74
	Week 2 & 3	26.69	5.77	0.75		Week 2 & 3	0.11	1.65	0.71
	Week 4	29.46	5.17	0.71		Week 4	0.11	1.57	0.7
Dec	Week 1 & 2	29.67	5.41	0.74	Dec	Week 1 & 2	0.11	1.53	0.71
	Week 3	27.69	5.31	0.77		Week 3	0.1	1.52	0.75

Fig 5.27: Model Performance for SE+ED Model-10 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:

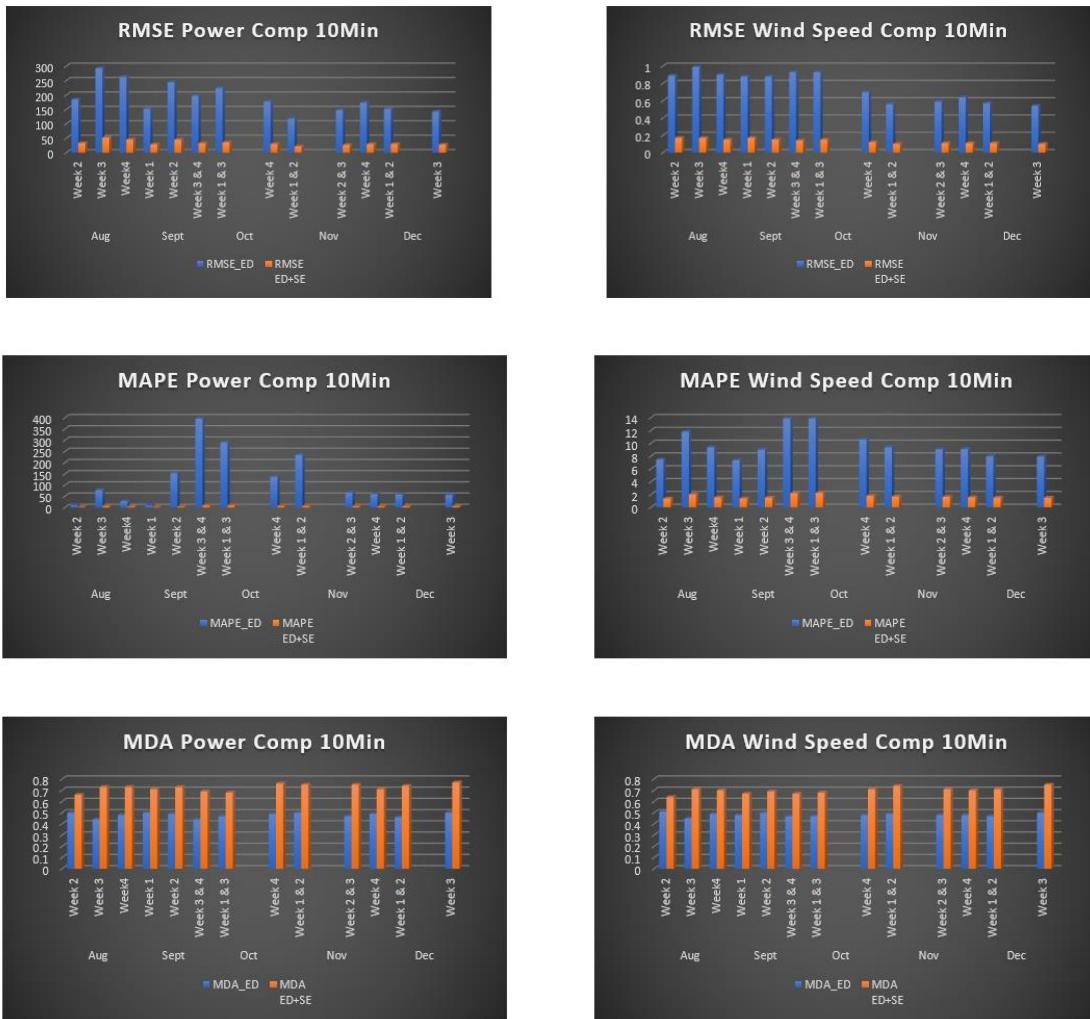
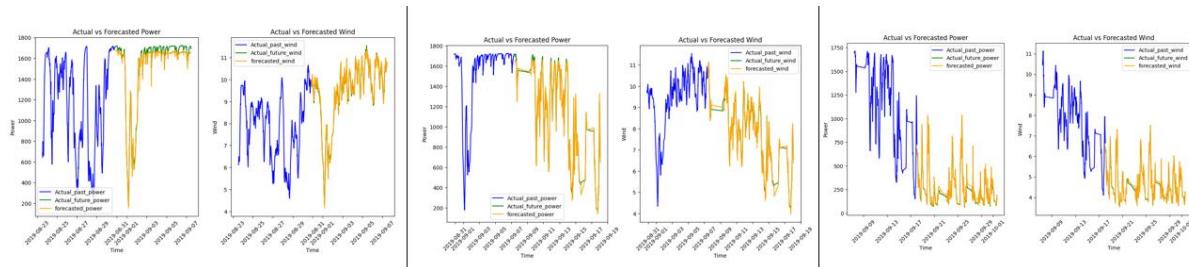


Fig 5.28: Performance comparison of ED and SE+ED architecture

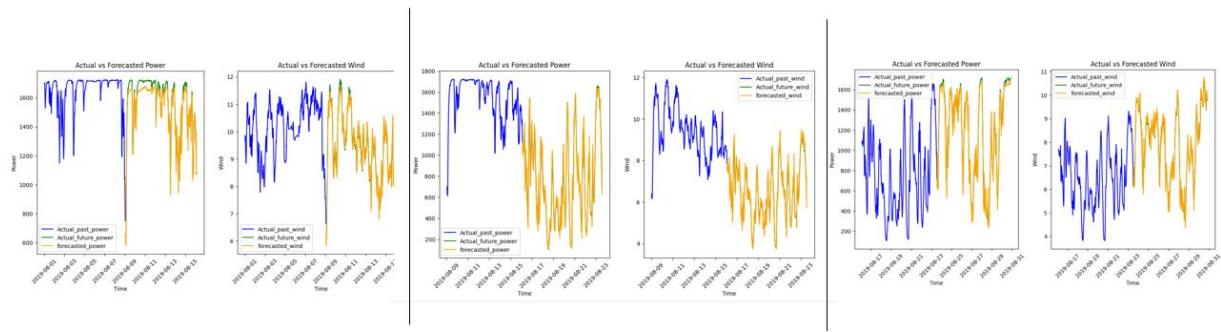
10Min Ahead Forecasting Window

- **30 min ahead forecasting results:**

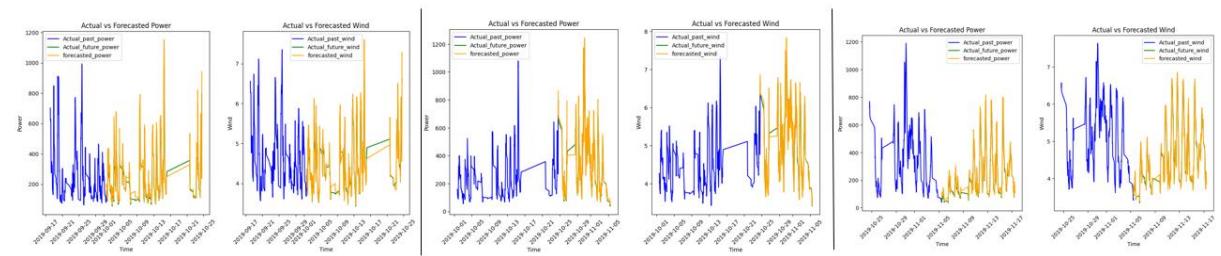
Aug:



Sept:



Oct-Nov:



Nov-Dec:

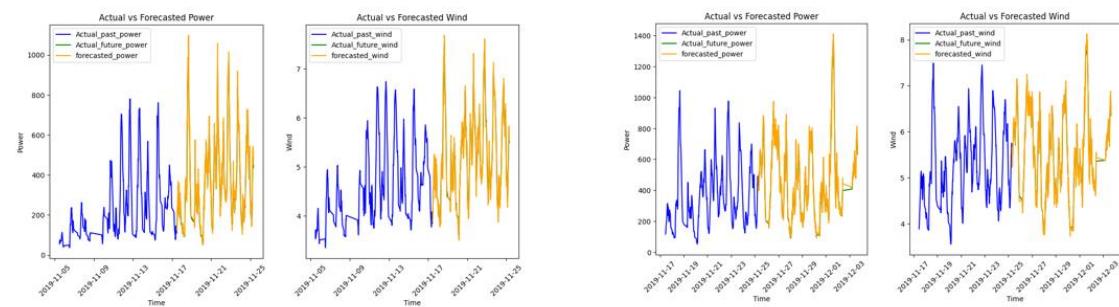
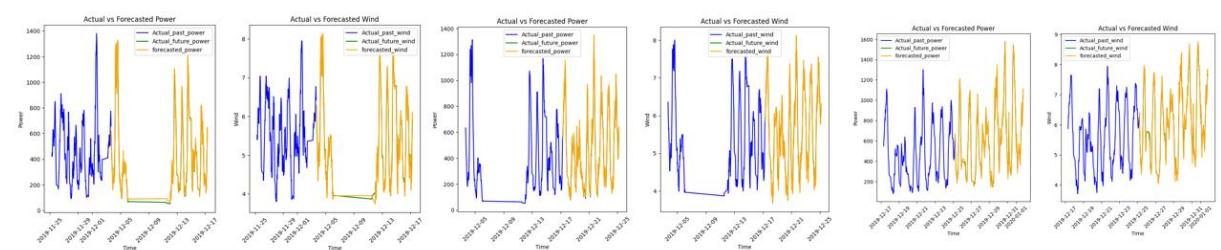


Fig 5.29: Weekly Prediction Plots using SE+ED Model for 30Min Forecasting Window

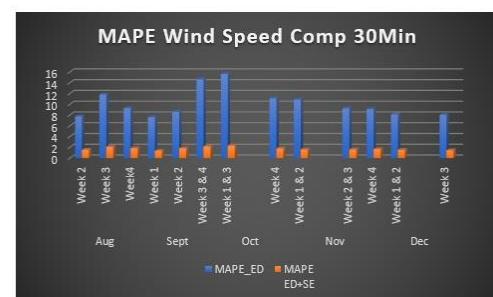
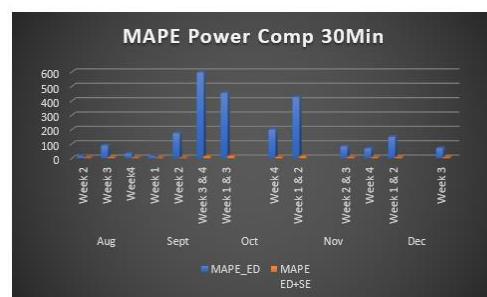
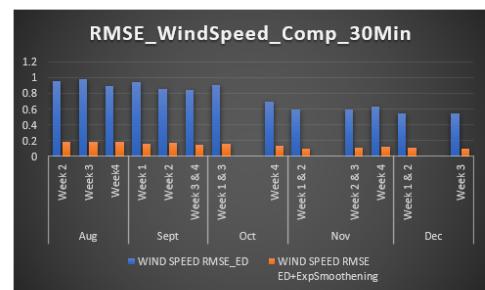
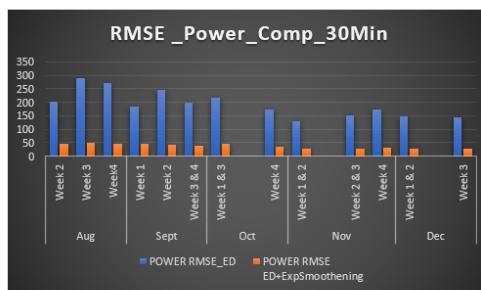
Model performance:

SIMPLE EXPONENTIAL SMOOTHENING				
ENCODER DECODER ARCHITECTURE RESULTS for 30 min				
WIND POWER				
Month	Week	RMSE	MAPE	MDA
Aug	Week 2	44.72	2.47	0.63
	Week 3	49.77	5.44	0.7
	Week4	46.64	3.6	0.68
Sept	Week 1	45.16	2.73	0.64
	Week 2	42.09	3.37	0.69
	Week 3 & 4	40.14	11.71	0.65
Oct	Week 1 & 3	45.72	13.31	0.66
	Week 4	34.46	8.38	0.7
Nov	Week 1 & 2	28.04	11.74	0.7
	Week 2 & 3	29.68	6.94	0.68
	Week 4	32.93	6.07	0.67
Dec	Week 1 & 2	28.82	6.53	0.7
	Week 3	27.44	5.64	0.71

WIND SPEED				
Month	Week	RMSE	MAPE	MDA
Aug	Week 2	0.18	1.55	0.62
	Week 3	0.18	2.09	0.67
	Week4	0.18	1.85	0.66
Sept	Week 1	0.16	1.37	0.64
	Week 2	0.17	1.81	0.64
	Week 3 & 4	0.14	2.17	0.64
Oct	Week 1 & 3	0.16	2.31	0.65
	Week 4	0.13	1.79	0.7
Nov	Week 1 & 2	0.1	1.61	0.7
	Week 2 & 3	0.11	1.62	0.66
	Week 4	0.12	1.67	0.65
Dec	Week 1 & 2	0.11	1.54	0.69
	Week 3	0.1	1.45	0.7

Fig 5.30: Model Performance for SE+ED Model-30 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



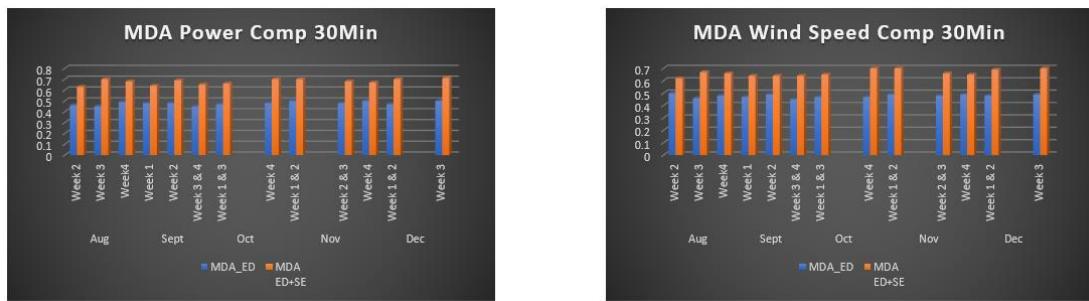
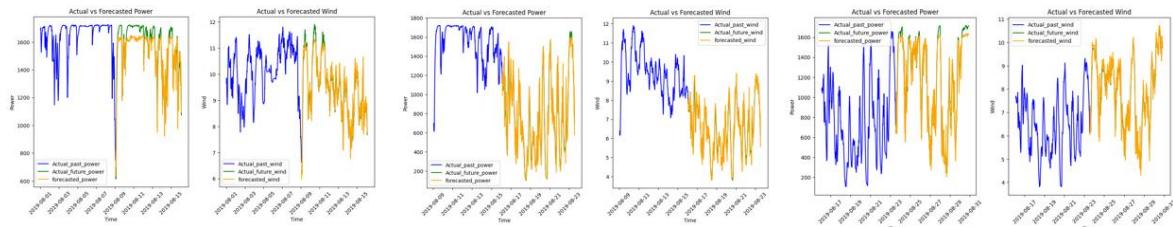


Fig 5.31: Performance comparison of ED and SE+ED architecture

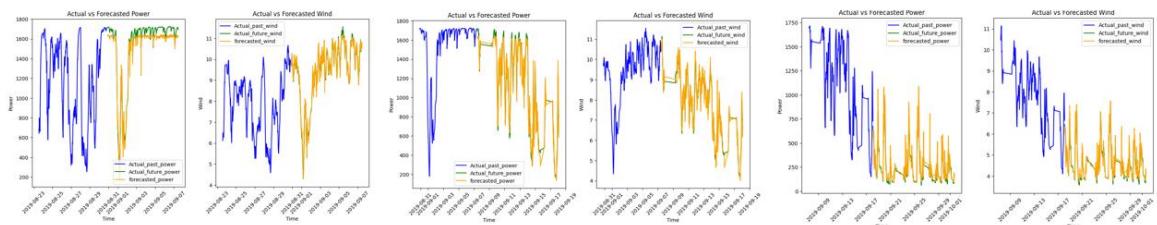
30Min Ahead Forecasting Window

- **60 min ahead forecasting:**

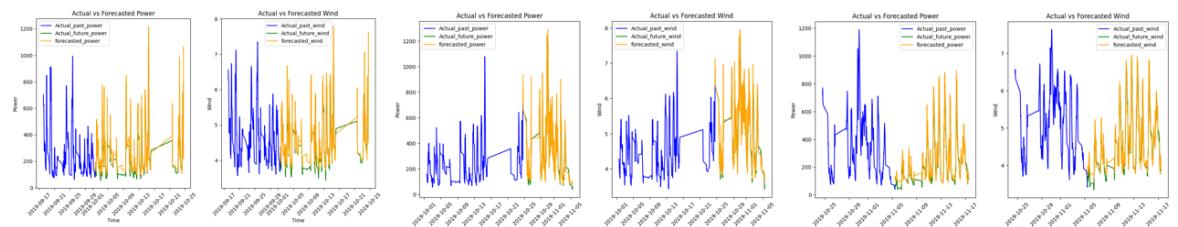
Aug:



Sept:



Oct-Nov:



Nov-Dec:

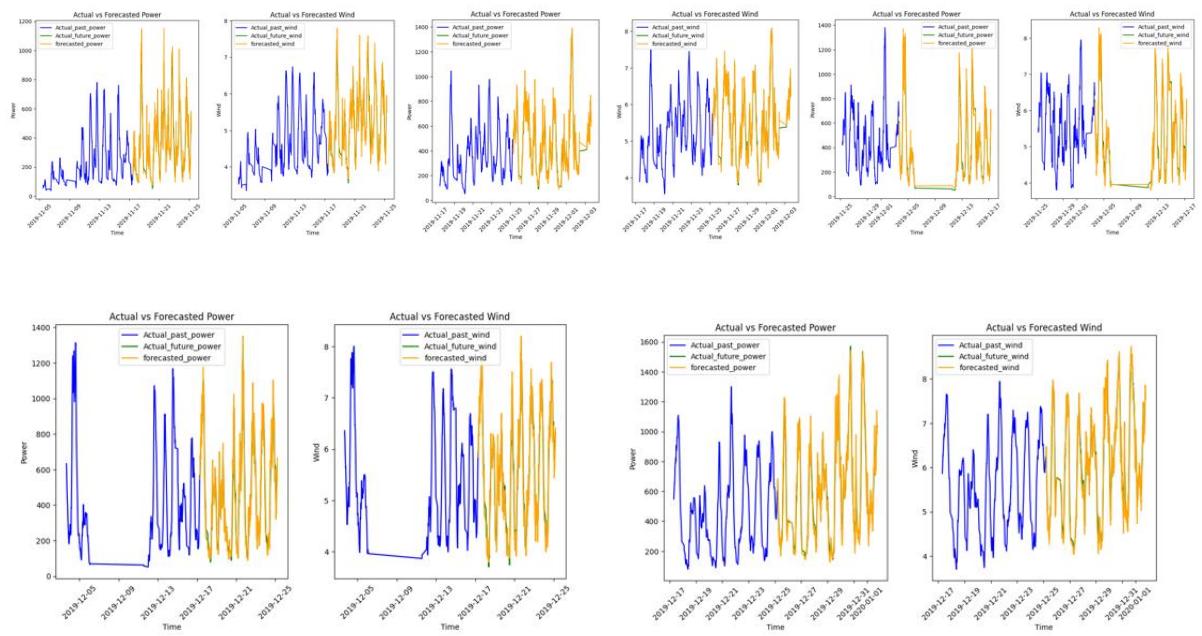


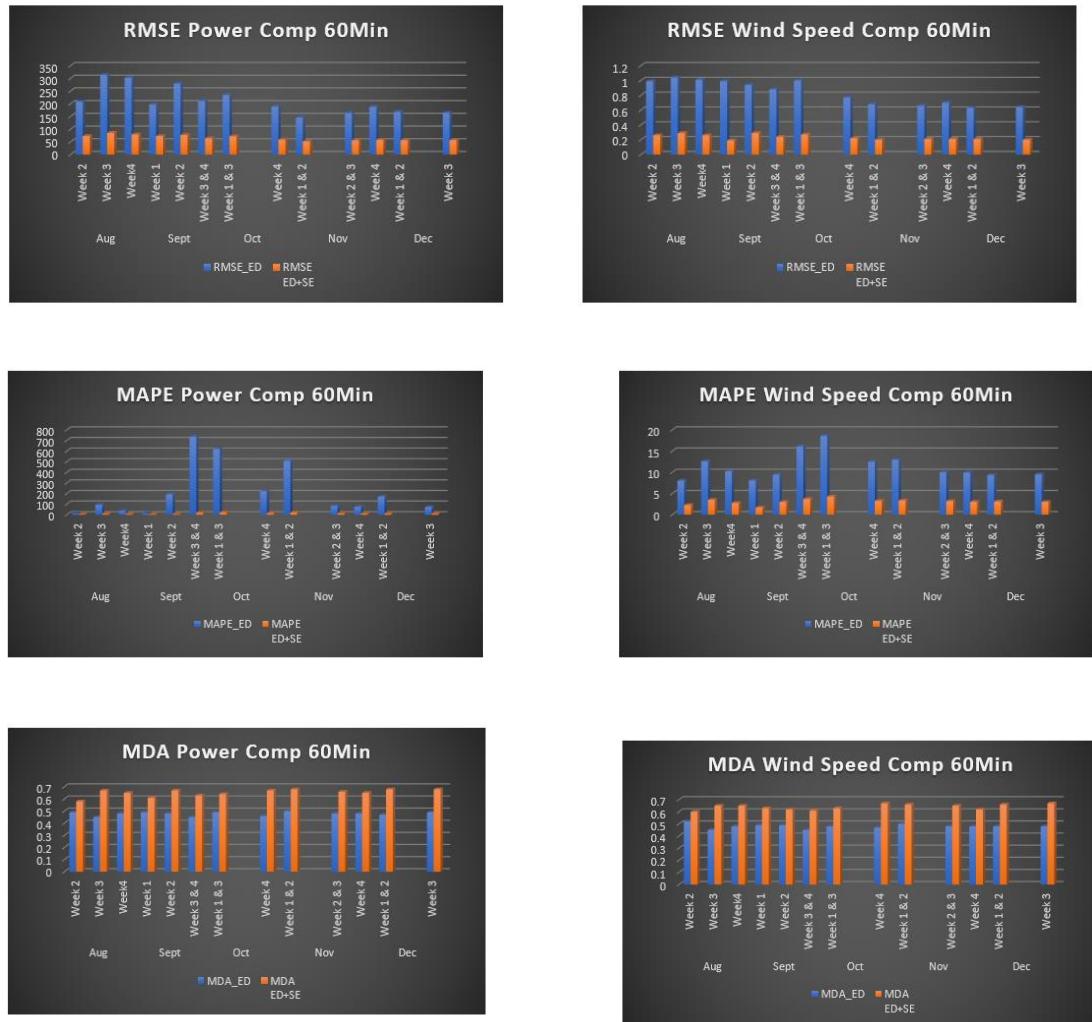
Fig 5.32: Weekly Prediction Plots using SE+ED Model for 60Min Forecasting Window

Model Performance:

SIMPLE EXPONENTIAL SMOOTHENING									
ENCODER DECODER ARCHITECTURE RESULTS for 60 min									
ACTIVW POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	72.75	4.14	0.58	Aug	Week 2	0.26	2.25	0.6
	Week 3	85.57	9.28	0.67		Week 3	0.29	3.49	0.65
	Week4	78.76	6.57	0.65		Week4	0.26	2.75	0.65
Sept	Week 1	72.71	4.79	0.61	Sept	Week 1	0.19	1.68	0.63
	Week 2	77.07	6.35	0.67		Week 2	0.29	3	0.62
	Week 3 & 4	63.88	18.05	0.63		Week 3 & 4	0.24	3.71	0.61
Oct	Week 1 & 3	71.83	22.01	0.64	Oct	Week 1 & 3	0.27	4.28	0.63
	Week 4	58.79	14.21	0.67		Week 4	0.22	3.27	0.67
Nov	Week 1 & 2	51.82	20.62	0.68	Nov	Week 1 & 2	0.2	3.35	0.66
	Week 2 & 3	55.63	12.26	0.66		Week 2 & 3	0.21	3.19	0.65
	Week 4	57.68	10.67	0.65		Week 4	0.21	3.01	0.62
Dec	Week 1 & 2	56.34	11.34	0.68	Dec	Week 1 & 2	0.21	3.09	0.66
	Week 3	56.61	10.97	0.68		Week 3	0.2	2.99	0.67

Fig 5.33: Model Performance for SE+ED Model-60 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



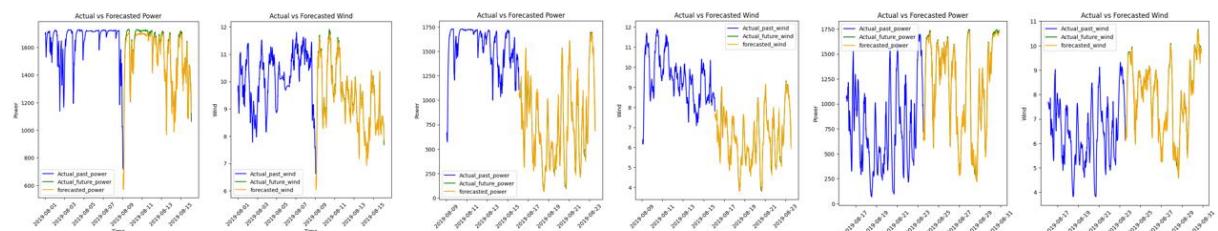
**Fig 5.34: Performance comparison of ED and SE+ED architecture
60Min Ahead Forecasting Window**

B] Double Exponential Smoothing:

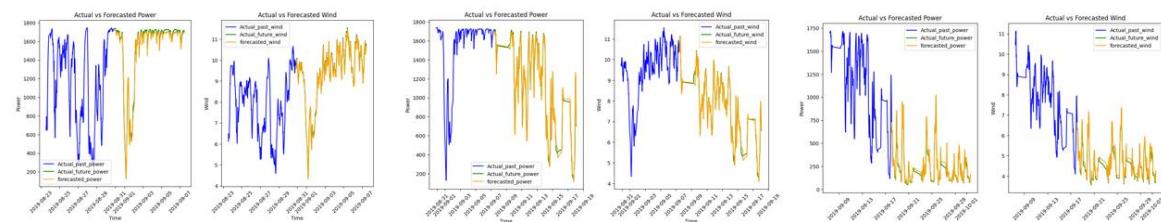
After smoothing the data, the SNR value for the wind speed feature signal becomes 10.21 dB, which can be accepted.

- **10 min ahead forecasting results:**

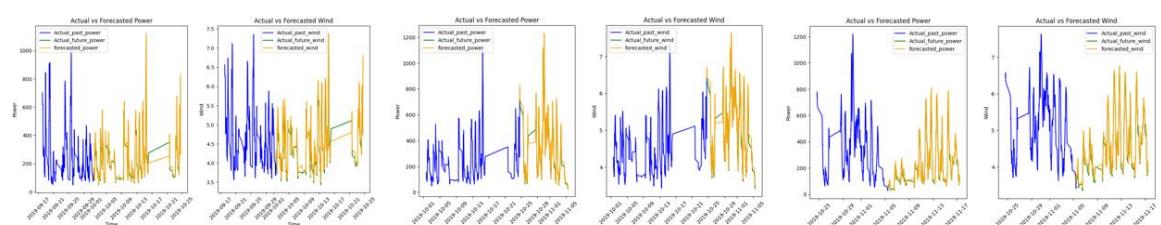
Aug:



Sept:



Oct-Nov:



Nov-Dec:

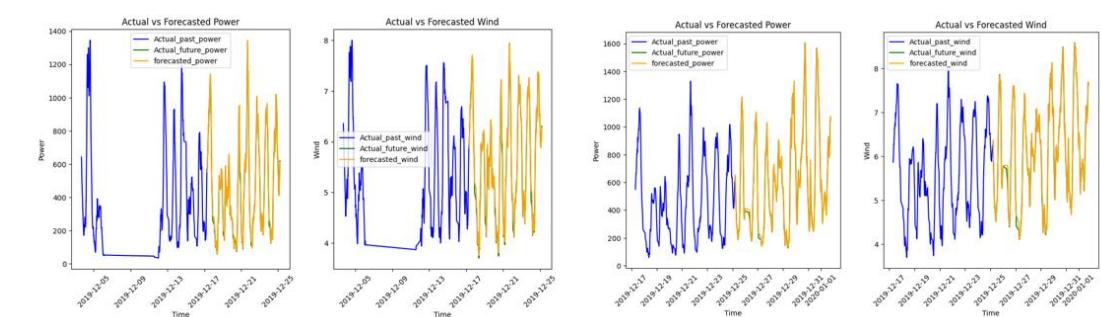
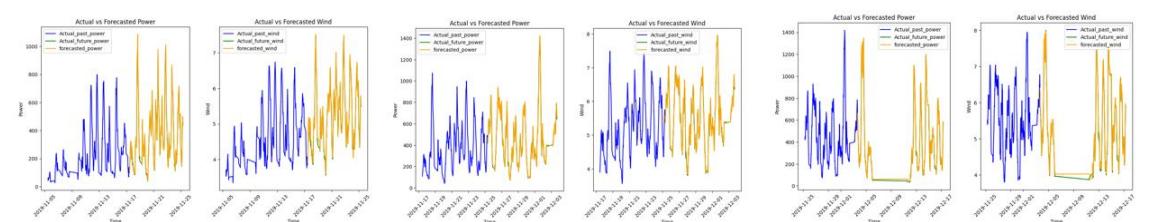


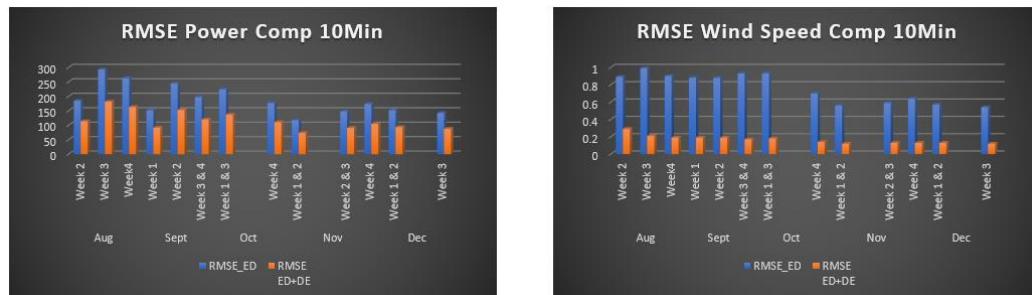
Fig 5.35: Weekly Prediction Plots using DE+ED Model for 10Min Forecasting Window

Model Performance:

DOUBLE EXPONENTIAL SMOOTHENING									
ENCODER DECODER ARCHITECTURE RESULTS for 10 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	113.5	5.35	0.58	Aug	Week 2	0.29	1.57	0.69
	Week 3	181.53	37.77	0.58		Week 3	0.21	2.59	0.76
	Week4	162.85	18.86	0.6		Week4	0.19	2	0.75
Sept	Week 1	91.59	5.39	0.57	Sept	Week 1	0.19	1.68	0.72
	Week 2	153.06	26.61	0.59		Week 2	0.19	1.99	0.77
	Week 3 & 4	119.7	334.57	0.57		Week 3 & 4	0.17	2.83	0.75
Oct	Week 1 & 3	136.76	346.66	0.57	Oct	Week 1 & 3	0.18	2.8	0.75
	Week 4	110.62	84.78	0.61		Week 4	0.14	2.23	0.78
Nov	Week 1 & 2	73.31	108.65	0.63	Nov	Week 1 & 2	0.12	2.15	0.8
	Week 2 & 3	90.75	104.33	0.59		Week 2 & 3	0.13	1.58	0.79
	Week 4	104.68	33.63	0.61		Week 4	0.13	1.9	0.75
Dec	Week 1 & 2	93.48	44.35	0.59	Dec	Week 1 & 2	0.13	1.83	0.8
	Week 3	87.42	52.43	0.64		Week 3	0.12	1.86	0.76

Fig 5.36: Model Performance for DE+ED Model-10 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



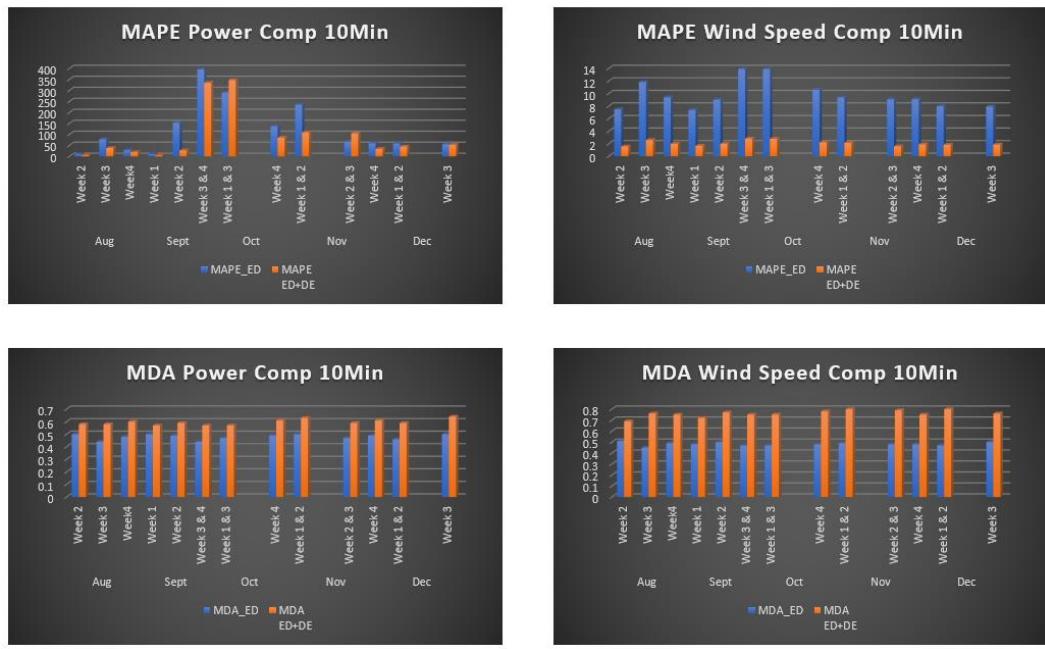
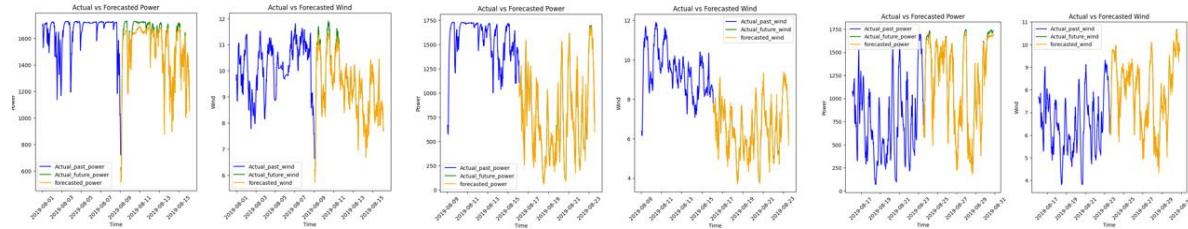


Fig 5.37: Performance comparison of ED and DE+ED architecture

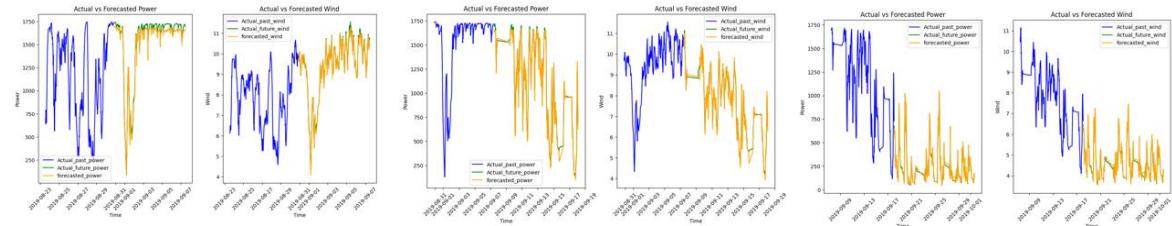
10Min Ahead Forecasting Window

- **30 min ahead forecasting results:**

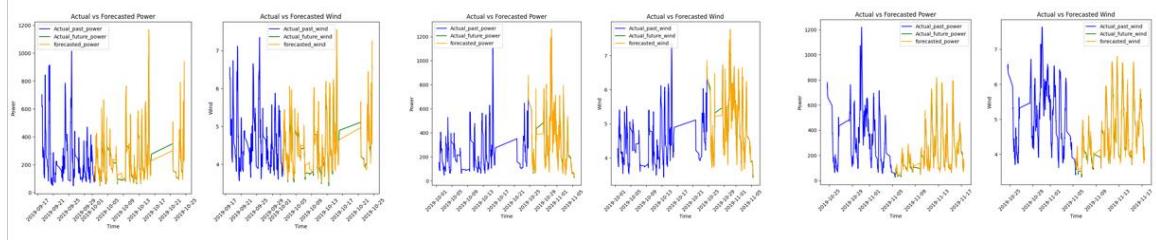
Aug:



Sept:



Oct-Nov:



Nov-Dec:

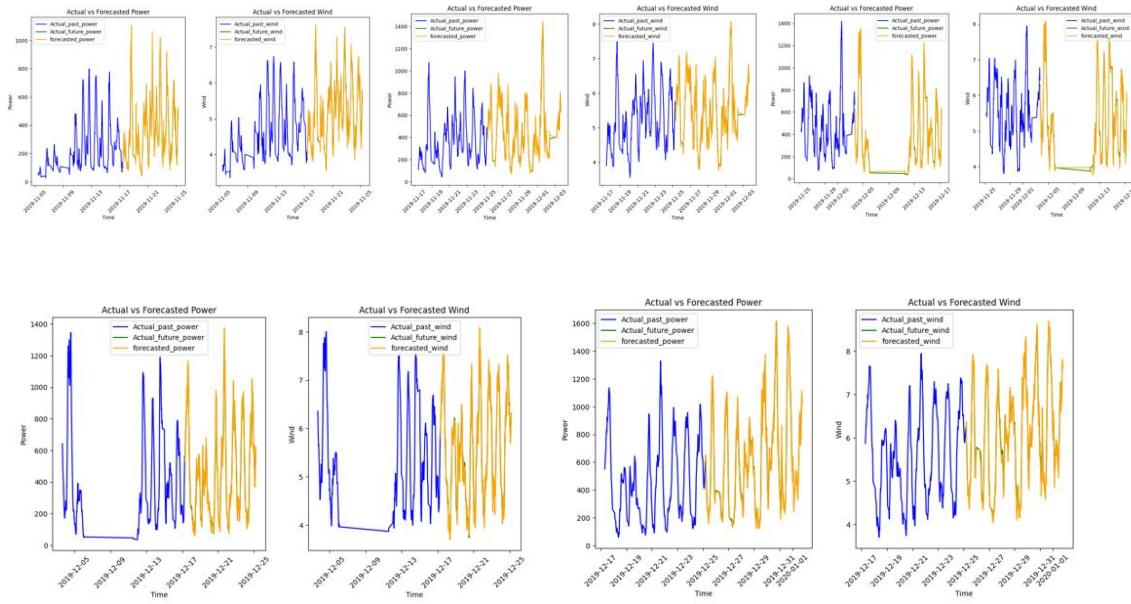


Fig 5.38: Weekly Prediction Plots using DE+ED Model for 30Min Forecasting Window

Model Performance:

DOUBLE EXPONENTIAL SMOOTHENING									
ENCODER DECODER ARCHITECTURE RESULTS for 30 min									
WIND POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week			
Aug	Week 2	109.61	5.3	0.57	Aug	Week 2	0.17	1.39	0.68
	Week 3	176.71	31.13	0.57		Week 3	0.18	2.12	0.74
	Week4	160.48	19.45	0.6		Week4	0.16	1.6	0.74
Sept	Week 1	88.8	5.12	0.57	Sept	Week 1	0.16	1.36	0.7
	Week 2	147.86	25.5	0.59		Week 2	0.16	1.64	0.74
	Week 3 & 4	120.23	125.73	0.57		Week 3 & 4	0.15	2.37	0.72
Oct	Week 1 & 3	134.91	79.55	0.57	Oct	Week 1 & 3	0.16	2.51	0.72
	Week 4	107.65	52.37	0.61		Week 4	0.12	1.83	0.76
Nov	Week 1 & 2	71.73	55.03	0.63	Nov	Week 1 & 2	0.1	1.71	0.78
	Week 2 & 3	88.19	52.97	0.58		Week 2 & 3	0.1	1.6	0.76
	Week 4	104.88	28.13	0.61		Week 4	0.11	1.6	0.72
Dec	Week 1 & 2	90.53	26.49	0.58	Dec	Week 1 & 2	0.1	1.42	0.77
	Week 3	84.89	52.45	0.63		Week 3	0.09	1.41	0.74

Fig 5.39: Model Performance for DE+ED Model-30 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



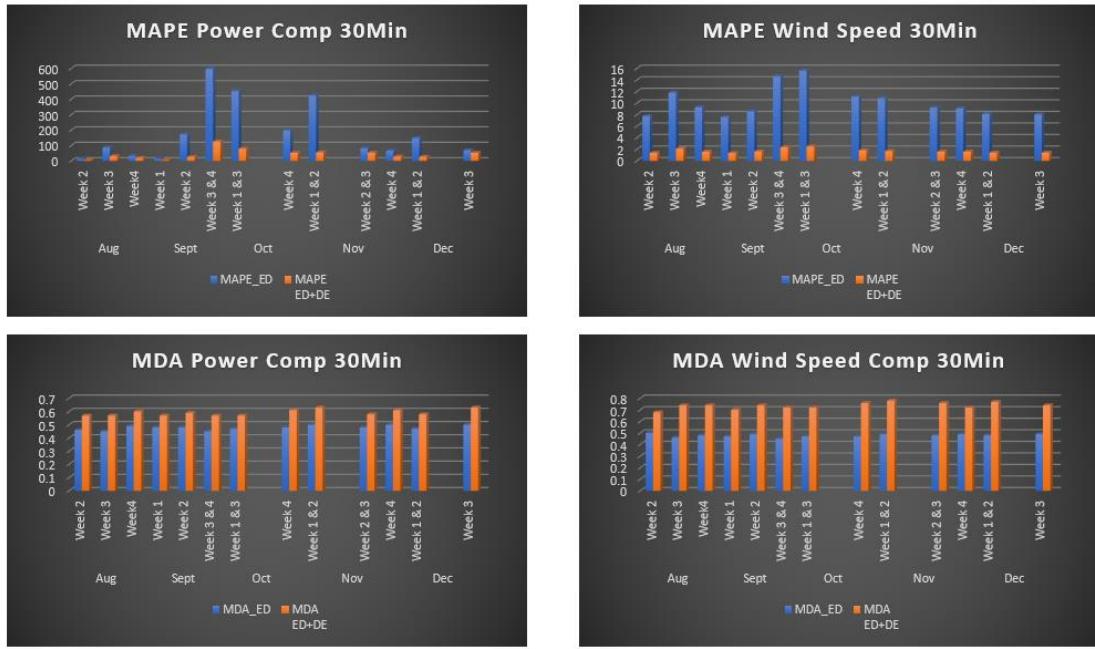
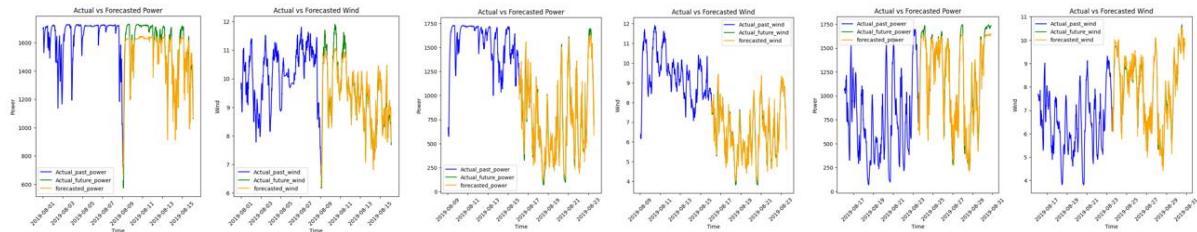


Fig 5.40: Performance comparison of ED and DE+ED architecture

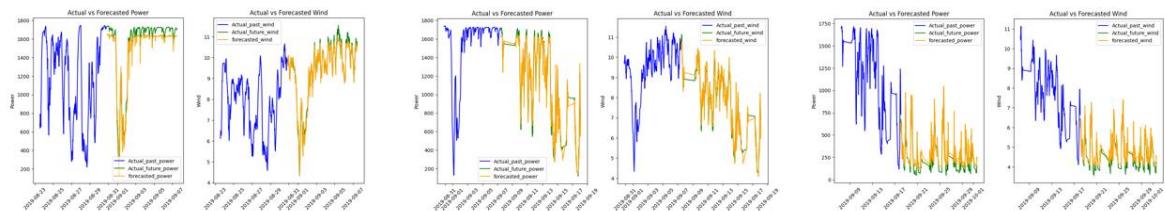
30Min Ahead Forecasting Window

- **60 Min ahead forecasting results:**

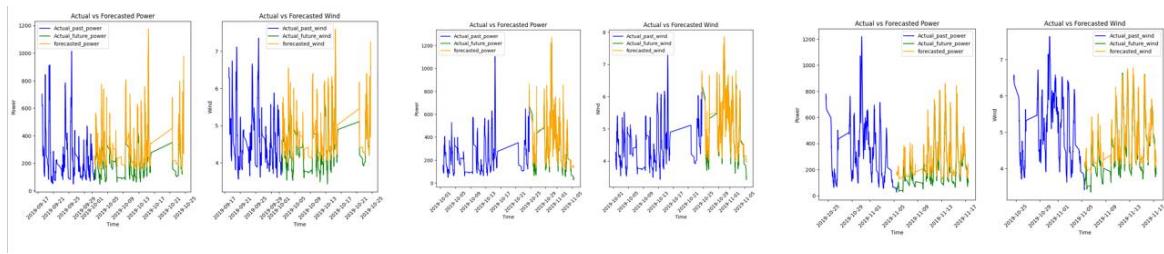
Aug:



Sept:



Oct-Nov:



Nov-Dec:

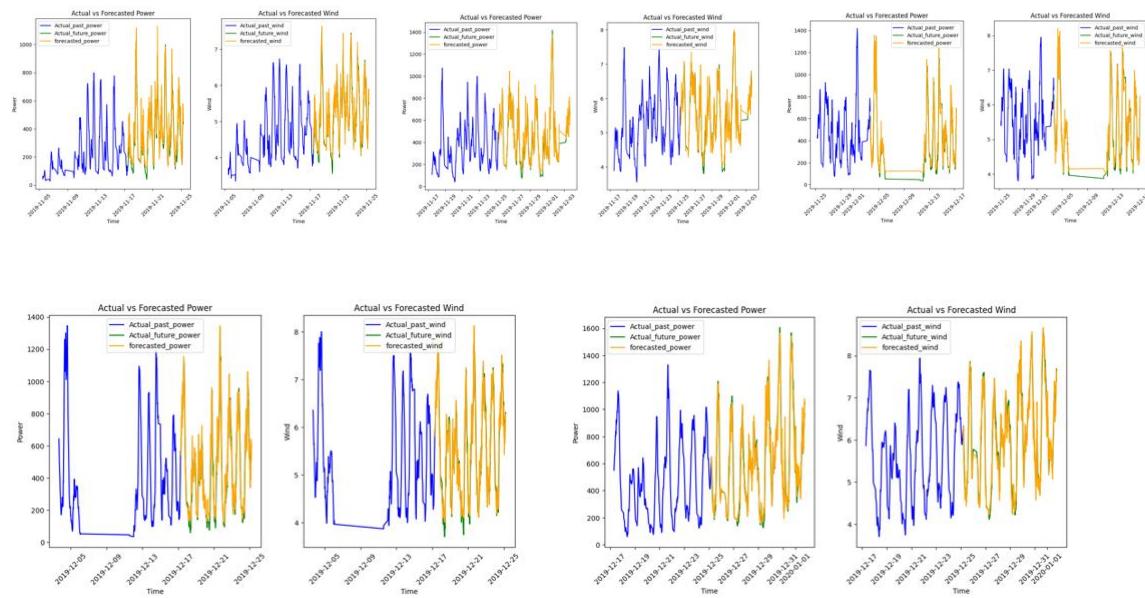


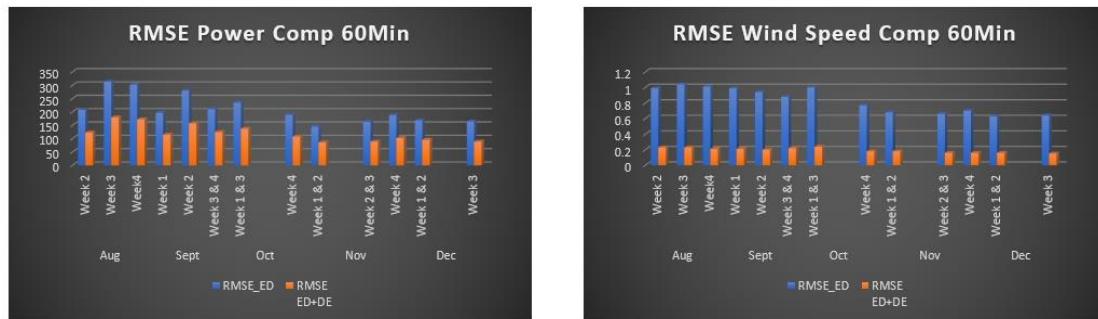
Fig 5.41: Weekly Prediction Plots using DE+ED Model for 60Min Forecasting Window

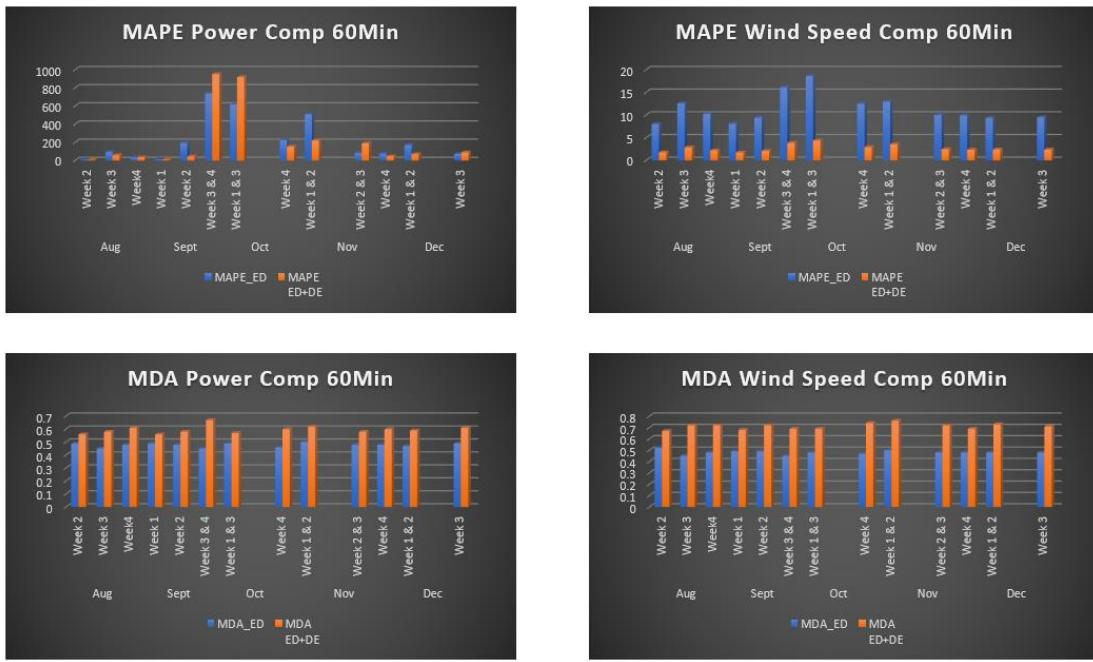
Model Performance:

DOUBLE EXPONENTIAL SMOOTHENING									
ENCODER DECODER ARCHITECTURE RESULTS for 60 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week			
Aug	Week 2	122.78	7.29	0.56	Aug	Week 2	0.23	1.78	0.67
	Week 3	179.74	58.09	0.58		Week 3	0.23	2.87	0.72
	Week4	172.35	34.01	0.61		Week4	0.21	2.21	0.72
Sept	Week 1	114.74	11.31	0.56	Sept	Week 1	0.21	1.74	0.68
	Week 2	157.18	43.18	0.58		Week 2	0.2	2.09	0.72
	Week 3 & 4	124.45	949.46	0.67		Week 3 & 4	0.22	3.85	0.69
Oct	Week 1 & 3	137.35	916.5	0.57	Oct	Week 1 & 3	0.24	4.39	0.69
	Week 4	107.82	152.71	0.6		Week 4	0.18	2.92	0.74
Nov	Week 1 & 2	85.76	219.49	0.62	Nov	Week 1 & 2	0.18	3.56	0.76
	Week 2 & 3	88.48	188.25	0.58		Week 2 & 3	0.16	2.45	0.72
	Week 4	102.51	45.79	0.6		Week 4	0.16	2.35	0.69
Dec	Week 1 & 2	95.94	69.16	0.59	Dec	Week 1 & 2	0.16	2.42	0.73
	Week 3	88.77	87.75	0.61		Week 3	0.15	2.37	0.71

Fig 5.42: Model Performance for DE+ED Model-60 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:





**Fig 5.43: Performance comparison of ED and DE+ED architecture
60Min Ahead Forecasting Window**

C] Fast Fourier Transform:

It is the technique used to transform time series data from the time domain to the frequency domain. It basically clears low amplitude signal (high frequency), which is known to be noise, and smoothens the data.

After implementing FFT on original data, the SNR of both the features- Active power and wind speed are crossing the threshold values of 5dB.

SNR of Active Power = 6.705 dB

SNR of Wind Speed = 9.515 dB (a value greater than 10 dB could bring more information from the signal).

To smoothen the data, we pick up threshold frequency, which helps in removing high-frequency signals. Threshold frequency depends on the energy threshold value, which can be manipulated.

The energy threshold value is kept at 95%, which helps to retain the frequencies that together account for 95% of the total energies of the signal. Frequencies beyond this threshold are considered as noise.

Below are the cumulative energy plots for both the features:

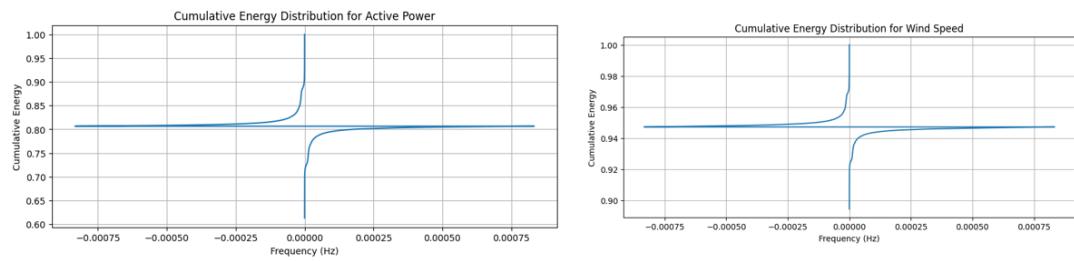
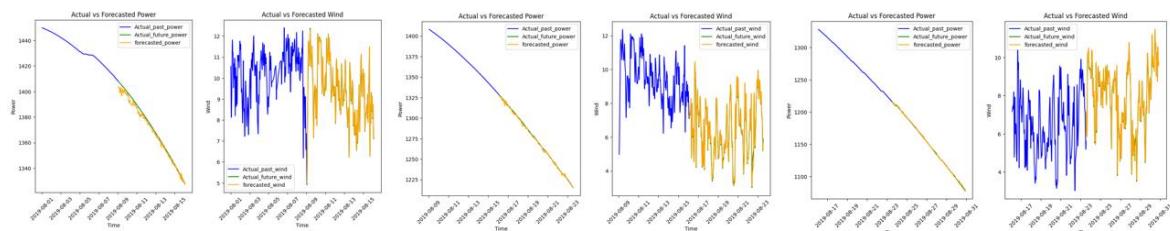


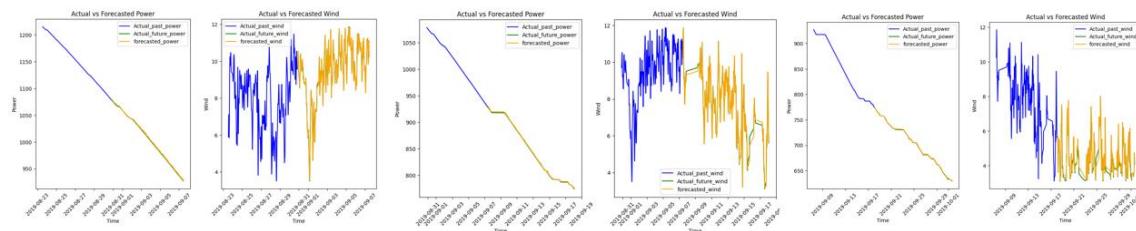
Fig 5.44 Cumulative Energy Plots

- **10 Min ahead forecasting Results:**

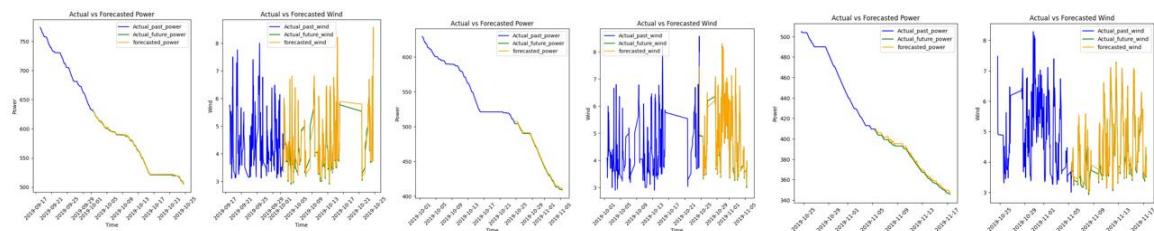
Aug:



Sept:



Oct-Nov:



Nov-Dec:

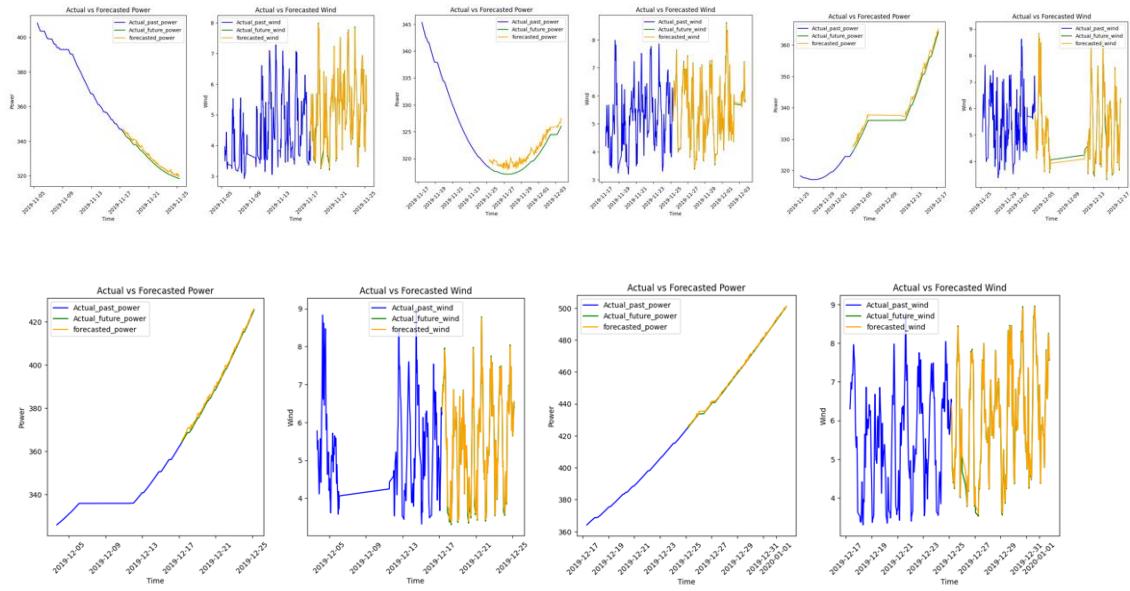


Fig 5.45: Weekly Prediction Plots using FFT+ED Model for 10Min Forecasting Window

Model Performance:

FFT				
ENCODER DECODER ARCHITECTURE RESULTS for 10 min				
ACTIVE POWER				
Month	Week	RMSE	MAPE	MDA
Aug	Week 2	2.68	0.15	0.65
	Week 3	1.18	0.07	0.75
	Week4	0.87	0.06	0.83
Sept	Week 1	1.46	0.13	0.77
	Week 2	1.15	0.12	0.8
	Week 3 & 4	0.73	0.09	0.89
Oct	Week 1 & 3	1.11	0.18	0.83
	Week 4	1.44	0.3	0.81
Nov	Week 1 & 2	1.71	0.44	0.7
	Week 2 & 3	1.69	0.5	0.66
	Week 4	1.57	0.48	0.53
Dec	Week 1 & 2	1.35	0.38	0.7
	Week 3	0.97	0.23	0.76

WIND SPEED				
Month	Week	RMSE	MAPE	MDA
Aug	Week 2	0.2	1.72	0.89
	Week 3	0.25	2.97	0.9
	Week4	0.22	2.31	0.89
Sept	Week 1	0.18	1.61	0.89
	Week 2	0.21	2.2	0.89
	Week 3 & 4	0.18	3.14	0.9
Oct	Week 1 & 3	0.2	3.32	0.89
	Week 4	0.16	2.57	0.9
Nov	Week 1 & 2	0.14	2.49	0.9
	Week 2 & 3	0.14	2.19	0.9
	Week 4	0.14	2.05	0.9
Dec	Week 1 & 2	0.13	1.91	0.92
	Week 3	0.13	2.02	0.91

Fig 5.46: Model Performance for FFT+ED Model-10 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:

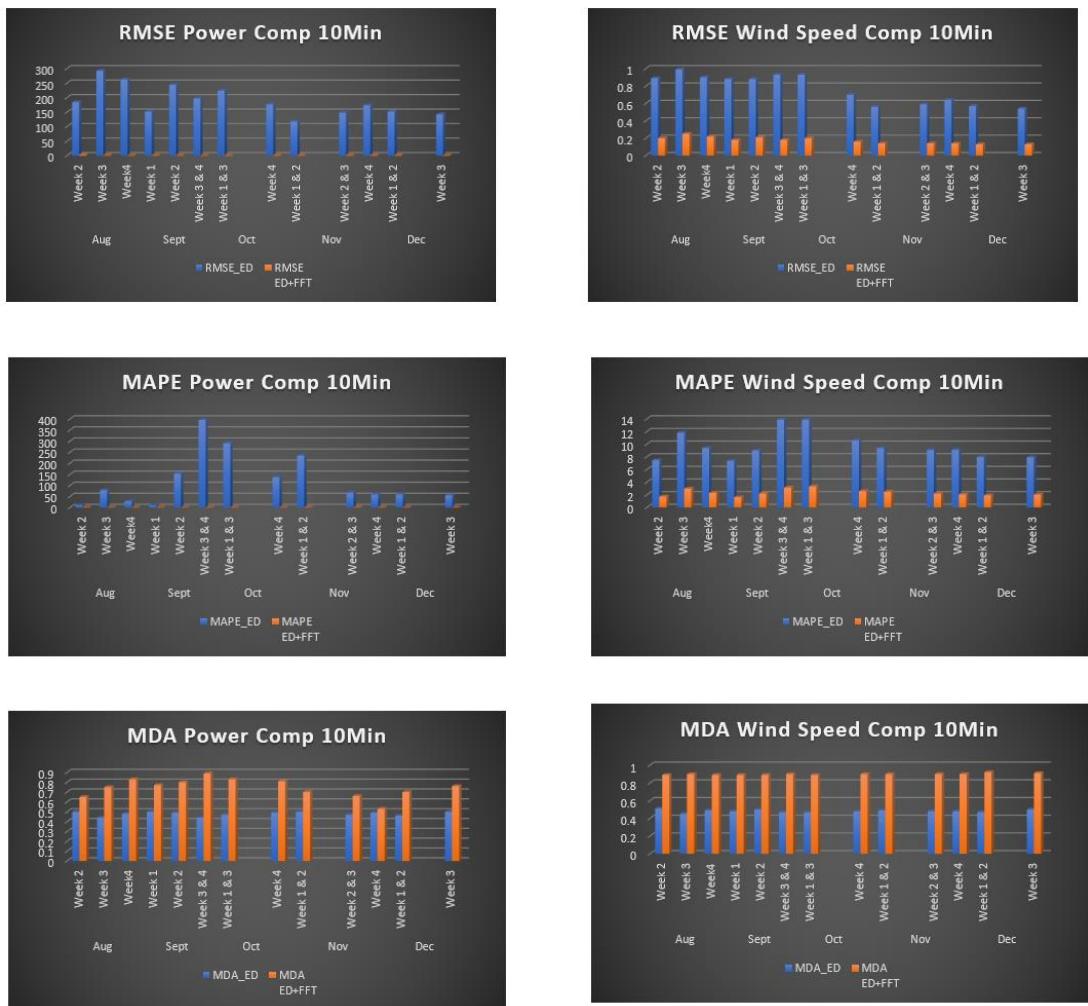
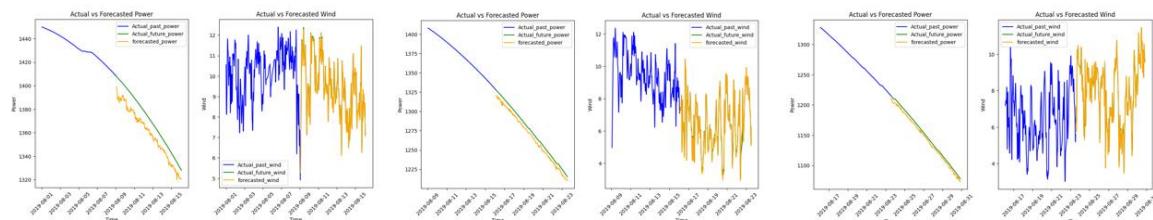


Fig 5.47: Performance comparison of ED and FFT+ED architecture

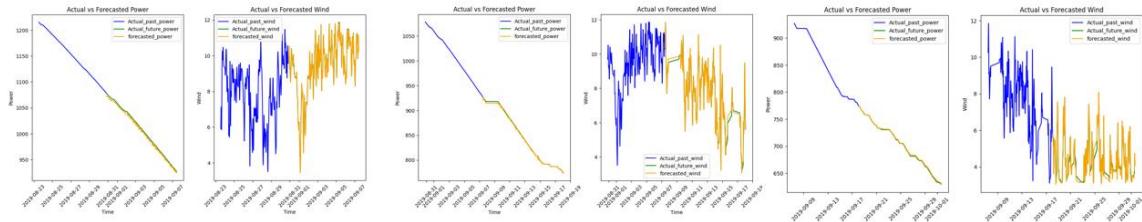
10Min Ahead Forecasting Window

- **30 min ahead forecasting results:**

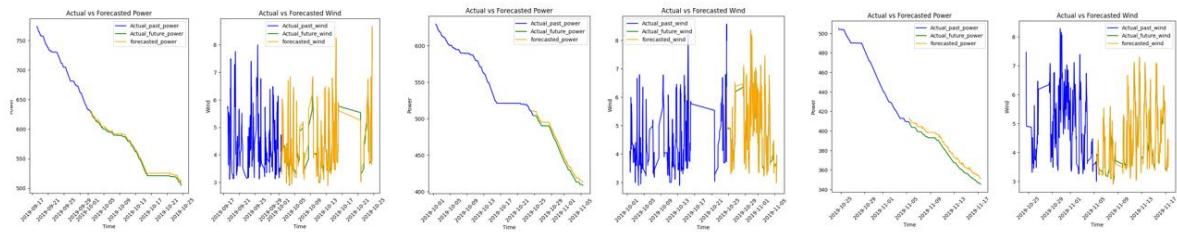
Aug:



Sept:



Oct-Nov:



Nov-Dec:

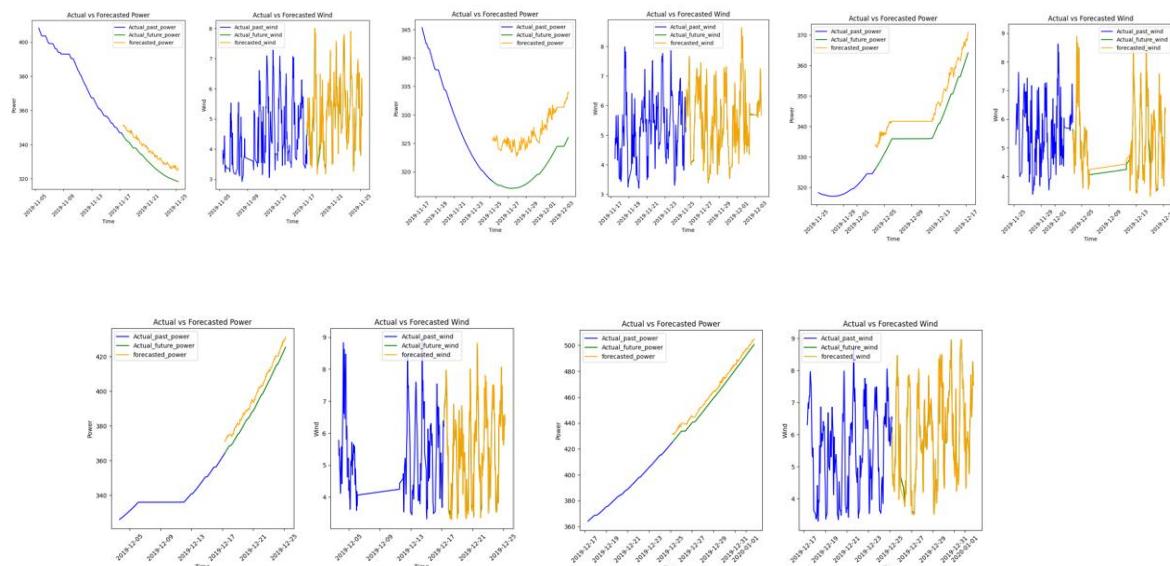


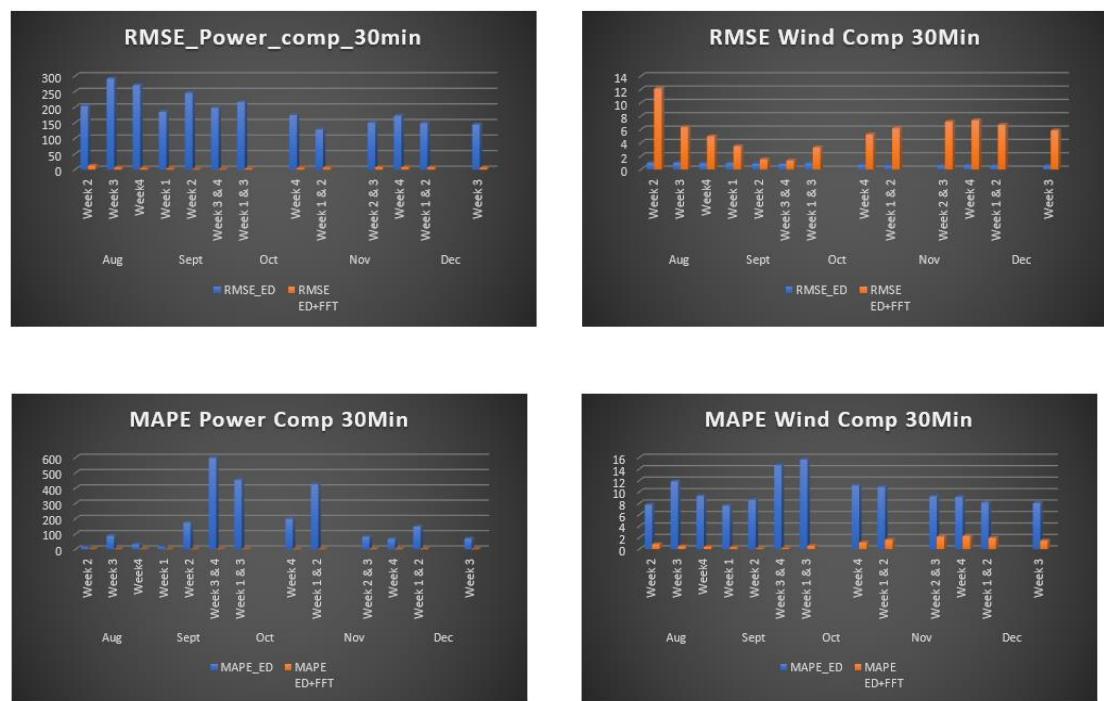
Fig 5.48: Weekly Prediction Plots using FFT+ED Model for 30Min Forecasting Window

Model Metric Performance:

FFT					
ENCODER DECODER ARCHITECTURE RESULTS for 30 min					
ACTIVE POWER					WIND SPEED
Month	Week	RMSE	MAPE	MDA	
Aug	Week 2	12.11	0.86	0.61	
	Week 3	6.35	0.49	0.73	
	Week4	4.96	0.42	0.75	
Sept	Week 1	3.48	0.32	0.7	
	Week 2	1.51	0.14	0.71	
	Week 3 & 4	1.33	0.17	0.8	
Oct	Week 1 & 3	3.3	0.57	0.82	
	Week 4	5.25	1.15	0.78	
Nov	Week 1 & 2	6.16	1.64	0.71	
	Week 2 & 3	7.16	2.17	0.66	
	Week 4	7.37	2.3	0.52	
Dec	Week 1 & 2	6.68	1.94	0.62	
	Week 3	5.84	1.48	0.69	

Fig 5.49: Model Performance for FFT+ED Model-30 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



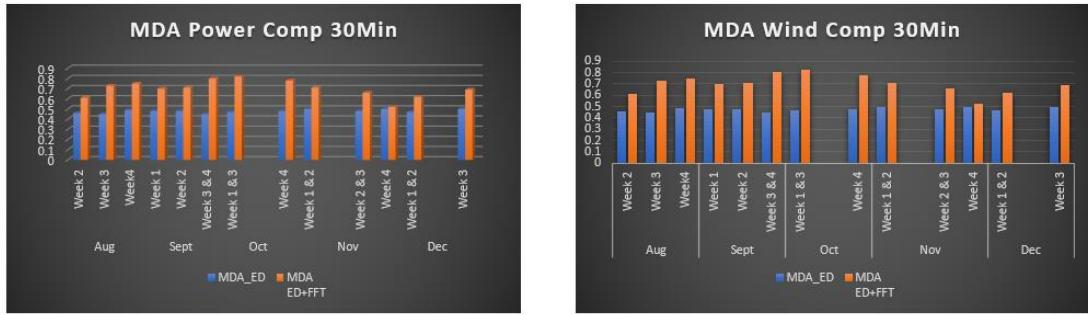
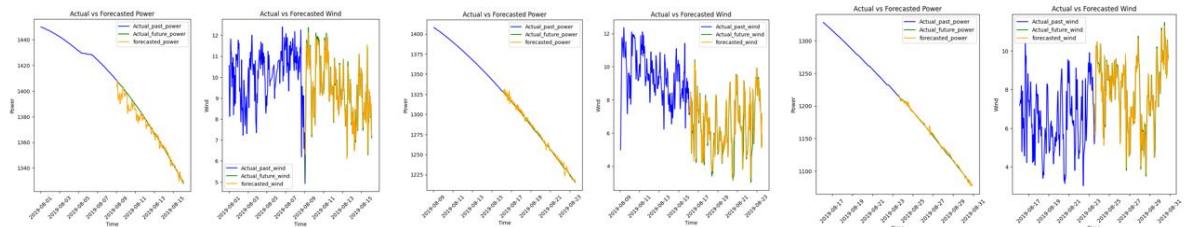


Fig 5.50: Performance comparison of ED and FFT+ED architecture

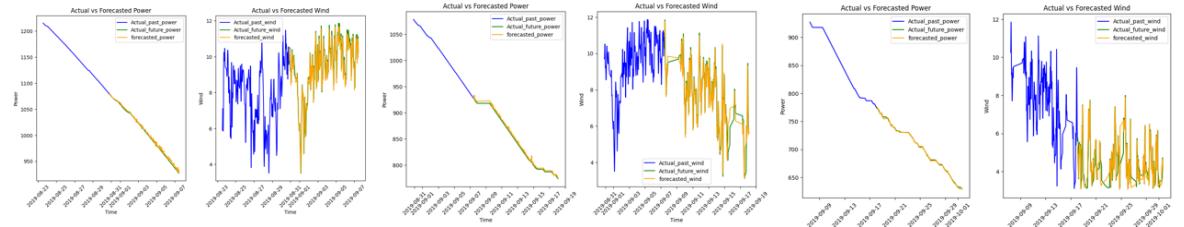
30Min Ahead Forecasting Window

- **60 Min ahead forecasting results:**

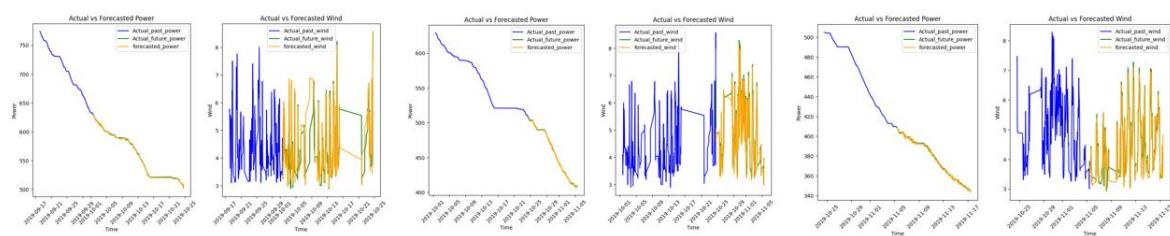
Aug:



Sept:



Oct-Nov:



Nov-Dec:

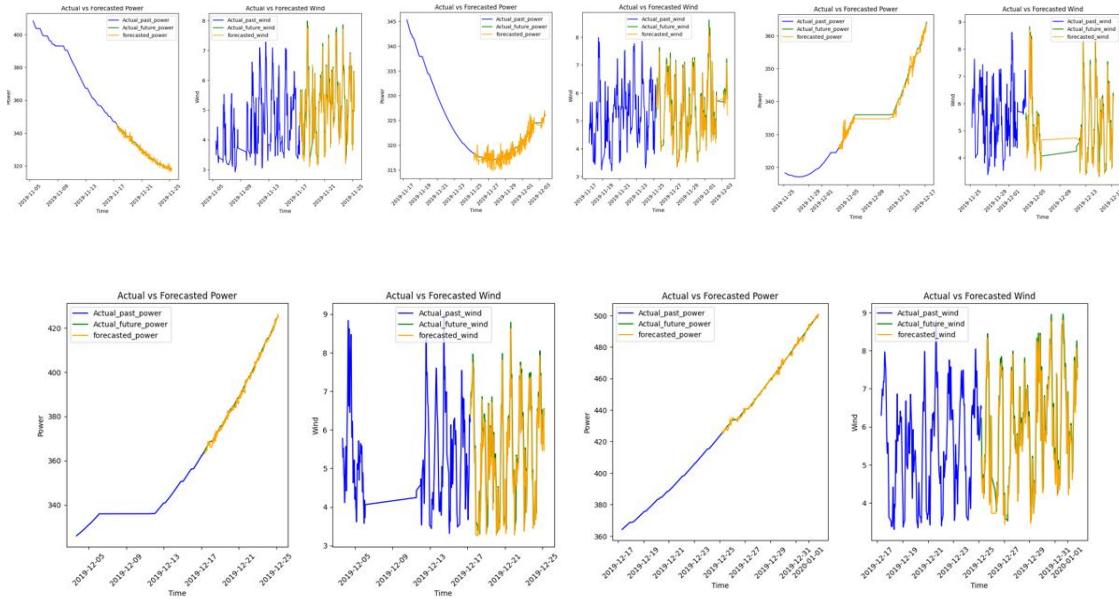


Fig 5.51: Weekly Prediction Plots using FFT+ED Model for 60Min Forecasting Window

Model Performance:

FFT									
ENCODER DECODER ARCHITECTURE RESULTS for 60 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	4.69	0.23	0.55	Aug	Week 2	0.63	5.35	0.62
	Week 3	2.34	0.15	0.59		Week 3	0.83	9.9	0.61
	Week4	1.81	0.12	0.67		Week4	0.75	7.9	0.58
Sept	Week 1	3.11	0.26	0.67	Sept	Week 1	0.61	5.52	0.6
	Week 2	3.58	0.38	0.62		Week 2	0.77	7.84	0.58
	Week 3 & 4	1.06	0.12	0.75		Week 3 & 4	0.62	10.57	0.62
Oct	Week 1 & 3	1.15	0.17	0.7	Oct	Week 1 & 3	0.73	12.01	0.57
	Week 4	0.94	0.16	0.61		Week 4	0.61	9.46	0.62
Nov	Week 1 & 2	1.26	0.28	0.56	Nov	Week 1 & 2	0.52	8.95	0.63
	Week 2 & 3	1.13	0.28	0.5		Week 2 & 3	0.54	8.26	0.6
	Week 4	1.07	0.27	0.52		Week 4	0.53	7.62	0.64
Dec	Week 1 & 2	1.12	0.25	0.56	Dec	Week 1 & 2	0.51	6.95	0.67
	Week 3	0.98	0.2	0.63		Week 3	0.51	7.49	0.64

Fig 5.48: Model Performance for FFT+ED Model-60 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:

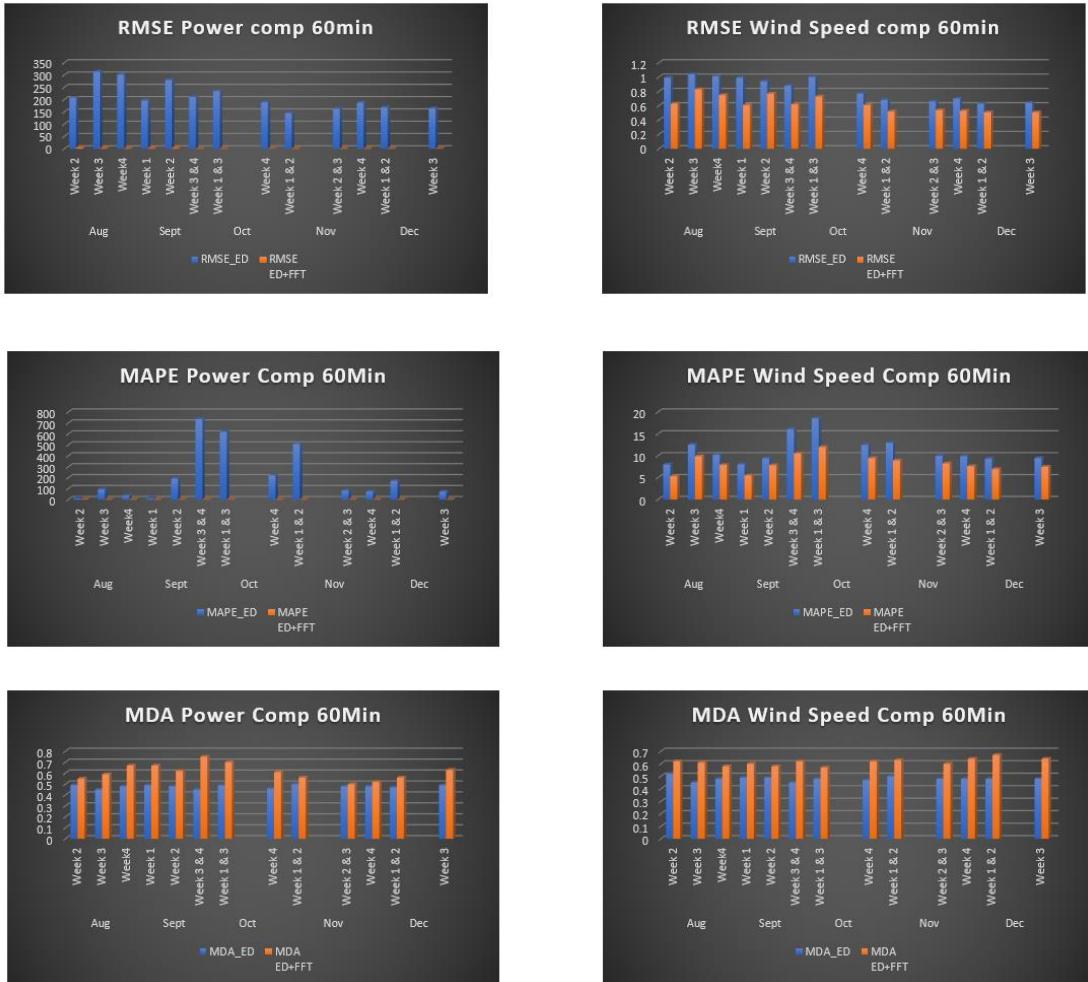


Fig 5.53: Performance comparison of ED and FFT+ED architecture

60Min Ahead Forecasting Window

D] Wavelet Transform:

Wavelet transform is an advanced technique that provides both time-frequency signal representations. It provides good time resolution for high-frequency components and good frequency resolution for low-frequency components.

After implementing wavelet transform on original data, the SNR of both features- Active power and wind speed- crosses the threshold value of 5dB.

SNR of Active Power = 5.81 dB

SNR of Wind Speed = 12.02 dB

An improvement in SNR for both features is observed.

After smoothening, the time series plot for both features are compared with original dataset:

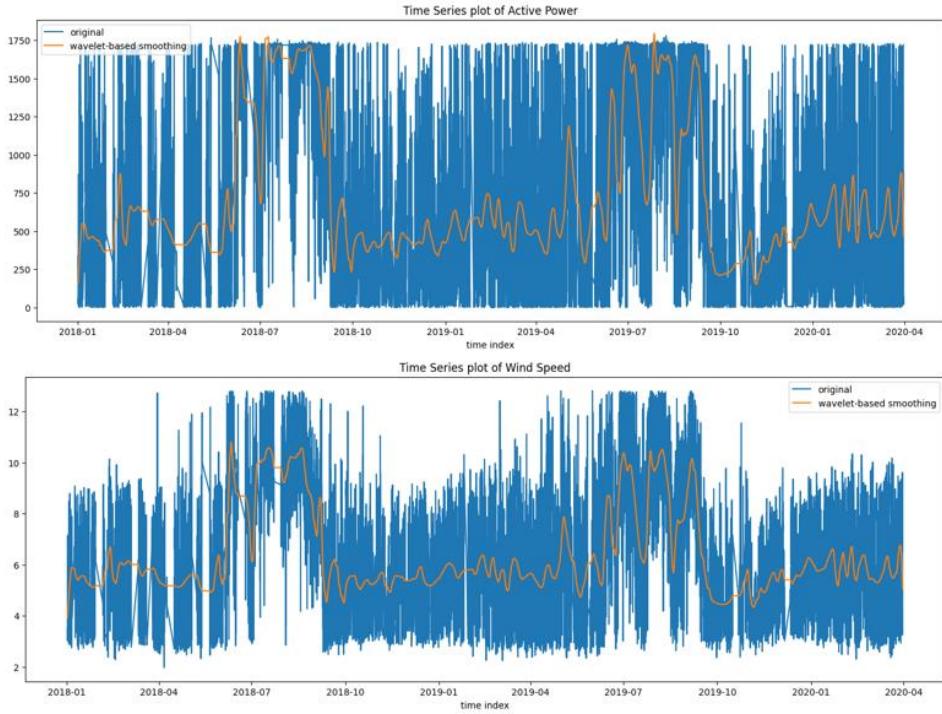
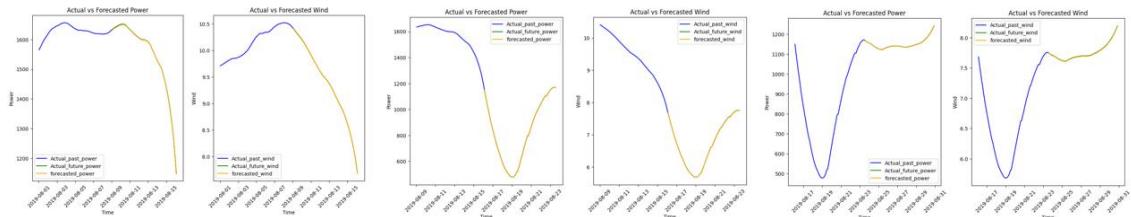


Fig 5.54: Time Series Plots for Wind Speed and Active Power After Smoothing

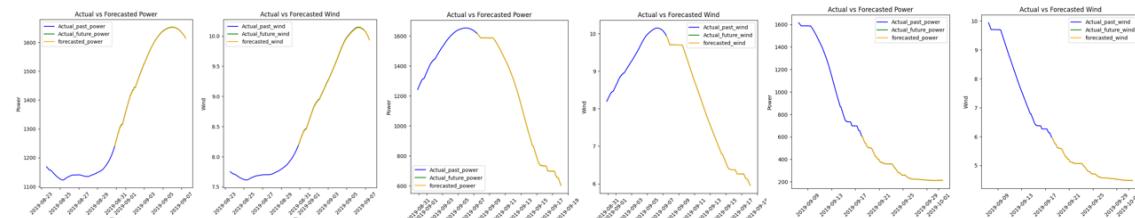
Forecasting results on the test dataset from Aug to Dec for each week are shown below:

- **10 Min Forecasting Ahead:**

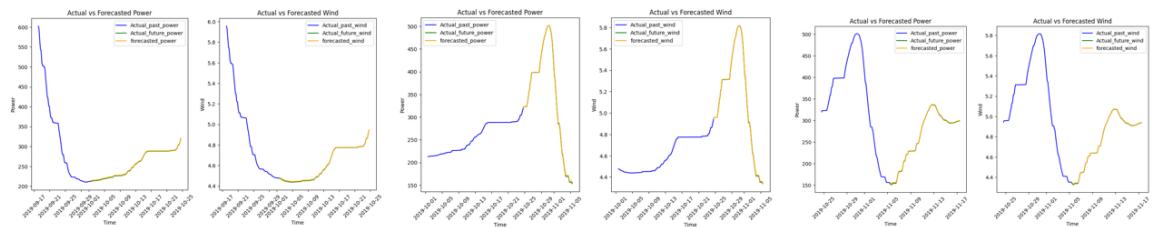
Aug:



Sept:



Oct-Nov:



Nov-Dec:

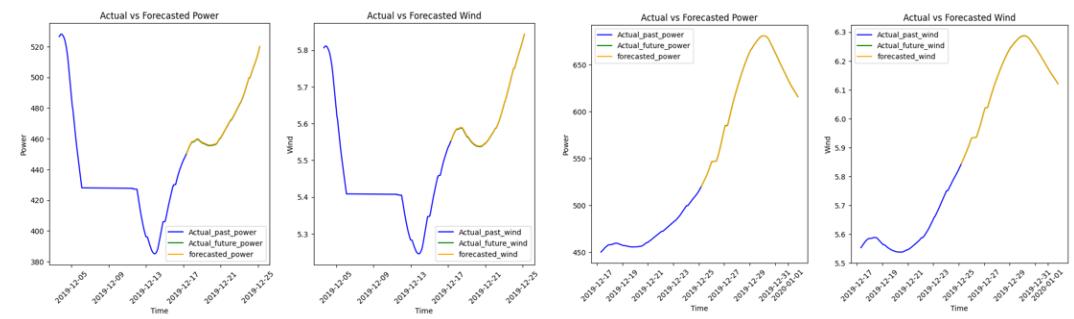
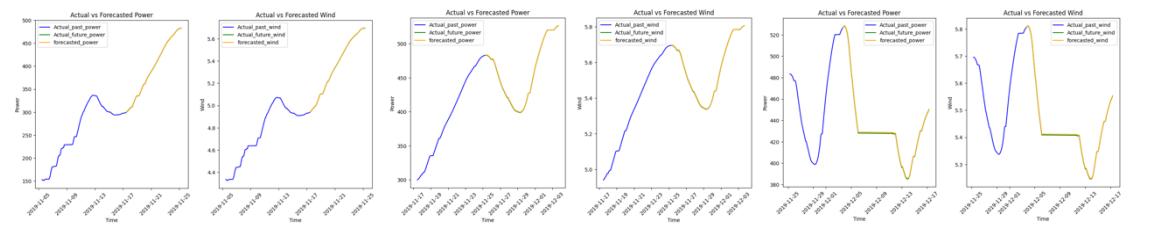


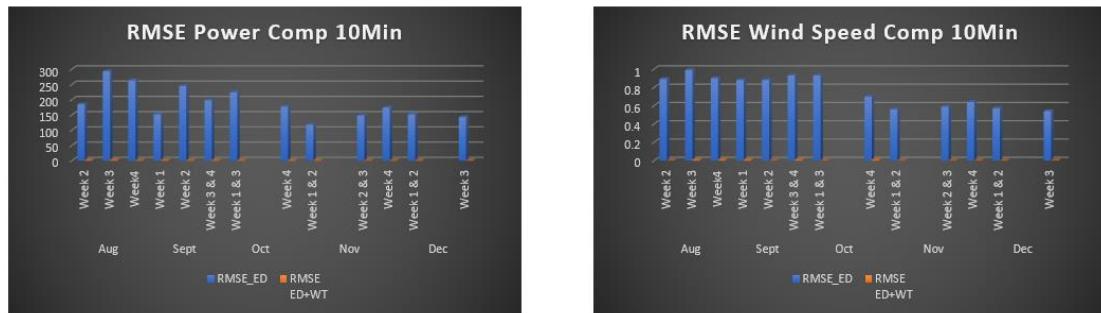
Fig 5.55: Weekly Prediction Plots using WT+ED Model for 10Min Forecasting Window

Model Performance:

WAVELET TRANSFORM									
ENCODER DECODER ARCHITECTURE RESULTS for 10 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week	RMSE	MAPE	MDA
Aug	Week 2	1.48	0.08	0.99	Aug	Week 2	0.001	0.03	1
	Week 3	2.11	0.26	1		Week 3	0.01	0.08	1
	Week4	1	0.08	0.98		Week4	0.01	0.08	1
Sept	Week 1	1.57	0.1	1	Sept	Week 1	0.01	0.1	1
	Week 2	0.9	0.08	1		Week 2	0.001	0.02	1
	Week 3 & 4	1.58	0.58	1		Week 3 & 4	0.01	0.11	1
Oct	Week 1 & 3	1.16	0.46	1	Oct	Week 1 & 3	0.001	0.08	1
	Week 4	1.49	0.44	1		Week 4	0.01	0.09	1
Nov	Week 1 & 2	1.04	0.39	1	Nov	Week 1 & 2	0.001	0.05	0.99
	Week 2 & 3	0.43	0.11	1		Week 2 & 3	0.001	0.02	1
	Week 4	0.54	0.1	1		Week 4	0.001	0.03	1
Dec	Week 1 & 2	0.66	0.14	1	Dec	Week 1 & 2	0.001	0.04	1
	Week 3	0.37	0.08	1		Week 3	0.001	0.02	1

Fig 5.56: Model Performance for WT+ED Model-10 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



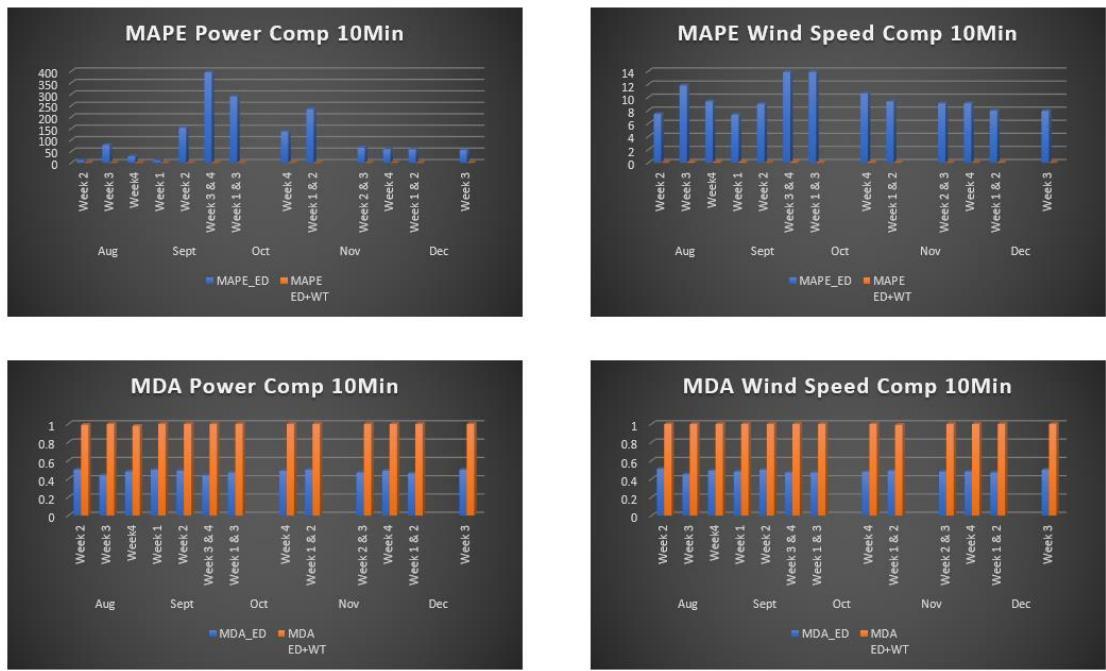
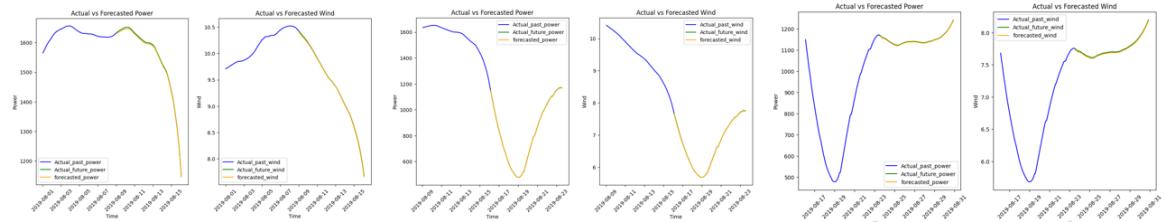


Fig 5.57: Performance comparison of ED and WT+ED architecture

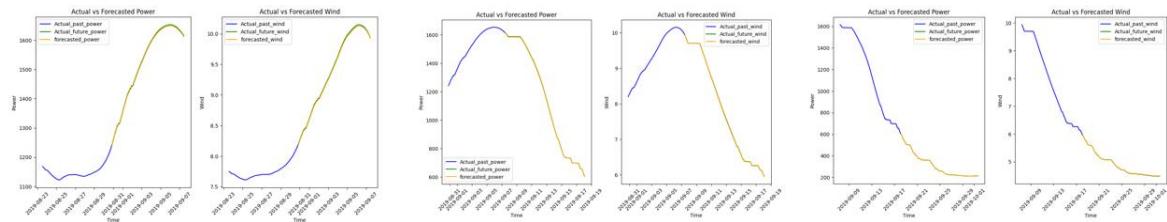
10Min Ahead Forecasting Window

- **30 Min Forecasting ahead**

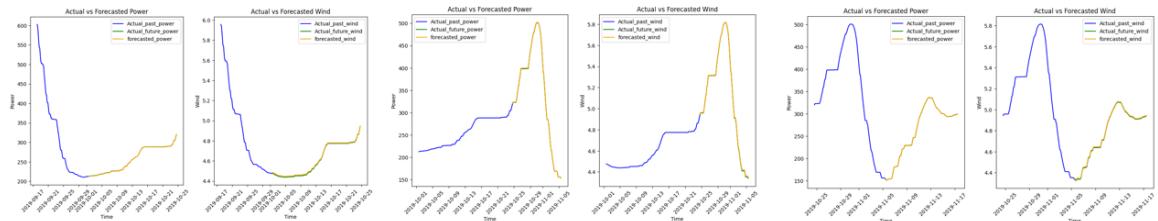
Aug:



Sept:



Oct-Nov:



Nov-Dec:

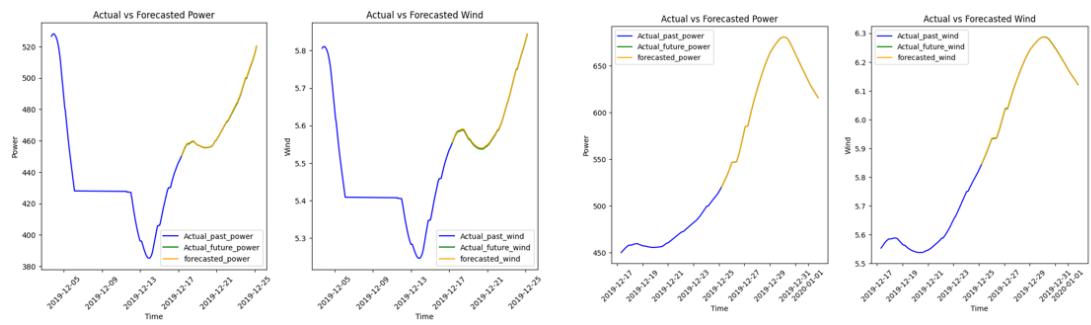
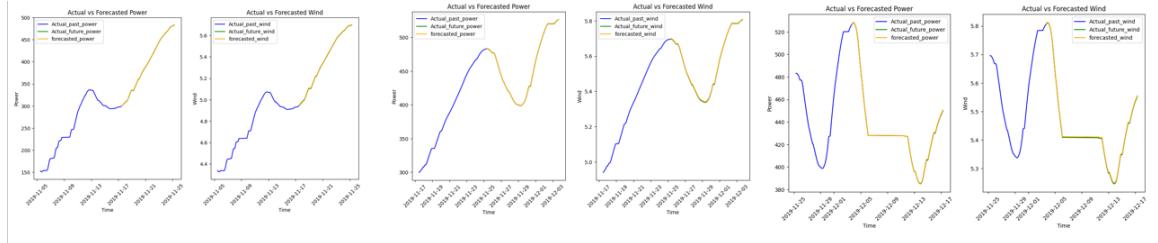


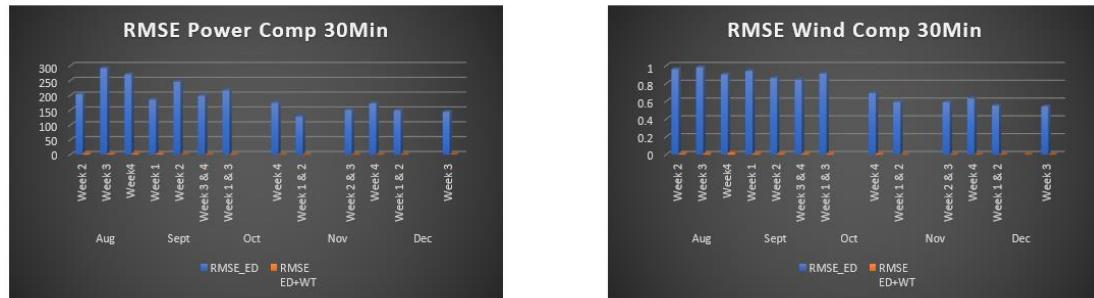
Fig 5.58: Weekly Prediction Plots using WT+ED Model for 30Min Forecasting Window

Performance Results for metrics:

WAVELET TRANSFORM									
ENCODER DECODER ARCHITECTURE RESULTS for 30 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week			
Aug	Week 2	3.58	0.23	0.98	Aug	Week 2	0.01	0.12	1
	Week 3	2.12	0.18	0.98		Week 3	0.01	0.16	1
	Week4	3.25	0.21	1		Week4	0.02	0.17	1
Sept	Week 1	2.97	0.21	1	Sept	Week 1	0.01	0.14	1
	Week 2	0.26	0.07	1		Week 2	0.01	0.12	1
	Week 3 & 4	0.42	0.15	0.99		Week 3 & 4	0.01	0.16	1
Oct	Week 1 & 3	0.79	0.16	0.99	Oct	Week 1 & 3	0.01	0.1	0.99
	Week 4	0.45	0.16	0.99		Week 4	0.01	0.13	0.99
Nov	Week 1 & 2	0.66	0.17	1	Nov	Week 1 & 2	0.001	0.07	1
	Week 2 & 3	0.41	0.09	1		Week 2 & 3	0.001	0.05	1
	Week 4	0.45	0.09	1		Week 4	0.001	0.05	1
Dec	Week 1 & 2	0.48	0.09	1	Dec	Week 1 & 2	0.001 0.001	0.04	1
	Week 3	0.17	0.02	1		Week 3	0.001	0.02	1

Fig 5.59: Model Performance for WT+ED Model-30 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



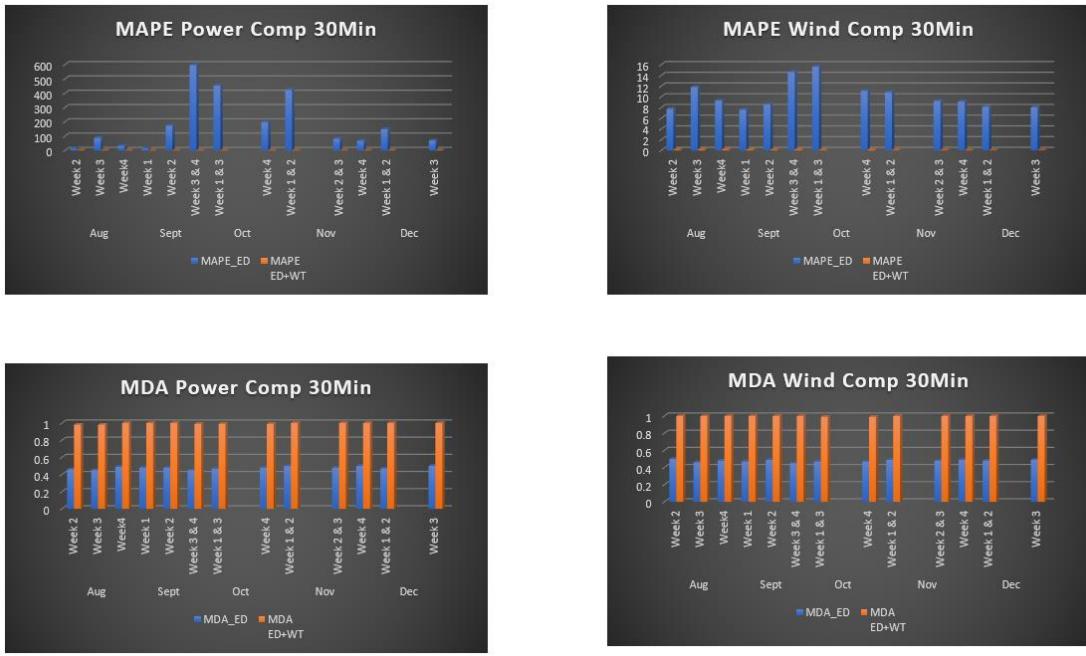
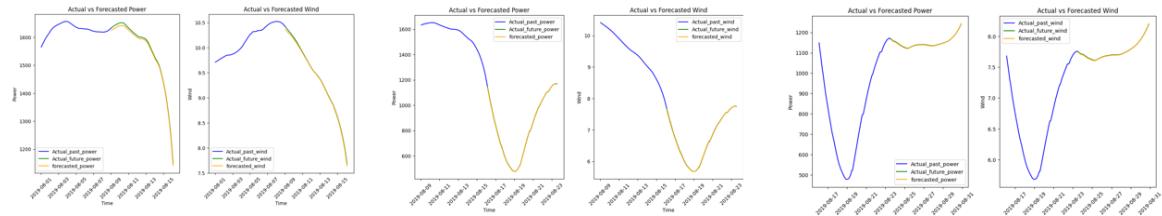


Fig 5.60: Performance comparison of ED and WT+ED architecture

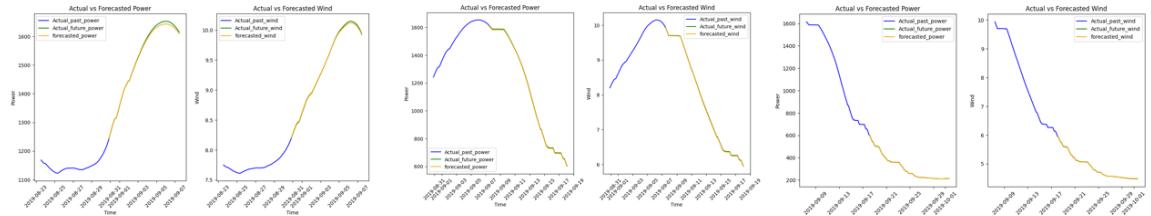
30Min Ahead Forecasting Window

- **60 Min ahead forecasting results:**

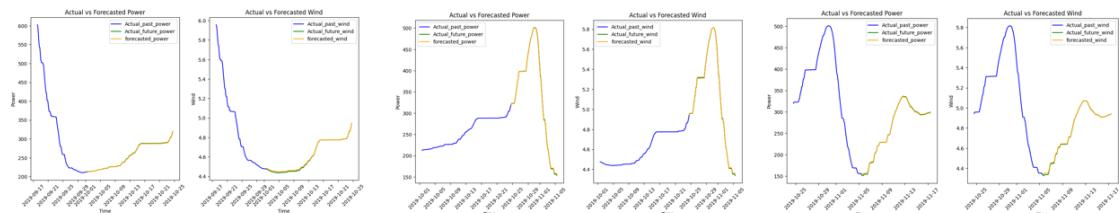
Aug:



Sept:



Oct-Nov:



Nov-Dec:

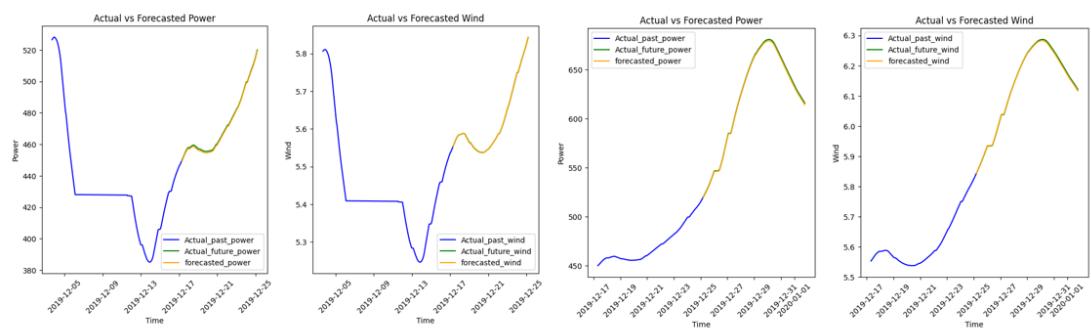
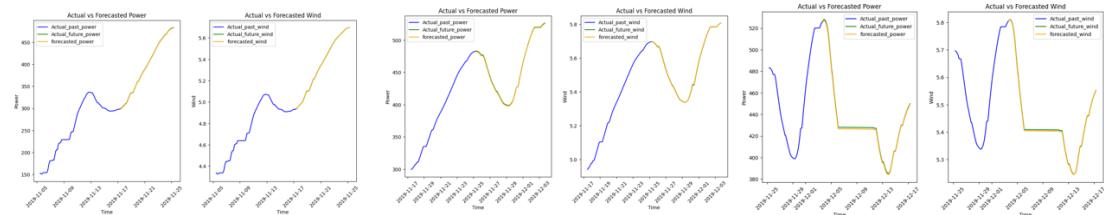


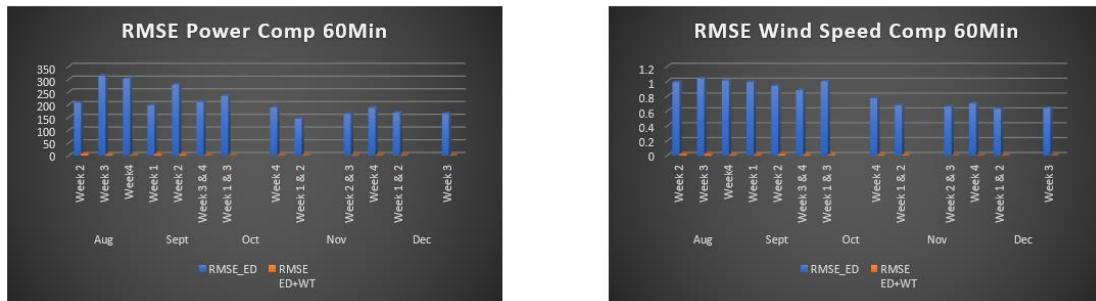
Fig 5.61: Weekly Prediction Plots using WT+ED Model for 60Min Forecasting Window

Model Performance:

WAVELET TRANSFORM									
ENCODER DECODER ARCHITECTURE RESULTS for 60 min									
ACTIVE POWER					WIND SPEED				
Month	Week	RMSE	MAPE	MDA	Month	Week			
Aug	Week 2	6.73	0.4	0.99	Aug	Week 2	0.02	0.19	1
	Week 3	3.98	0.41	0.99		Week 3	0.02	0.21	0.98
	Week4	1.36	0.11	0.96		Week4	0.01	0.06	0.95
Sept	Week 1	5.77	0.29	1	Sept	Week 1	0.01	0.12	1
	Week 2	4.58	0.42	1		Week 2	0.02	0.22	1
	Week 3 & 4	1.62	0.32	1		Week 3 & 4	0.01	0.12	1
Oct	Week 1 & 3	0.79	0.27	0.99	Oct	Week 1 & 3	0.01	0.12	1
	Week 4	1.5	0.38	0.99		Week 4	0.01	0.11	0.99
Nov	Week 1 & 2	1.29	0.47	0.98	Nov	Week 1 & 2	0.01	0.08	0.98
	Week 2 & 3	0.61	0.15	1		Week 2 & 3	0.001	0.03	0.99
	Week 4	0.92	0.18	0.99		Week 4	0.001	0.04	0.99
Dec	Week 1 & 2	1.11	0.23	0.99	Dec	Week 1 & 2	0.001	0.05	0.98
	Week 3	0.73	0.15	0.99		Week 3	0.001	0.01	0.99

Fig 5.62: Model Performance for WT+ED Model-60 Min Forecasting Window

Comparison of results with Encoder Decoder architecture without smoothening technique:



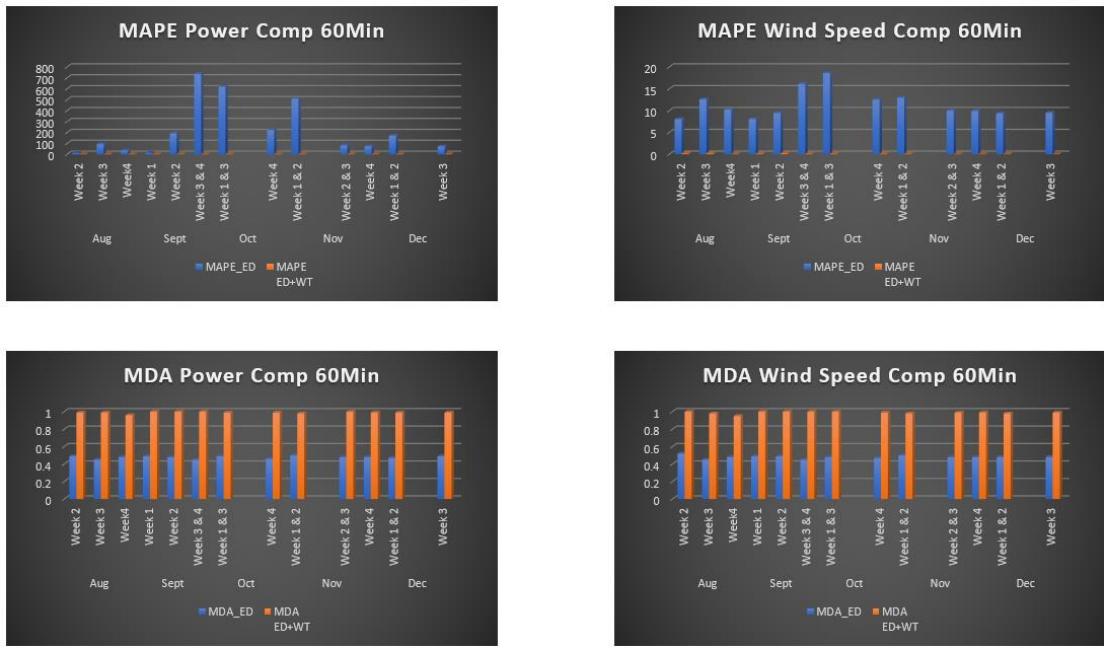
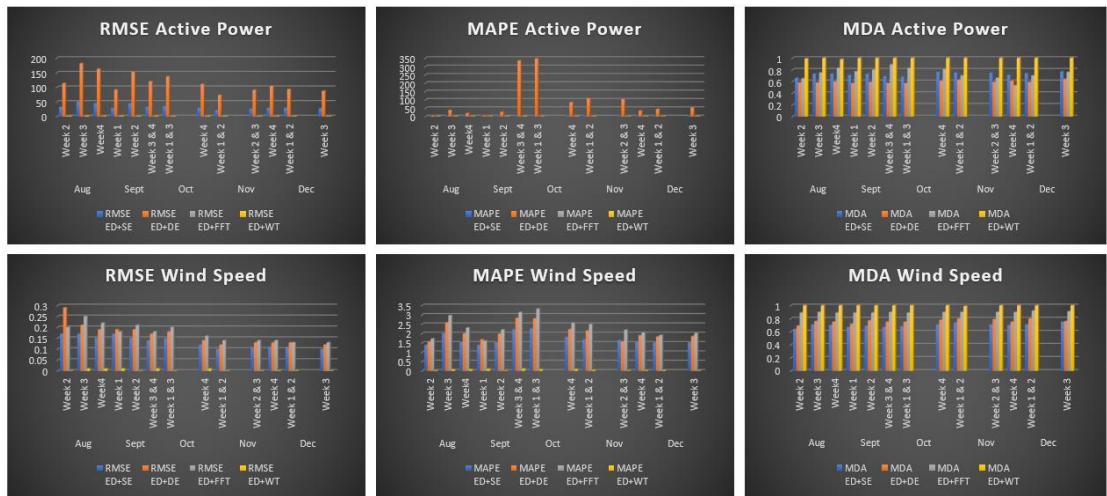


Fig 5.63: Performance comparison of ED and WT+ED architecture

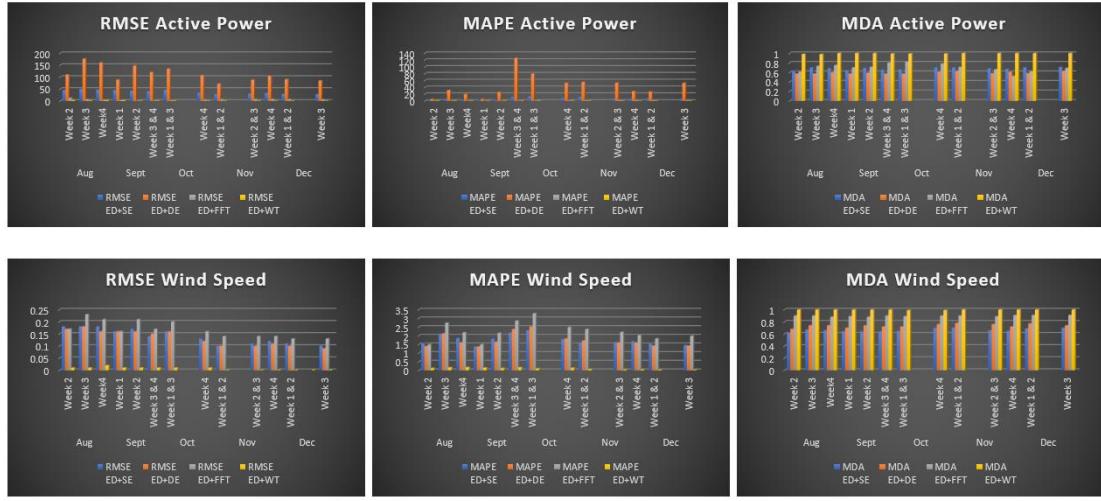
60Min Ahead Forecasting Window

5.2.6 Overall Comparison Among All Smoothening Techniques:

- **10 Min Of Forecasting Window:**



- **30 Min of Forecasting Window:**



- 60 Min of Forecasting Window:**

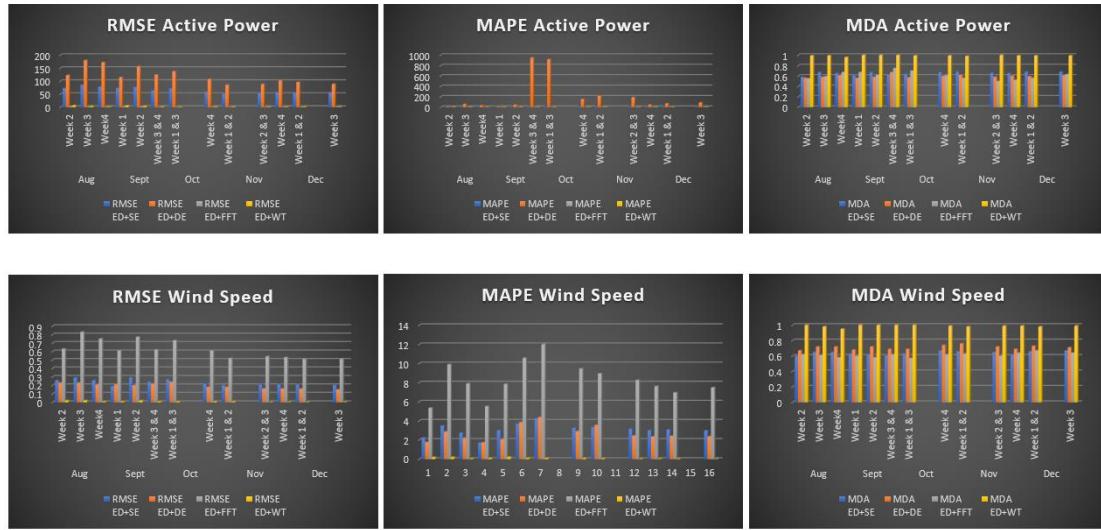


Fig 5.64: Overall Performance Comparison of All Smoothing Techniques

5.3 Results Performed on Dataset-2 based on Ireland Region:

Dataset description is already highlighted in section 4.1.2. Based on required features and preprocessing of data, data is made available for further training.

Model training has been done on a data from Jan 2006 to March 2008. This training involves 2 seasonal periods (Oct- Feb).

Results are classified into two sets of test data. 1st test dataset captures observations from April to August, and 2nd test data captures observations from September to January.

Below is the detailed summary of dataset classification:

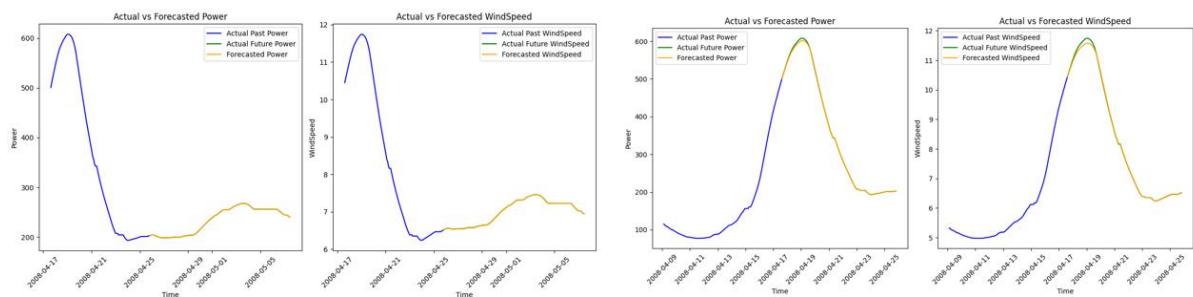
Data	Observations	Training data points
Train	Jan 2006 to March 2008	75000
Test-1	April 2008 to Aug 2009	50000
Test-2	Sept 2009 to Jan 2010	30000

Table 5.3 Dataset-2 training and testing classification

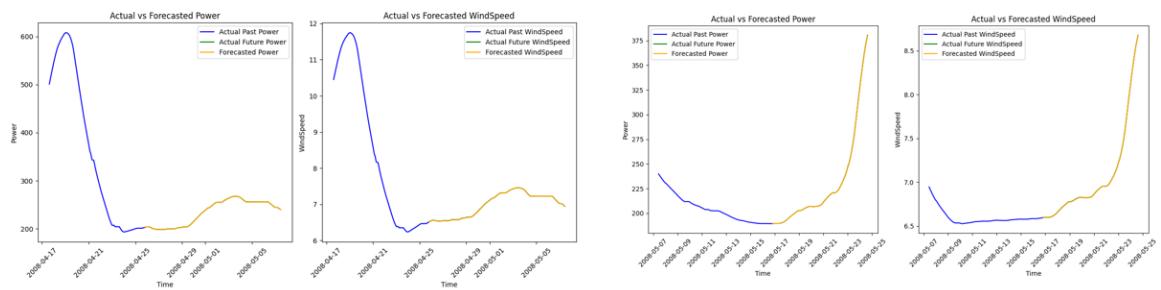
5.3.1 1 step ahead (10 min in future) Forecasting:

Plots for alternate week from every month:

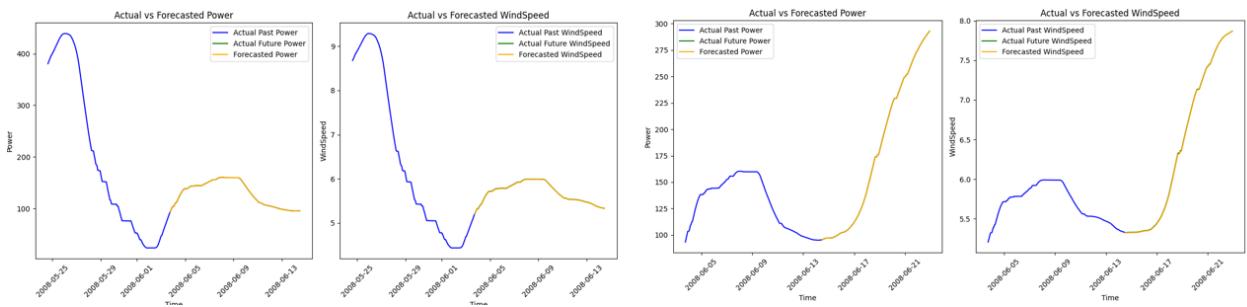
April:



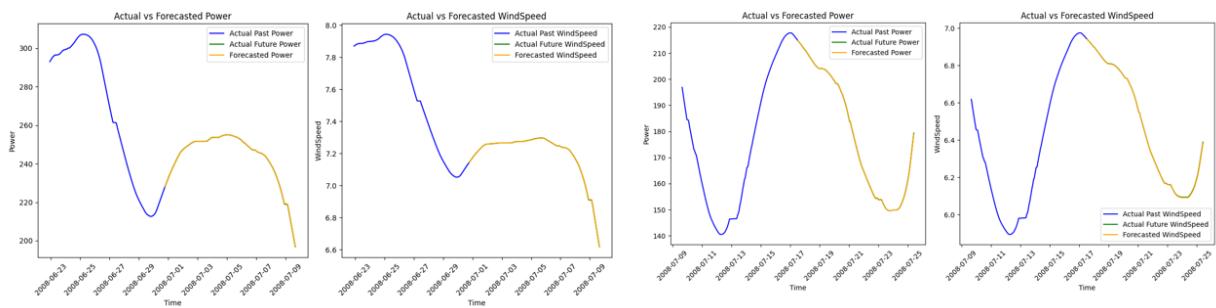
May:



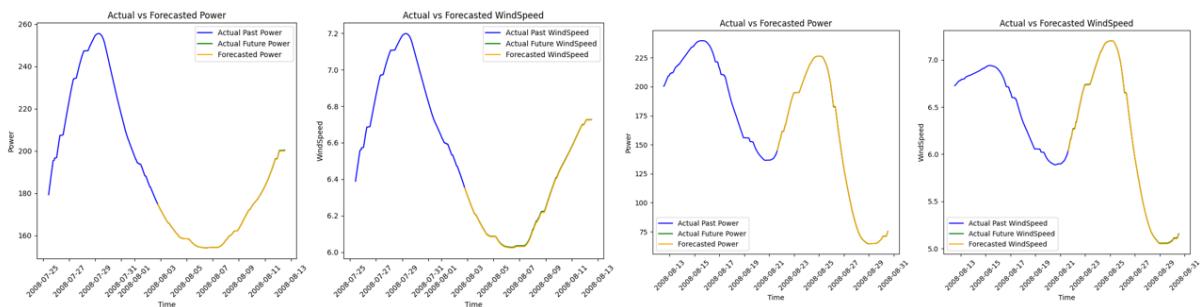
June:



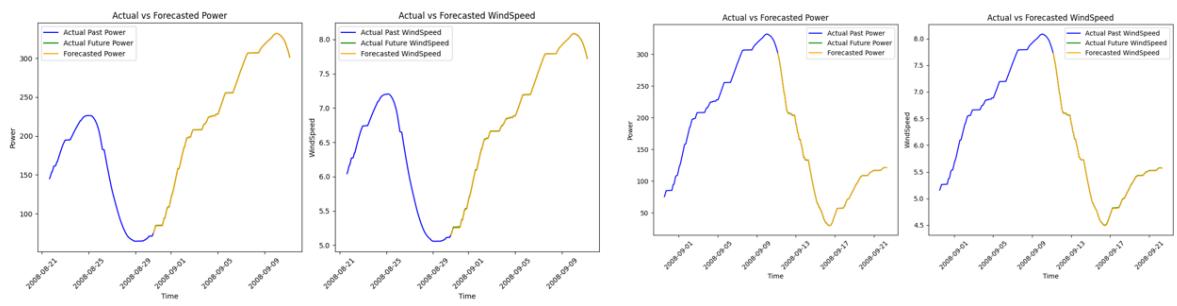
July:



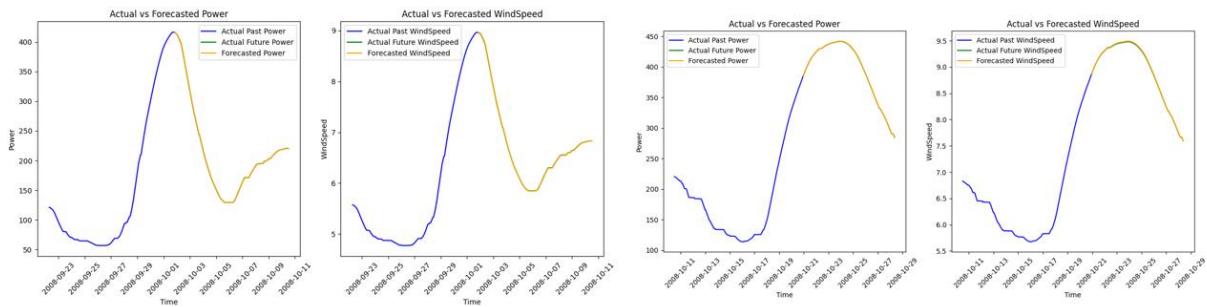
Aug:



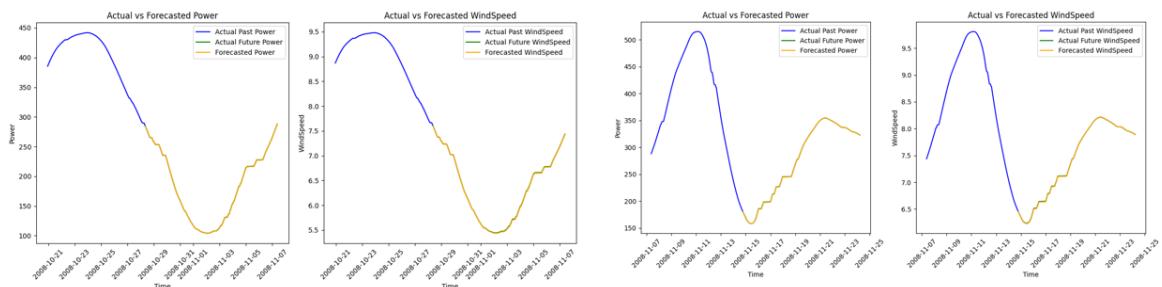
Sept:



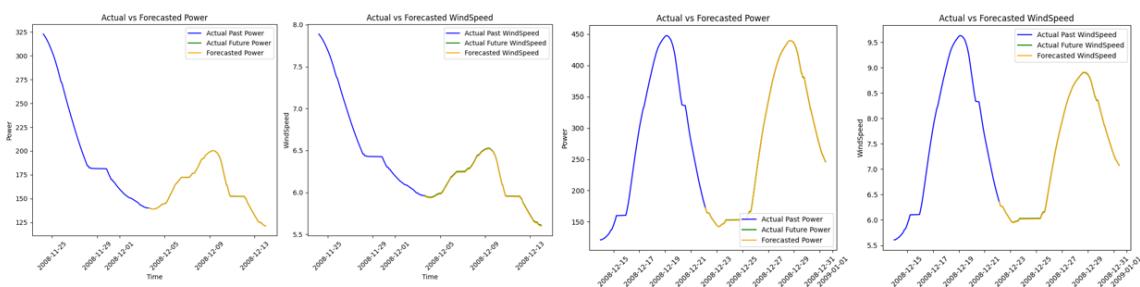
Oct:



Nov:



Dec:



Jan:

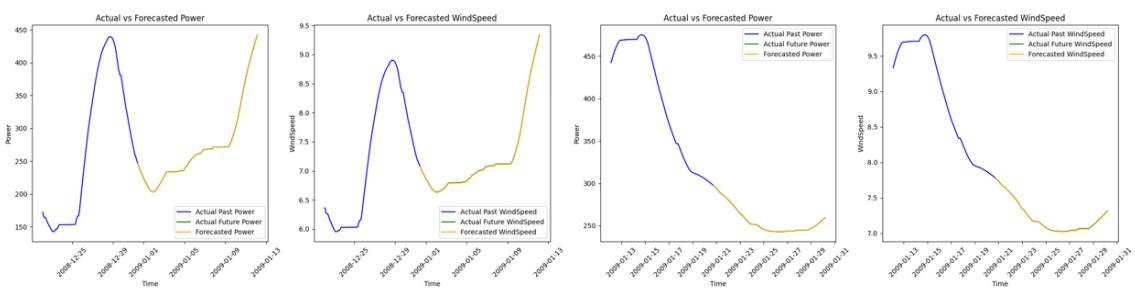


Fig 5.65: Weekly Prediction Plots using WT+ED Model for 10Min Forecasting Window for Dataest-2

All metric results are tabulated below:

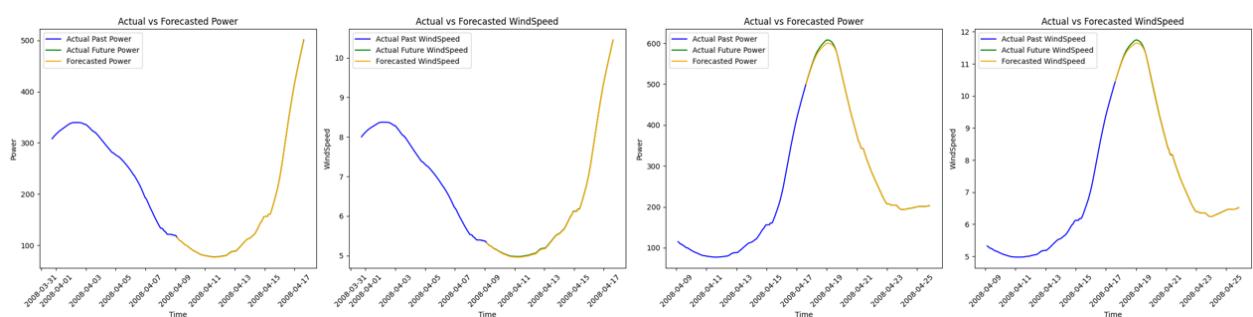
WAVELET TRANSFORM ENCODER DECODER ARCHITECTURE RESULTS for 10 min							
Duration		Power			Wind Speed		
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA
April	Week 2	0.51	0.17	1	0.01	0.07	1
	Week 4	1.76	0.24	0.99	0.03	0.2	0.99
May	Week 1	0.12	0.05	0.99	0.01	0.02	1
	Week 2	0.14	0.06	0.98	0.001	0.01	0.98
	Week 3	0.36	0.1	0.99	0.01	0.03	0.98
June	Week 1	0.74	0.56	1	0.01	0.12	1
	Week 3	0.26	0.18	1	0.01	0.03	0.99
July	Week 1	0.24	0.08	0.99	0.001	0.05	1
	Week 3	0.13	0.07	1	0.001	0.05	1
Aug	Week 1	0.22	0.09	1	0.001	0.03	1
	Week 3	0.28	0.13	1	0.001	0.03	1
Sept	Week 3 & 4	0.38	0.22	1	0.001	0.06	1
Oct	Week 1	0.45	0.17	1	0.001	0.04	1
	Week 3	0.65	0.3	1	0.01	0.15	1
	Week 4	0.33	0.08	0.99	0.001	0.02	0.99
Nov	Week1	0.24	0.07	0.99	0.01	0.06	0.99
	Week3	0.62	0.13	1	0.01	0.1	1
Dec	Week1	0.29	0.08	1	0.01	0.09	1
	Week2	0.31	15	1	0.02	0.03	1
	Week3	0.54	0.35	1	0.02	0.02	1
Jan	Week1	0.48	0.15	0.94	0.01	0.1	1
	Week3	0.66	0.23	0.99	0.01	0.15	0.98

Fig 5.66 : Performance metrics of ED + Wavelet Transform model for 1 step ahead (10 min ahead) forecasting for Dataset-2

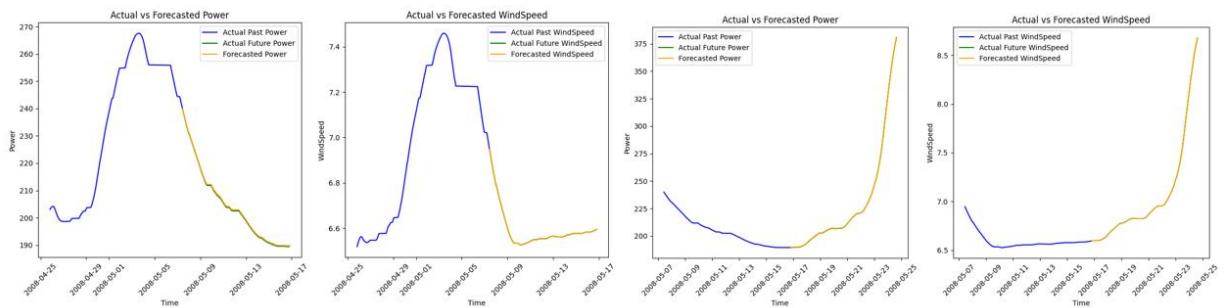
5.3.2 3 step ahead (30 min in future) Forecasting:

Plots for each week from every month:

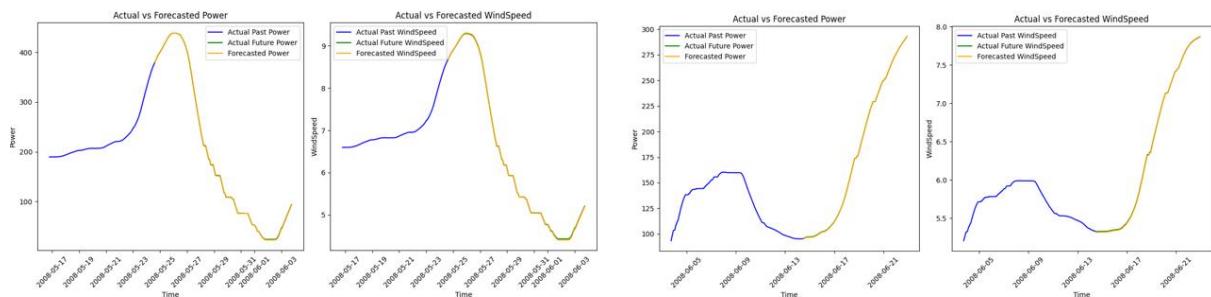
April:



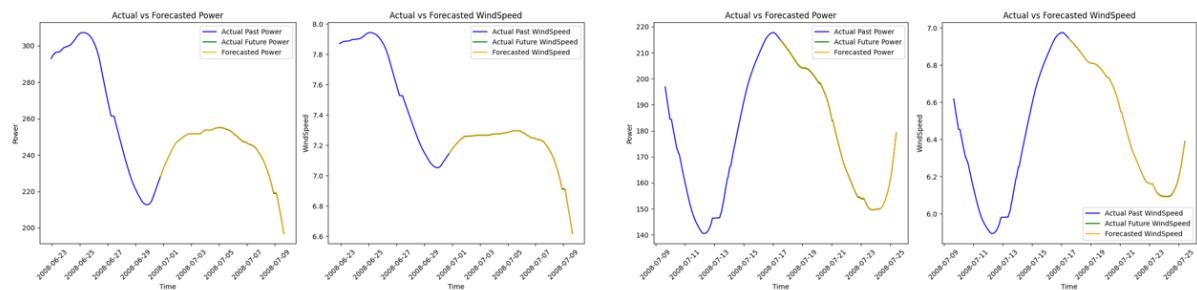
May:



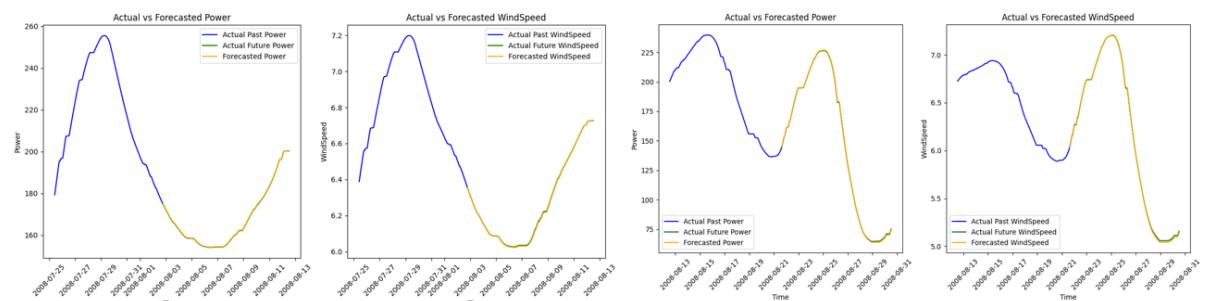
June:



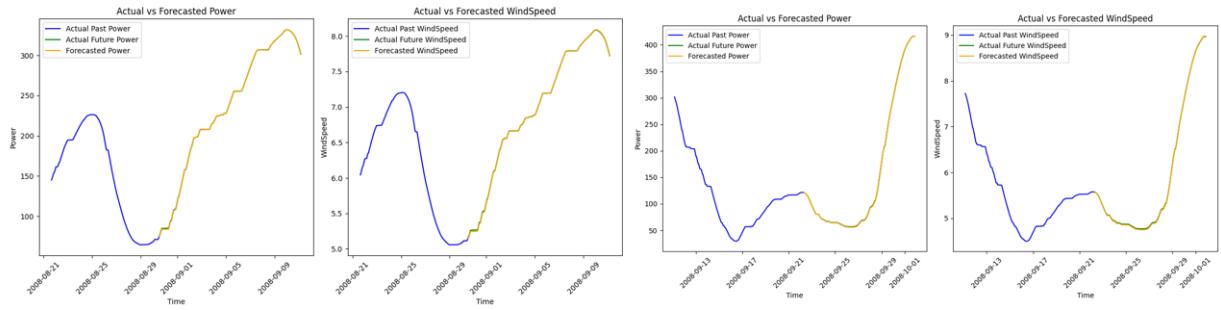
July:



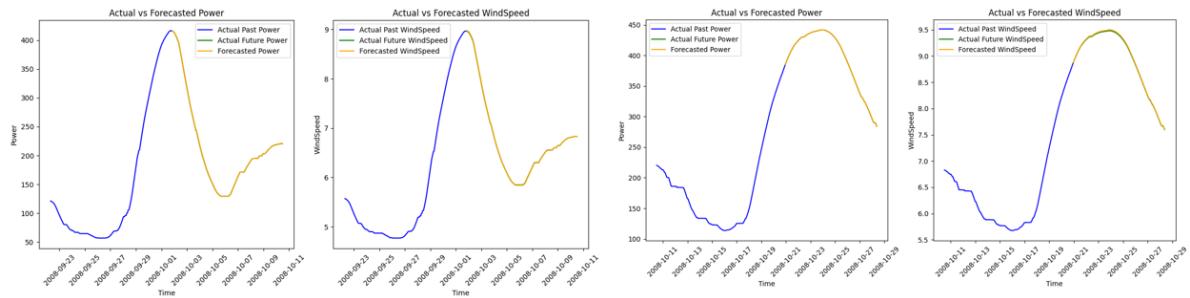
Aug:



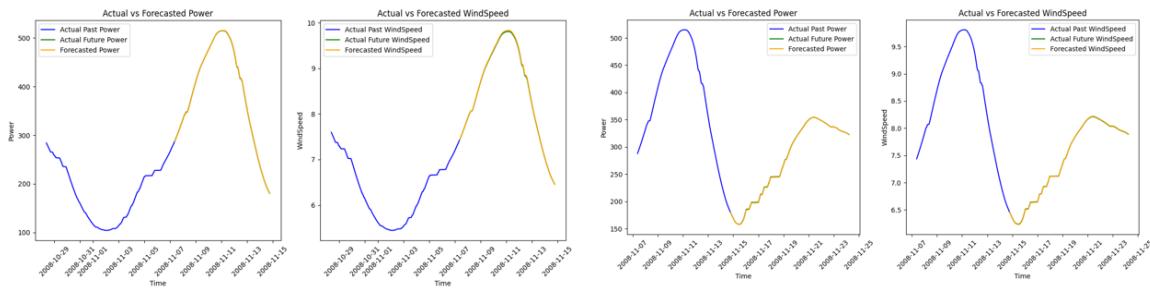
Sept:



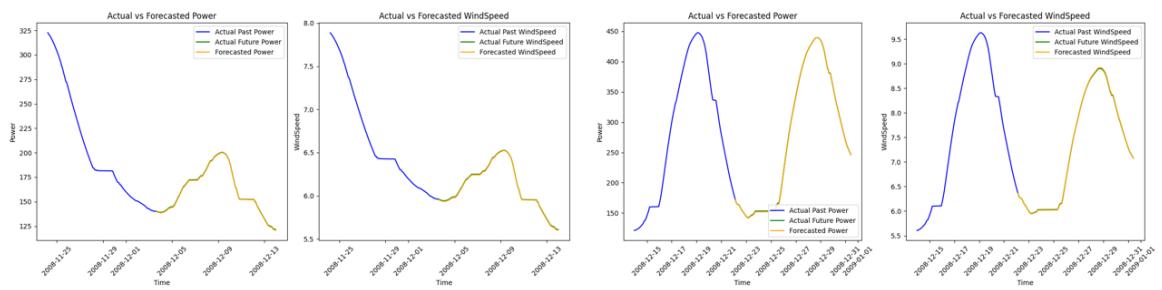
Oct:



Nov:



Dec:



Jan:

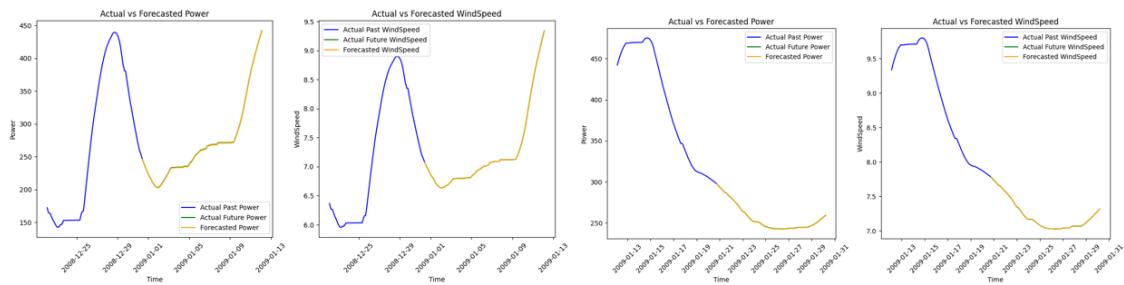


Fig 5.67: Weekly Prediction Plots using WT+ED Model for 30Min Forecasting Window for Dataest-2

All the metric results are tabulated:

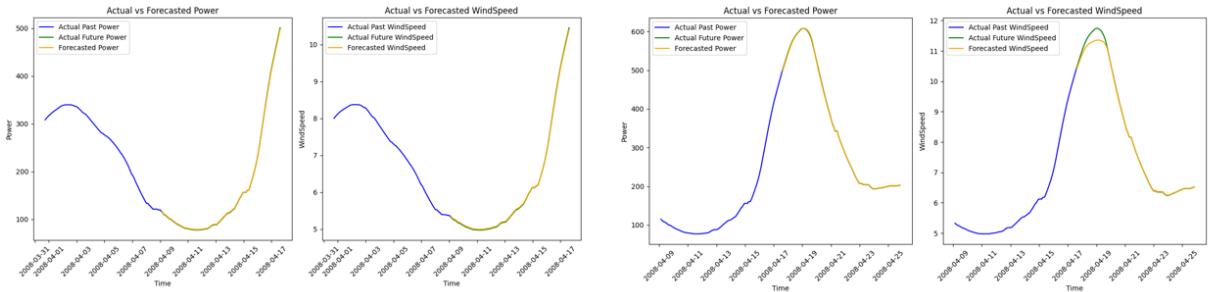
WAVELET TRANSFORM ENCODER DECODER ARCHITECTURE RESULTS for 30 min							
Duration		Power			Wind Speed		
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA
April	Week 2	0.84	0.27	1	0.01	0.2	1
	Week 4	1.49	0.18	0.99	0.06	0.33	0.99
May	Week 1	0.17	0.06	0.99	0.01	0.03	1
	Week 2	0.2	0.08	1	0.001	0.05	0.98
	Week 3	0.49	0.11	0.98	0.01	0.05	0.98
June	Week 1	0.67	0.48	1	0.01	0.21	0.99
	Week 3	0.25	0.16	1	0.01	0.15	0.99
July	Week 1	0.18	0.06	1	0.001	0.04	0.99
	Week 3	0.21	0.1	0.99	0.01	0.09	0.98
Aug	Week 1	0.23	0.08	1	0.001	0.03	1
	Week 3	0.08	0.03	1	0.01	0.07	1
Sept	Week 3 &	0.32	0.18	1	0.001	0.07	1
Oct	Week 1	0.72	0.28	1	0.01	0.08	0.98
	Week 3	1.23	0.47	0.99	0.01	0.14	1
	Week 4	0.77	0.2	0.99	0.001	0.02	1
Nov	Week1	0.41	0.13	0.99	0.01	0.04	0.99
	Week3	1.01	0.23	1	0.01	0.05	1
Dec	Week1	0.15	0.03	1	0.01	0.06	1
	Week2	0.2	0.04	1	0.01	0.08	1
	Week3	0.14	0.09	1	0.01	0.13	1
Jan	Week1	1.14	0.45	0.92	0.01	0.1	1
	Week3	0.51	0.17	0.99	0.01	0.16	0.98

Fig 5.68: Performance metrics of ED + Wavelet Transform model for 3 step ahead (30 min ahead) forecasting

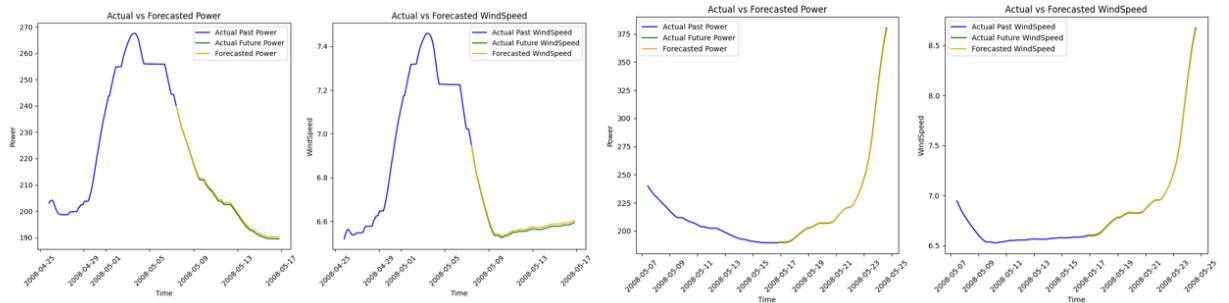
5.3.3 6 step ahead (60 min in future) Forecasting:

Plots for each week from every month:

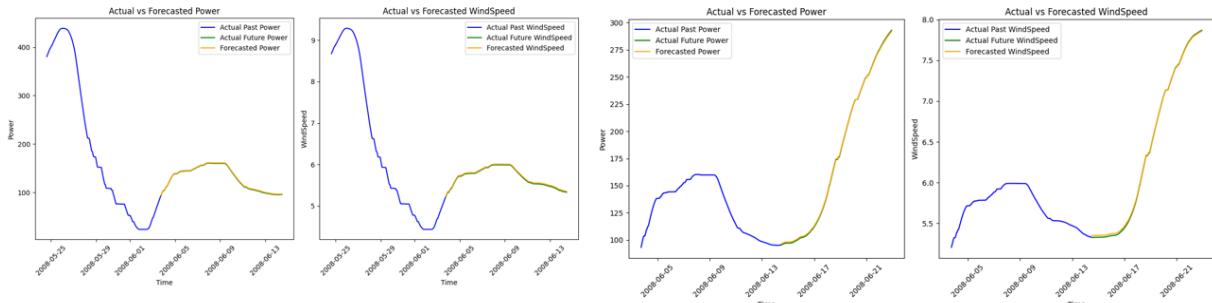
April:



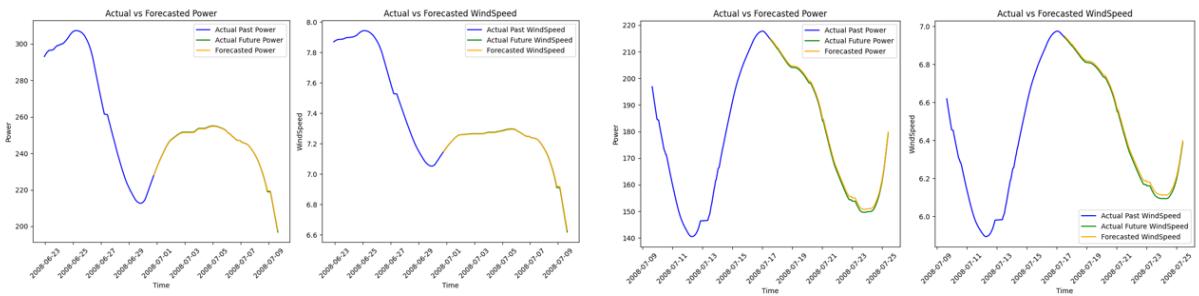
May:



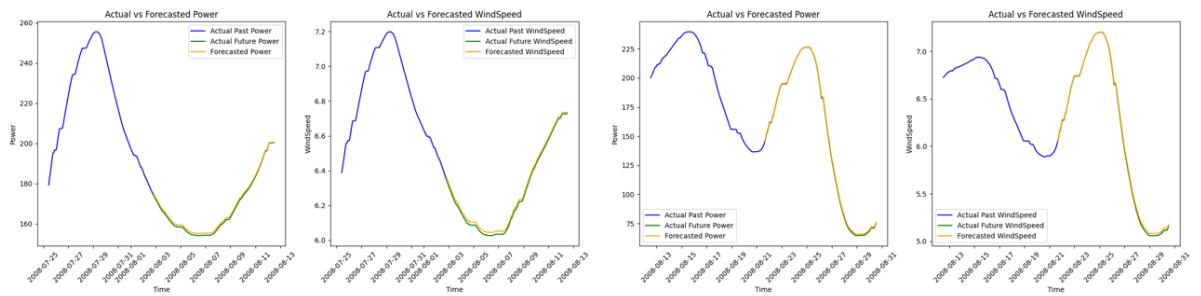
June:



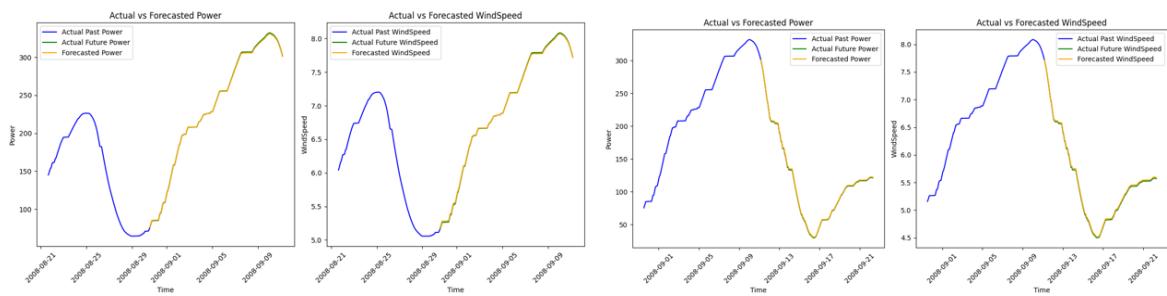
July:



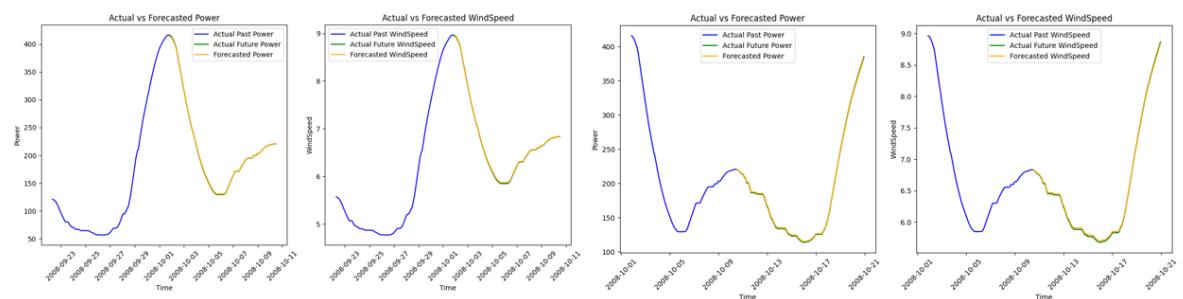
Aug:



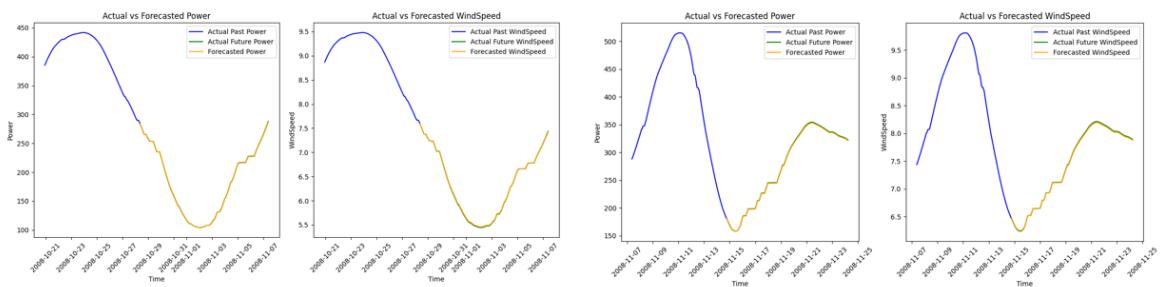
Sept:



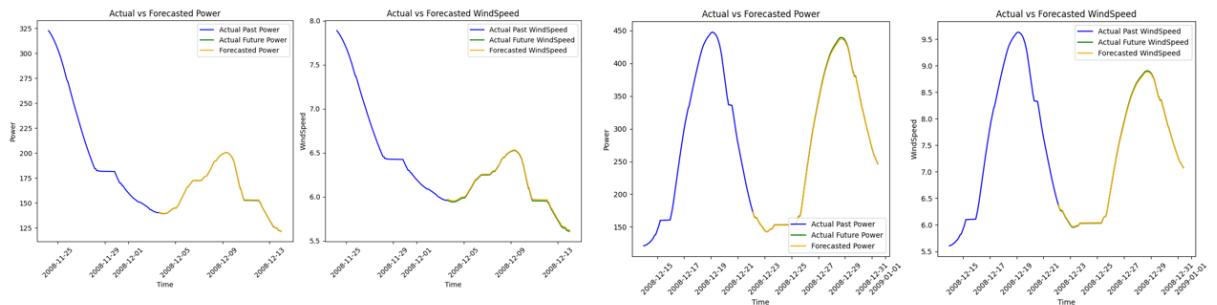
Oct:



Nov



Dec:



Jan:

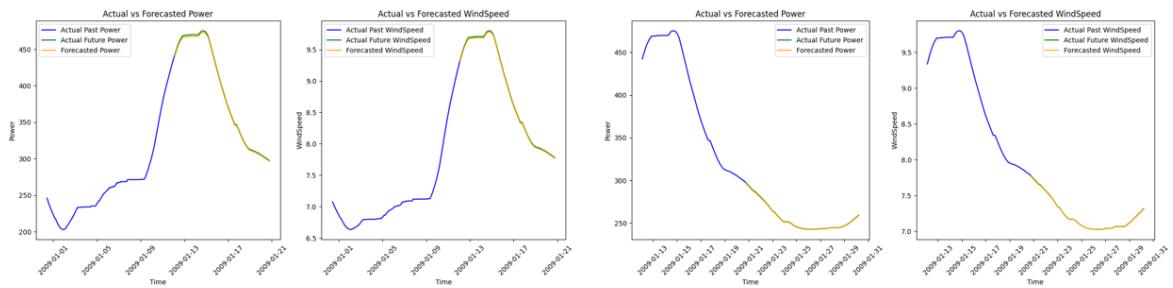


Fig 5.69: Weekly Prediction Plots using WT+ED Model for 60Min Forecasting Window for Dataest-2

All the metric results are tabulated below:

WAVELET TRANSFORM ENCODER DECODER ARCHITECTURE RESULTS for 60 min							
Duration		Power			Wind Speed		
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA
April	Week 2	2.71	0.39	1	0.06	0.73	1
	Week 4	3.93	0.64	0.99	0.22	1.3	0.99
May	Week 1	0.49	0.19	0.99	0.01	0.12	0.99
	Week 2	0.81	0.39	1	0.01	0.2	0.98
	Week 3	1.34	0.33	0.98	0.03	0.23	0.98
June	Week 1	3.16	1.12	1	0.06	0.83	0.99
	Week 3	0.49	0.17	1	0.03	0.38	1
July	Week 1	0.37	0.11	1	0.01	0.08	0.98
	Week 3	0.74	0.38	0.99	0.02	0.31	0.98
Aug	Week 1	0.85	0.32	1	0.01	0.14	1
	Week 3	0.67	0.33	0.99	0.02	0.26	1
Sept	Week 3 & 4	1.33	1.01	0.99	0.03	0.44	1
Oct	Week 1	1.43	0.63	1	0.03	0.33	0.98
	Week 3	5.28	1.28	0.99	0.09	0.96	1
	Week 4	5.59	1.15	0.99	0.09	0.95	0.99
Nov	Week1	2.19	0.63	0.99	0.05	0.59	0.99
	Week3	6.33	1.45	1	0.13	1.33	1
Dec	Week1	2.17	0.55	1	0.05	0.43	1
	Week2	0.35	0.13	1	0.01	0.18	1
	Week3	0.04	0.02	0.99	0.02	0.31	1
Jan	Week1	0.57	0.19	0.93	0.02	0.19	1
	Week3	2.58	0.73	0.99	0.07	0.67	0.98

Fig 5.70: Performance metrics of ED + Wavelet Transform model for 6 step ahead (60 min ahead) forecasting

5.4 Calculation of Potential Wind Power and Average power loss:

As per [46],

Wind Power Density (WPD), which represents the power available per unit area perpendicular to the wind flow. WPD depends on wind speed, air density, and the rotor's swept area and increases with the cube of wind velocity, formulated as:

where:

- P = is the power generation by the wind (W/m^2),

- ρ = is the air density (typically 1.225 kg/m^3 at sea level),
 - A = is the rotor swept area (m^2),
 - V = is the wind speed (m/s).

Since wind speed fluctuates, the average of the cube of different wind speeds can be significantly higher than the cube of the average speed. To improve accuracy, a correction factor known as the Wind Energy Pattern Factor (WEPF) is introduced, defined as:

where:

v_i represents individual wind speed observations

n is the number of data points

The WEPF typically ranges from 0 to 4, depending on how effectively a wind turbine can optimize energy extraction. Using the WEPF, the mean WPD at a location can be reformulated as:

This adjusted equation accounts for variability in wind speed, providing a more accurate estimate of potential wind power at a site.

So, potential power at each time step will be calculated as per equation 5.3. The difference between the Average of potential power and the average of calculated power will decide the average loss in the power generation.

Loss between potential power generation and one step ahead of forecasted power:

Specifications of Vestas wind turbine are tabulated below:

Characteristic	Details
Model Name	Vestas V52
Rated power	0.9 MW
Hub height	78.98 m
Rotor diameter (D)	52.2 m
Swept area (A)	2140 m ²
Cut-in wind speed	3.5 m/s
Rated wind speed	11 m/s
Cut-off wind speed	14.9 m/s

Table 5.4 Dataset-2 Characteristics of turbine

The methodology to calculate the loss is as shown in fig:

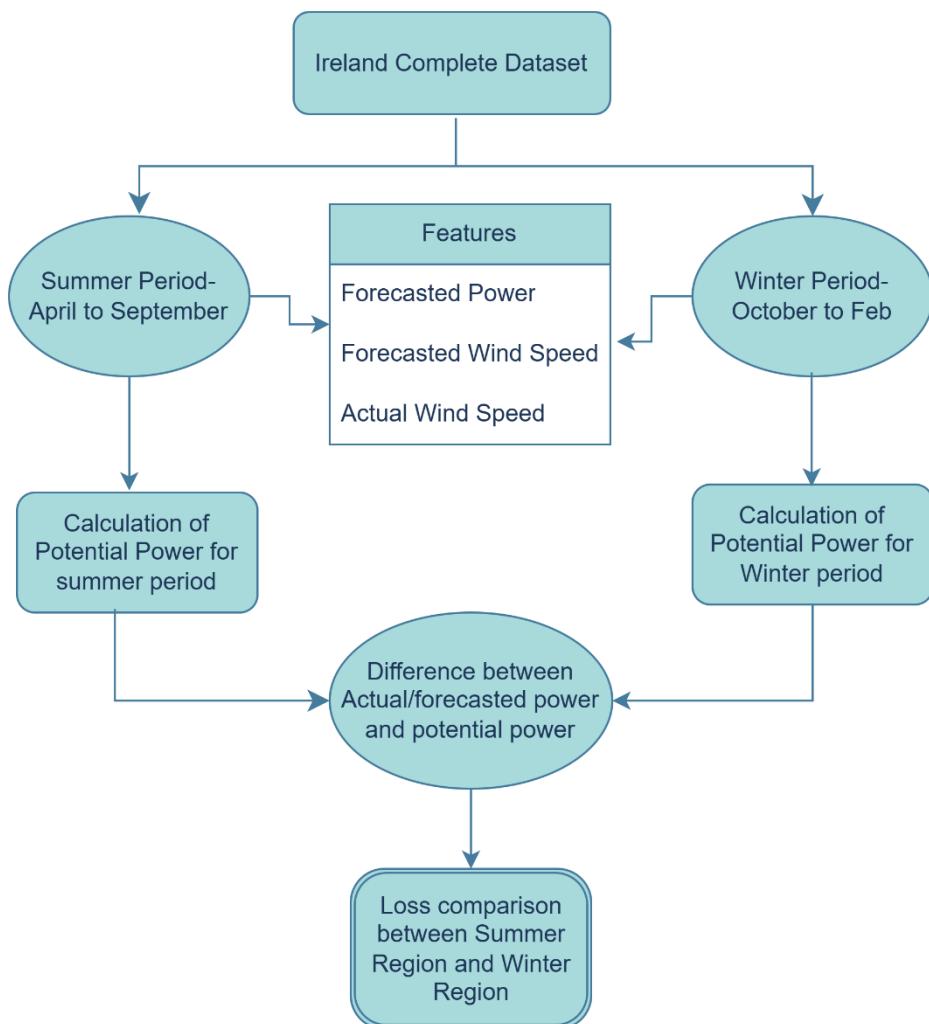


Fig 5.71: Methodology to understand loss in summer and winter region

As per the methodology, forecasted results have been compared with potential power that would have been possible.

Once potential power is calculated, it is compared with summer period forecasting and winter period forecasting to determine the loss.

Loss between potential power generation and one step ahead of forecasted power

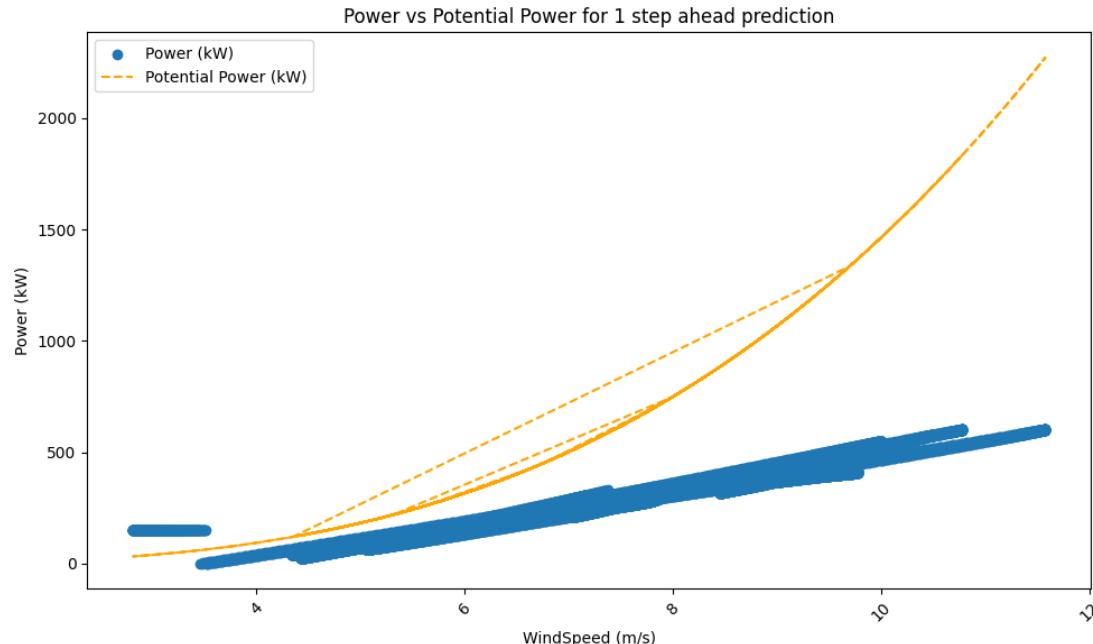


Fig 5.72 : Potential Power Vs Forecasted power generated in the summer period for 1 step ahead forecasting

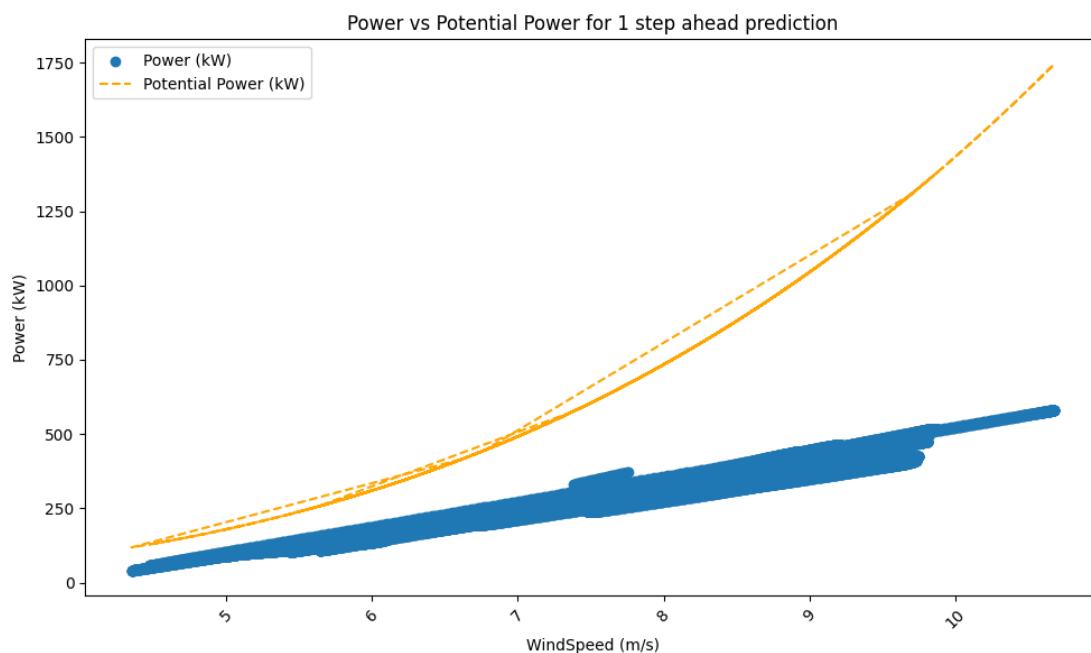


Fig 5.73: Potential Power Vs Forecasted power generated in the winter period for 1 step ahead forecasting

Loss between potential power generation and three steps ahead of forecasted power

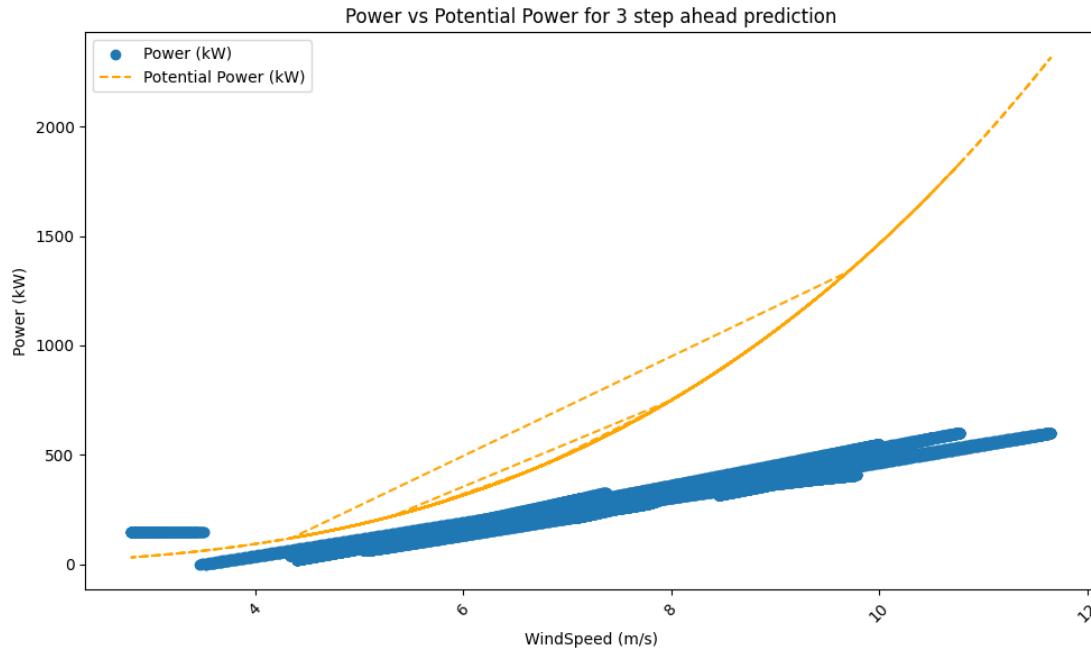


Fig 5.74: Potential Power Vs Forecasted power generated in the summer period for 3 step ahead forecasting

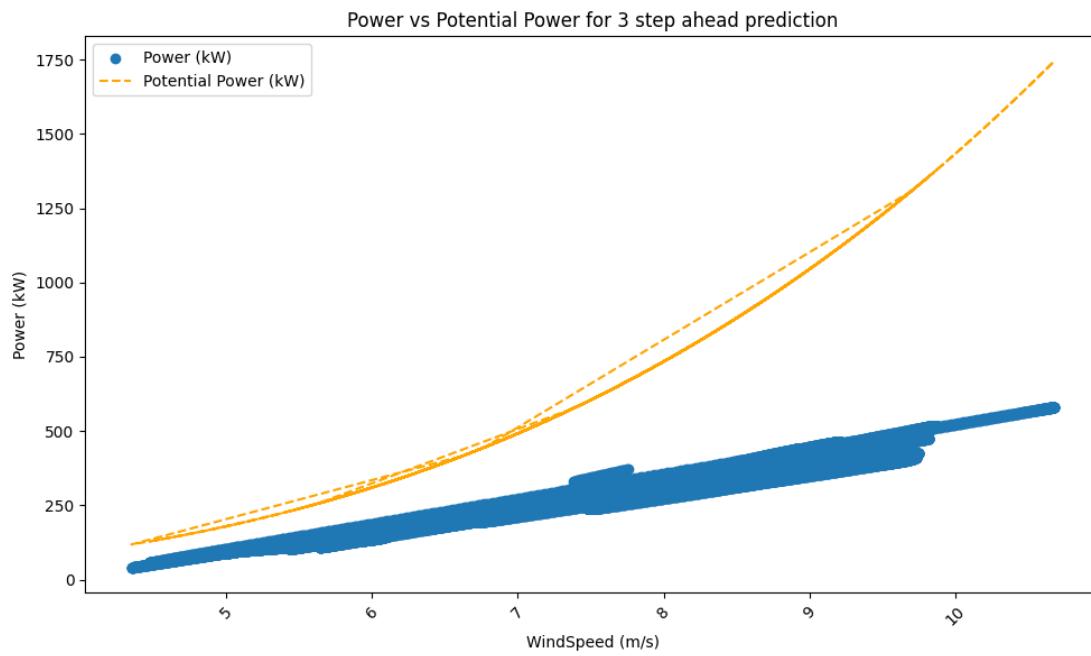


Fig 5.75: Potential Power Vs Forecasted power generated in the winter period for 3 step ahead forecasting

Loss between potential power generation and six steps ahead of forecasted power

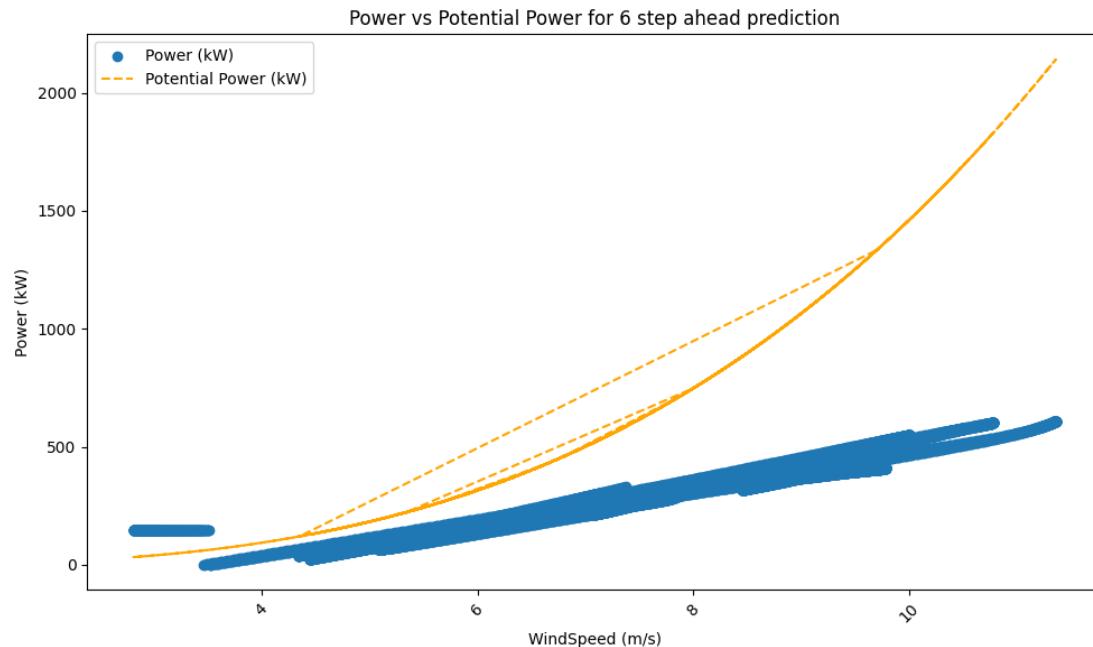


Fig 5.76: Potential Power Vs Forecasted power generated in the summer period for 6 step ahead forecasting

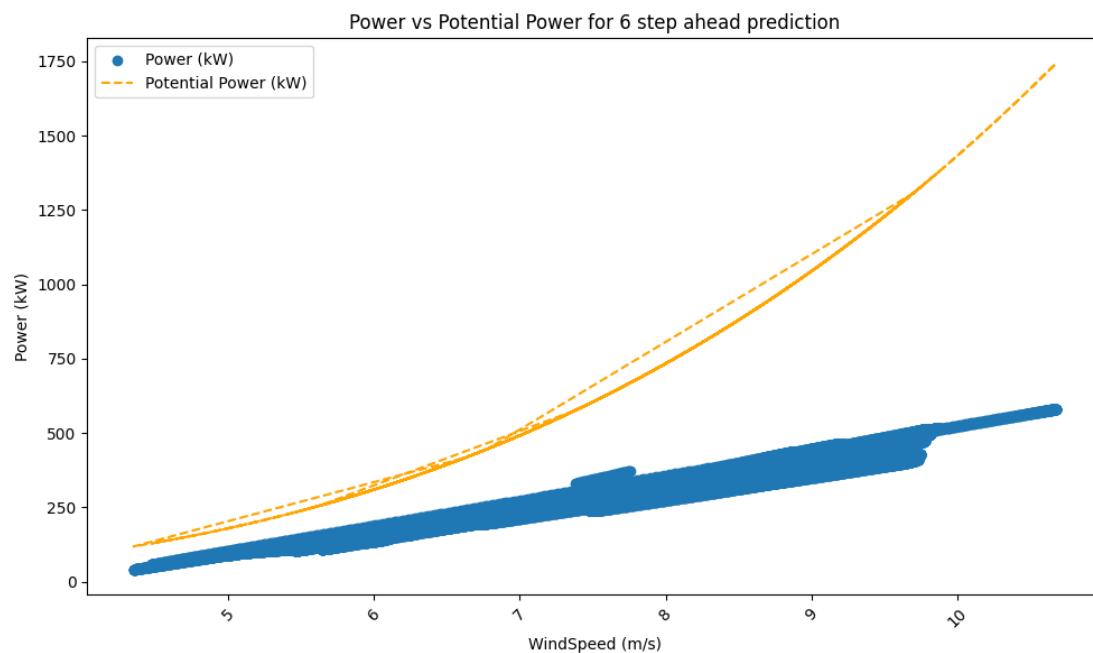


Fig 5.77: Potential Power Vs Forecasted power generated in the winter period for 6 step ahead forecasting

Average loss calculations:

DATASET-2 AVERAGE LOSS			
Forecasting steps	Summer Period loss	Winter Period loss	% Loss difference
1 step	51.23	54.89	3.66
3 step	50.24	54.93	4.69
6 step	51.19	54.88	3.69

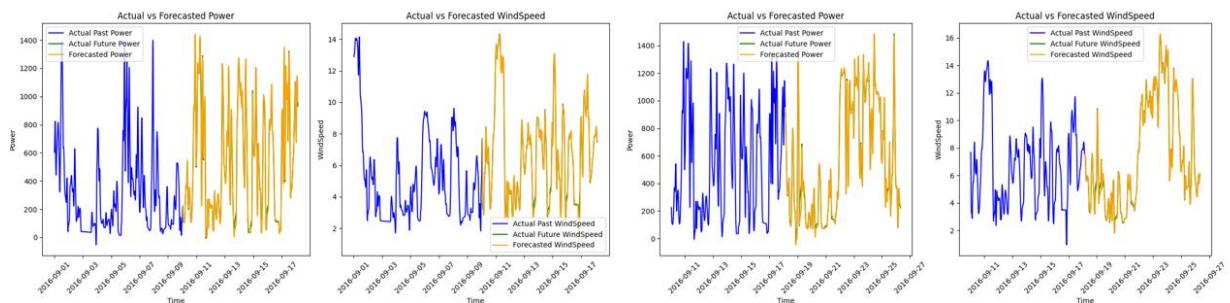
Table 5.5: Average loss calculations for dataset-2

5.5 Results Performed on Dataset-3 based on Kuwait (High-Temperature Region) Region:

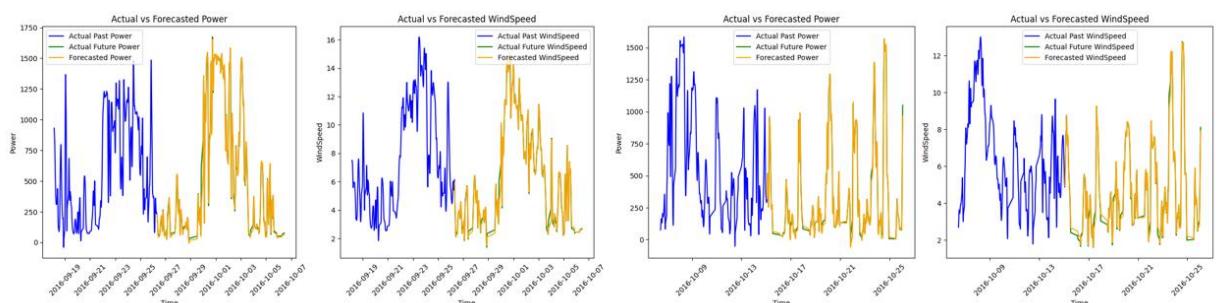
5.5.1 1 step (10 min ahead of forecasting) Plots:

Plots for each week from every month:

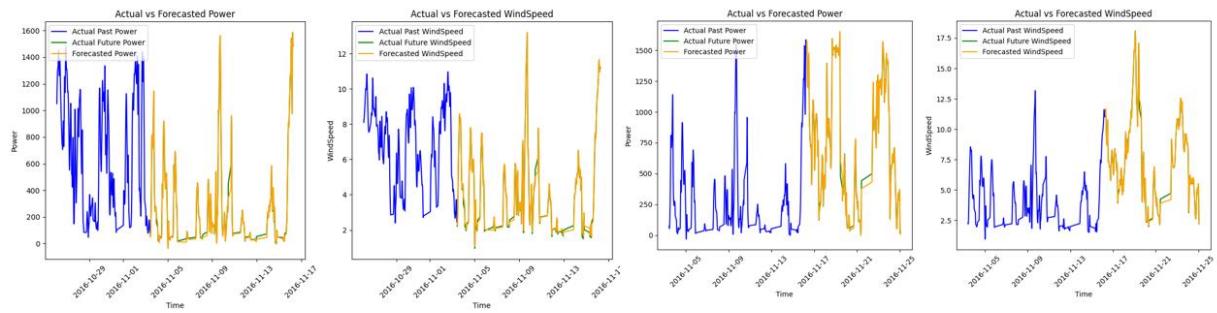
Sept:



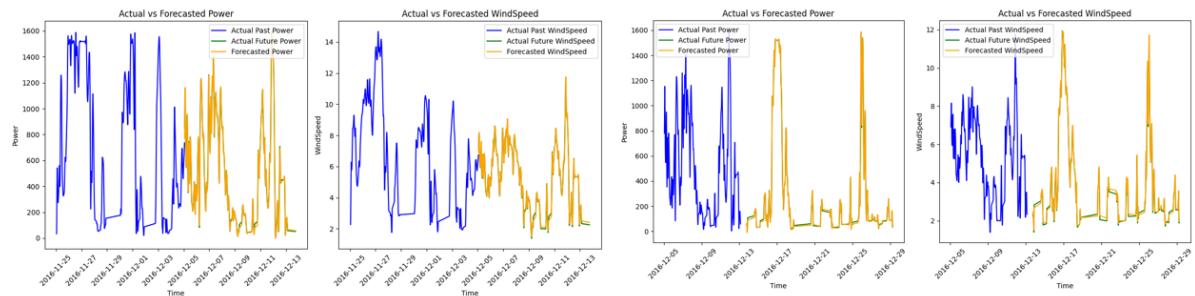
Oct:



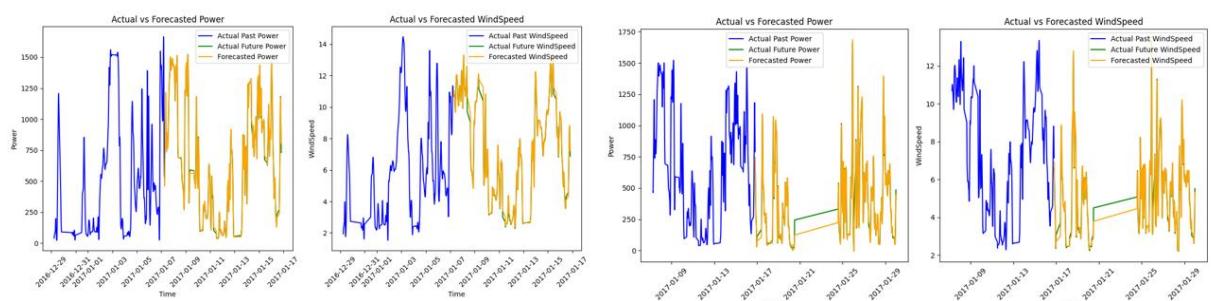
Nov:



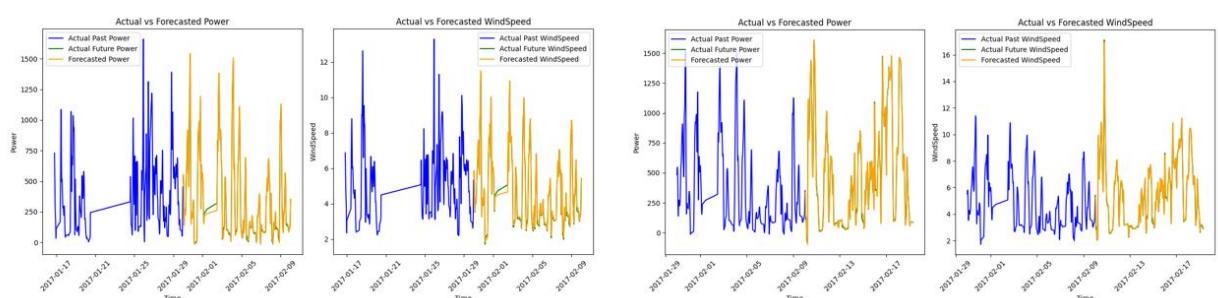
Dec:



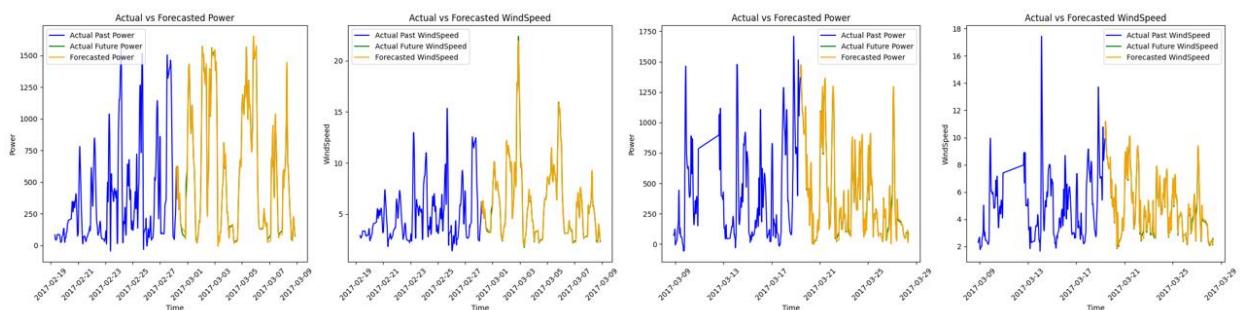
Jan:



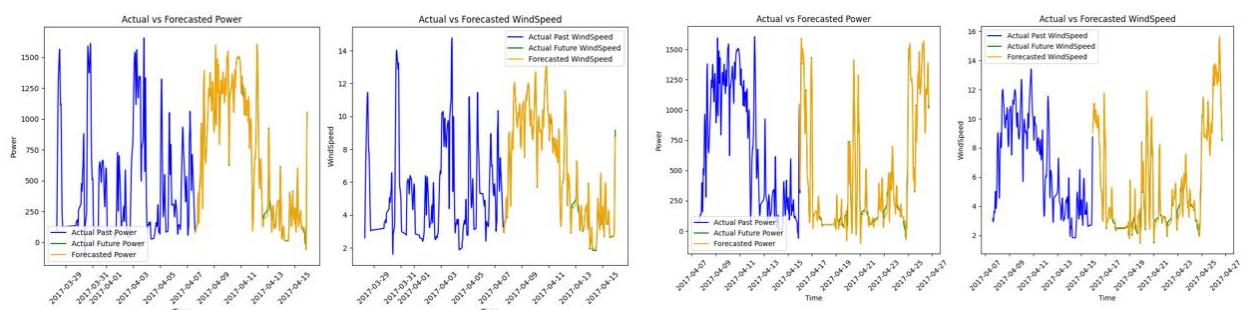
Feb:



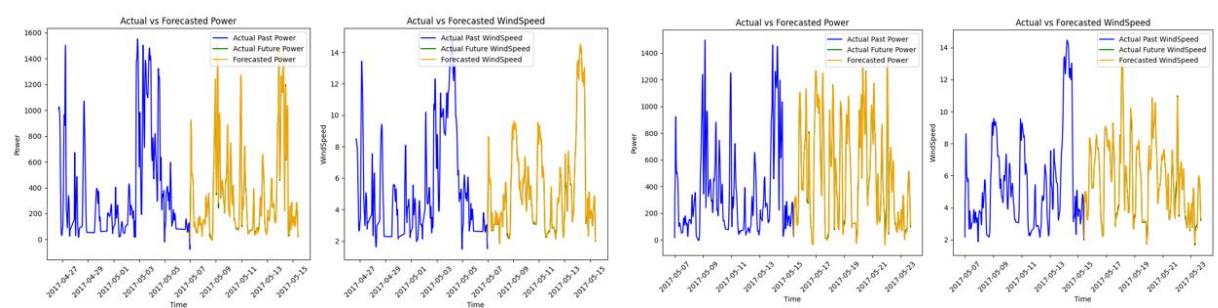
March:



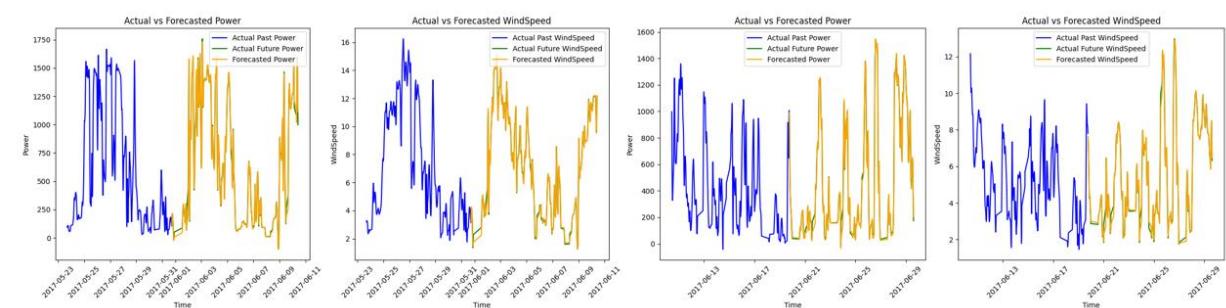
April:



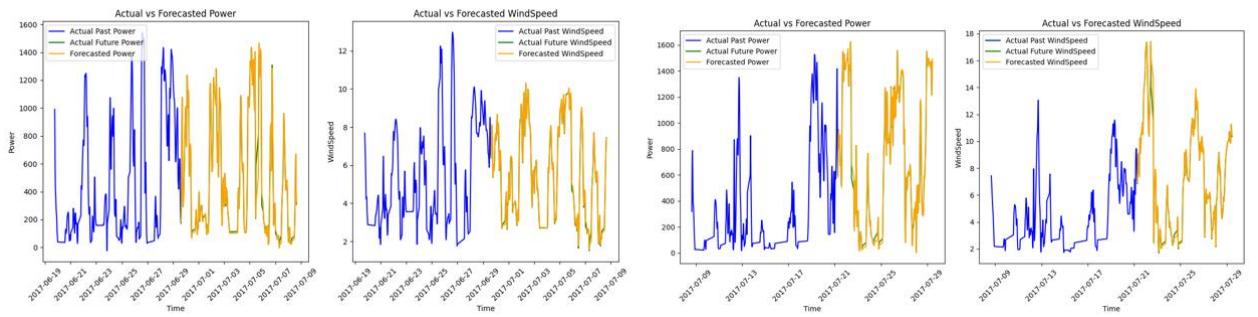
May:



June:



July:



Aug:

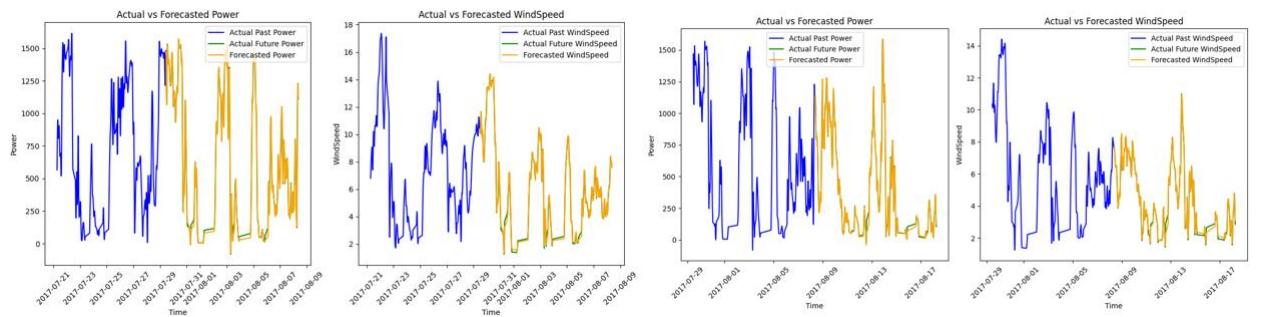


Fig 5.78: Weekly Prediction Plots using ED Model for 10Min Forecasting Window for Dataest-3

Results of metrics for 1 step ahead forecasting is shown:

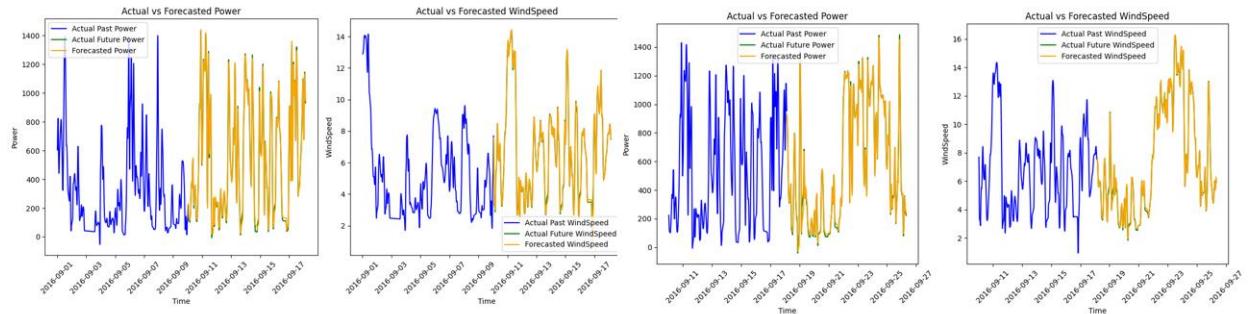
ENCODER DECODER ARCHITECTURE RESULTS for 10 min								
Duration		Power			Wind Speed			
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA	
Sept	Week 2	58.71	10.7	0.88	0.29	3.74	0.89	
	Week 4	47.76	9.36	0.87	0.27	3	0.89	
Oct	Week 1	50.08	8.9	0.9	0.29	3.83	0.87	
	Week 2	43.01	9.41	0.88	0.22	2.91	0.88	
	Week 3	41.04	13.19	0.87	0.26	4.45	0.9	
Nov	Week 1	34.57	12.33	0.87	0.23	4.41	0.89	
	Week 3	40.08	6.19	0.87	0.25	2.67	0.88	
Dec	Week 1	44.48	6.97	0.84	0.22	2.95	0.89	
	Week 3	30.63	8.97	0.87	0.19	3.67	0.88	
Jan	Week 1	46.04	6.22	0.89	0.25	2.67	0.89	
	Week 3	53.63	12.16	0.87	0.33	4.25	0.87	
Feb	Week 1	41.73	16.79	0.89	0.25	3.84	0.88	
	Week 3	44.5	9.54	0.89	0.28	3.39	0.9	
March	Week1	42.3	10.88	0.89	0.29	3.29	0.91	
	Week3	40.48	12.54	0.9	0.24	3.91	0.87	
April	Week1	52	9.72	0.91	0.35	4.15	0.87	
	Week2	43.8	7.97	0.89	0.24	3.03	0.89	
	Week3	48.08	33.21	0.88	0.31	4.35	0.87	
May	Week1	45.89	12.42	0.88	0.31	4.59	0.87	
	Week3	52.87	20.07	0.9	0.28	3.8	0.89	
June	Week1	62.15	11.43	0.86	0.29	3.4	0.88	
	Week3	41.1	11.83	0.87	0.25	3.6	0.88	
July	Week1	49.04	9.4	0.87	0.2	3.26	0.9	
	Week3	36.74	13.07	0.92	0.22	3.69	0.88	
Aug	Week1	44.35	10.75	0.9	0.21	3.18	0.9	
	Week3	32.44	9.32	0.9	0.17	3.35	0.89	

Fig 5.79: Performance metrics of ED model for 1 step ahead (10 min ahead) forecasting

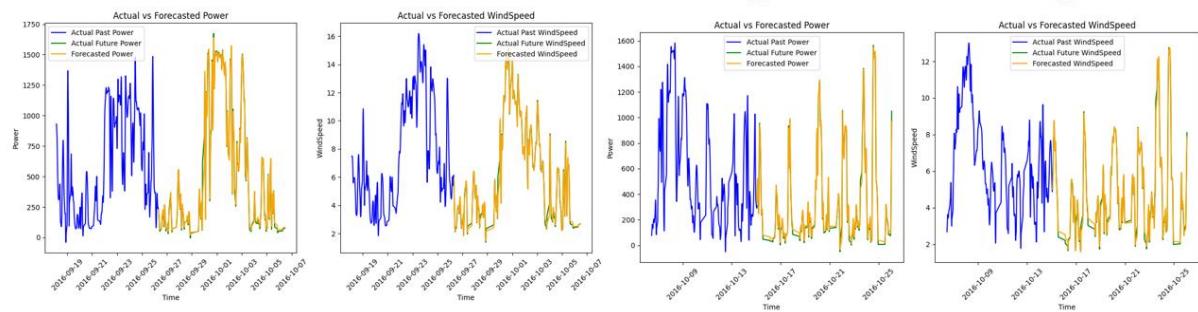
5.5.2 3 step (30 min ahead of forecasting) Plots:

Plots for each week from every month:

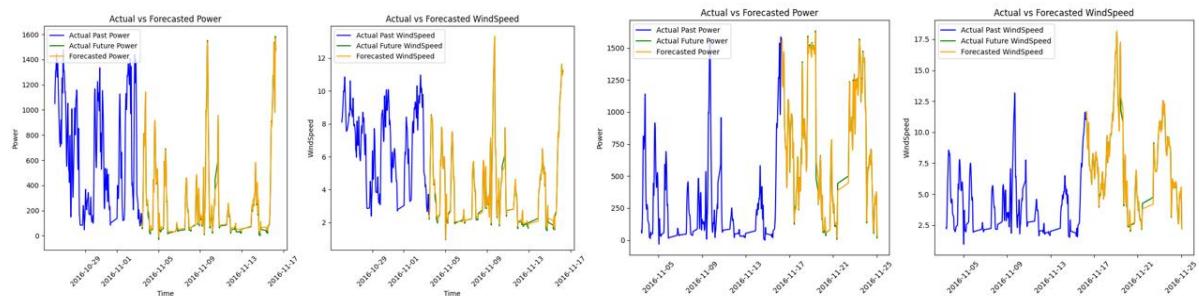
Sept:



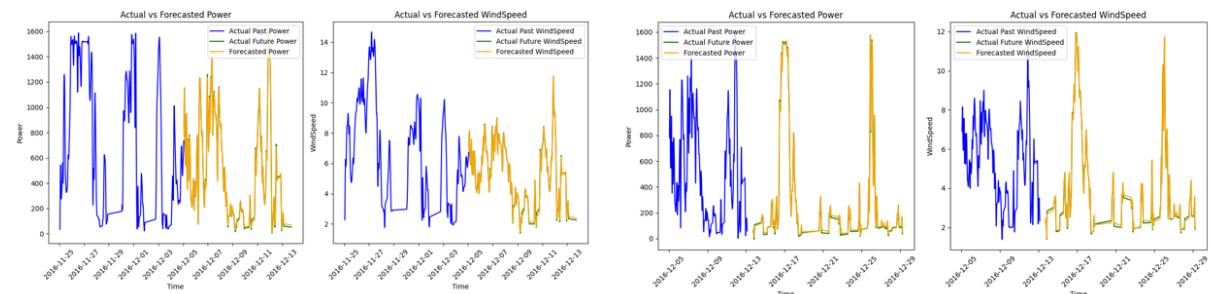
Oct:



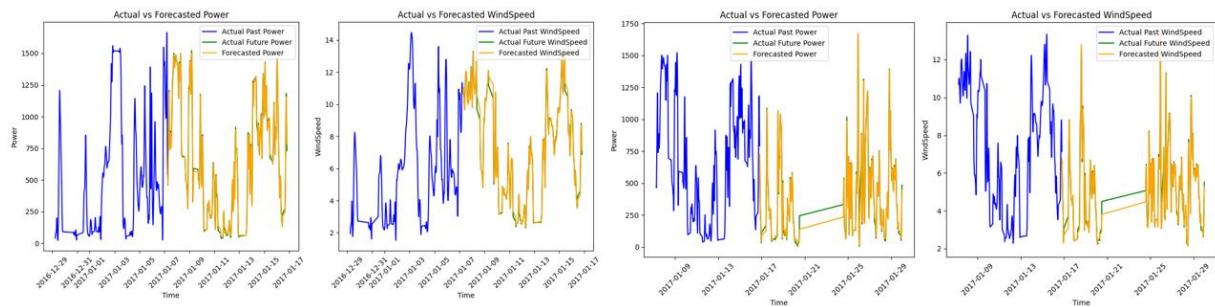
Nov:



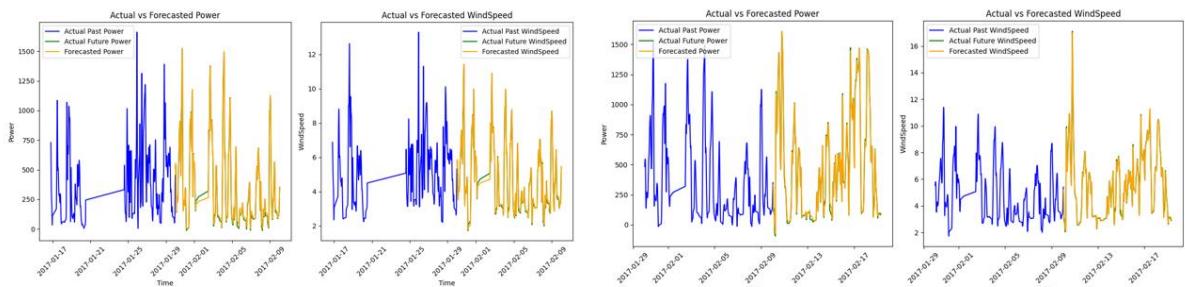
Dec:



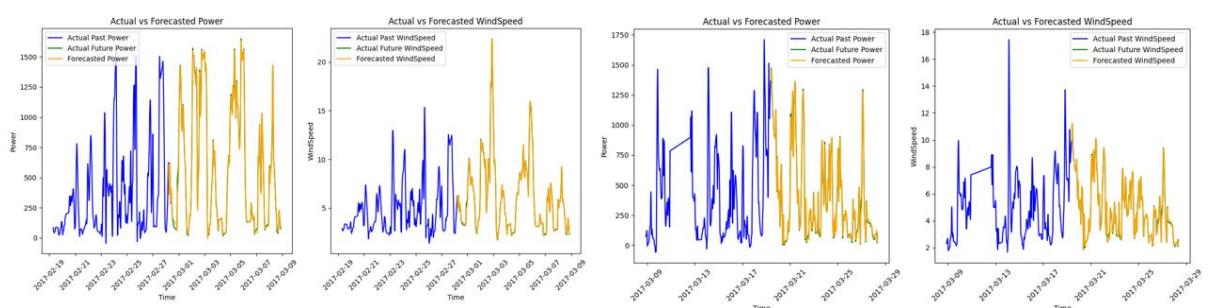
Jan:



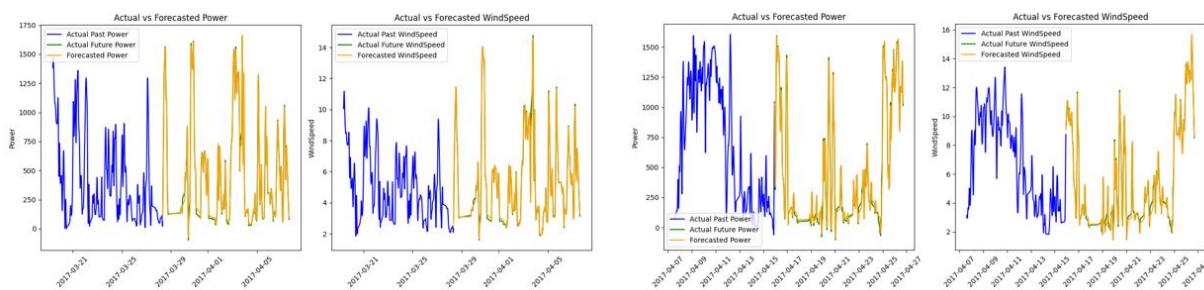
Feb:



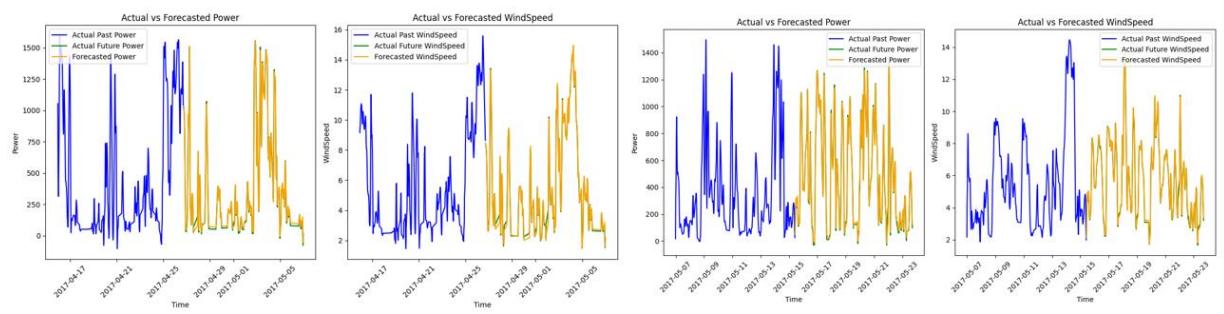
March:



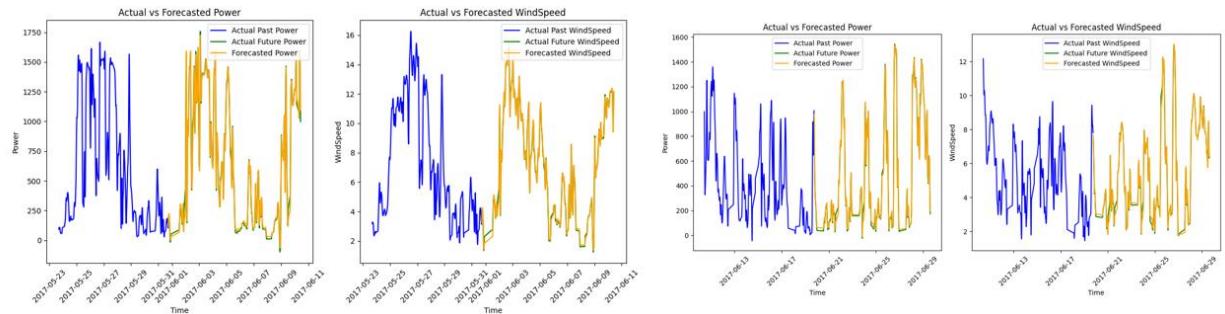
April:



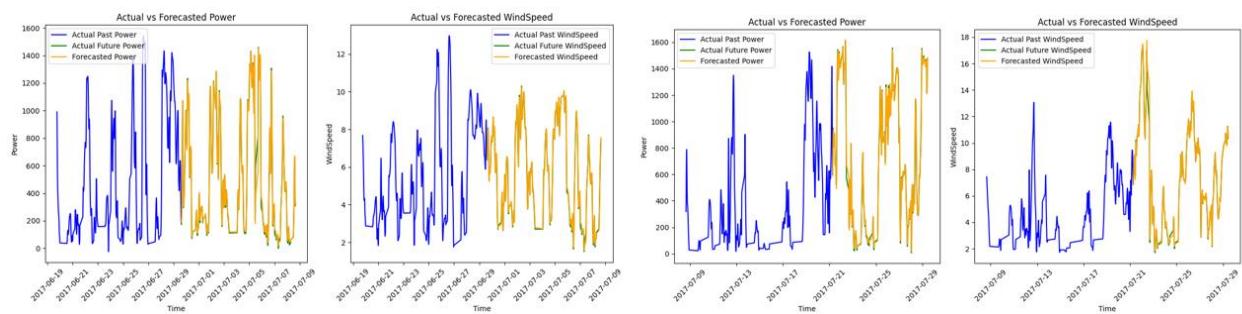
May:



June:



July:



Aug:

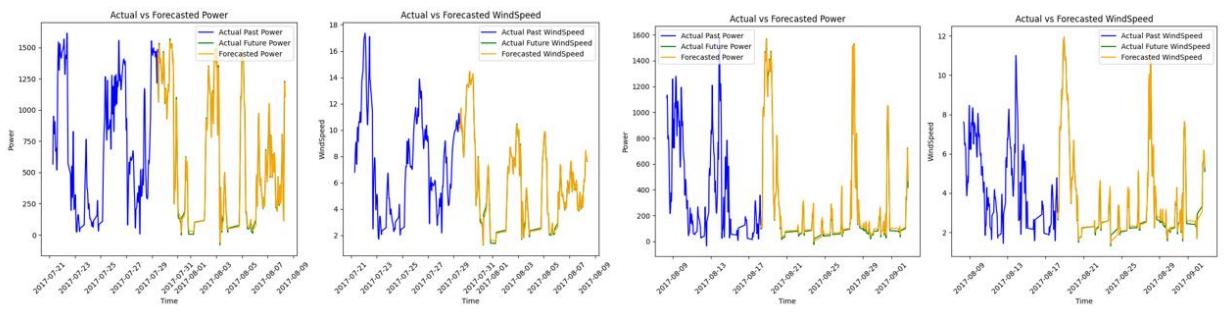


Fig 5.80: Weekly Prediction Plots using ED Model for 30Min Forecasting Window for Dataest-3

All the metric results are tabulated below:

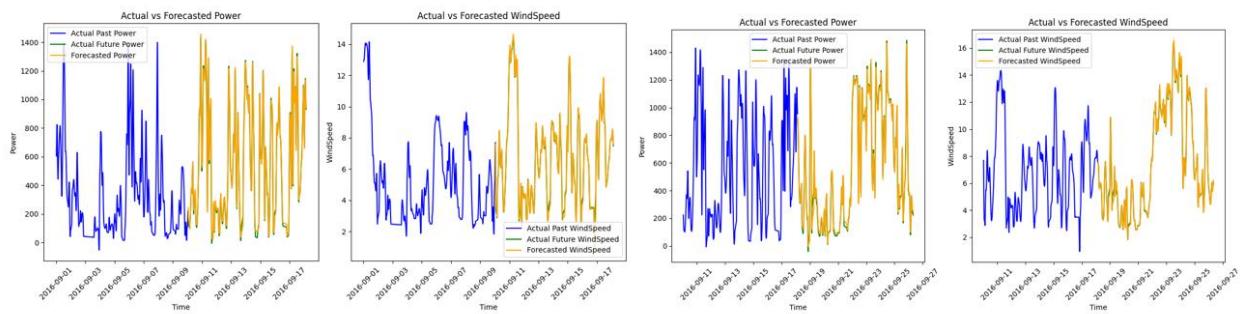
ENCODER DECODER ARCHITECTURE RESULTS for 30 min							
Duration		Power			Wind Speed		
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Sept	Week 2	55.17	11.81	0.9	0.27	3.58	0.9
	Week 4	45.21	10.8	0.89	0.25	2.86	0.91
Oct	Week 1	47.34	15.37	0.91	0.28	3.83	0.89
	Week 2	43.01	9.41	0.88	0.22	2.91	0.88
	Week 3	39.94	21.83	0.89	0.25	4.33	0.91
Nov	Week 1	34.29	22.38	0.88	0.22	4.46	0.91
	Week 3	38.34	7.74	0.87	0.24	2.6	0.89
Dec	Week 1	42.25	7.59	0.86	0.21	2.78	0.9
	Week 3	30.52	13.34	0.88	0.18	3.47	0.89
Jan	Week 1	44.05	8.97	0.86	0.23	3.14	0.9
	Week 3	50.35	14.01	0.88	0.31	3.89	0.9
Feb	Week 1	43.1	10.69	0.9	0.27	3.2	0.9
	Week 3	45.2	17.88	0.9	0.31	4.31	0.92
March	Week1	41.16	11.88	0.9	0.28	3.1	0.92
	Week3	48.61	17.56	0.9	0.31	3.86	0.89
April	Week1	49.41	10.41	0.91	0.32	3.89	0.89
	Week2	42.19	8.88	0.9	0.23	2.88	0.92
	Week3	46.63	33.37	0.9	0.29	4.1	0.89
May	Week1	44.11	15.54	0.9	0.3	4.48	0.88
	Week3	49.79	22.78	0.91	0.26	3.61	0.91
June	Week1	57.99	13.63	0.99	0.3	4.48	0.88
	Week3	49.79	22.78	0.88	0.28	3.36	0.89
July	Week1	46.55	18.35	0.89	0.2	3.19	0.9
	Week3	36.74	13.07	0.92	0.22	3.69	0.91
Aug	Week1	42.13	16.61	0.9	0.2	3.25	0.9
	Week3	31.81	14.43	0.91	0.17	3.48	0.91

Fig 5.81: Performance metrics of ED model for 3 step ahead (30 min ahead) forecasting

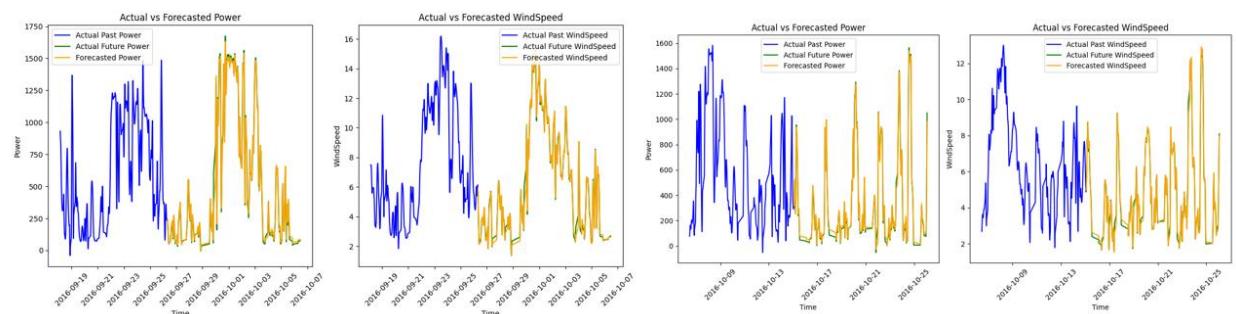
5.5.3 6 step (60 min ahead of forecasting) Plots:

Plots for each week from every month:

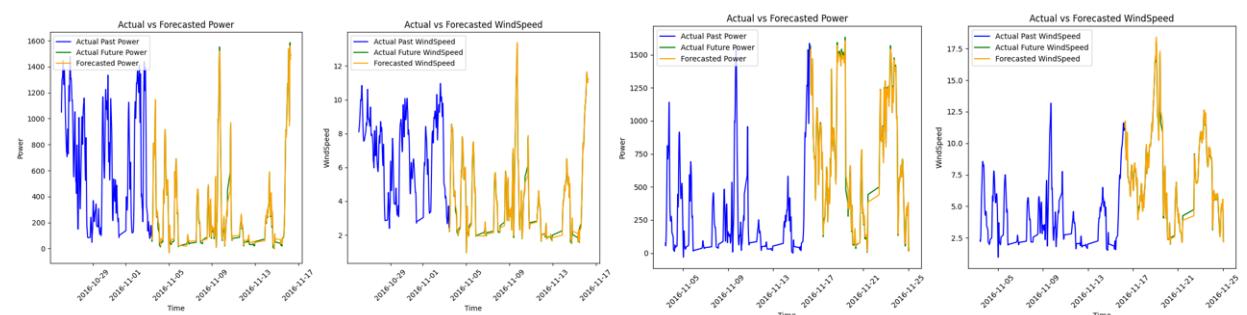
Sept:



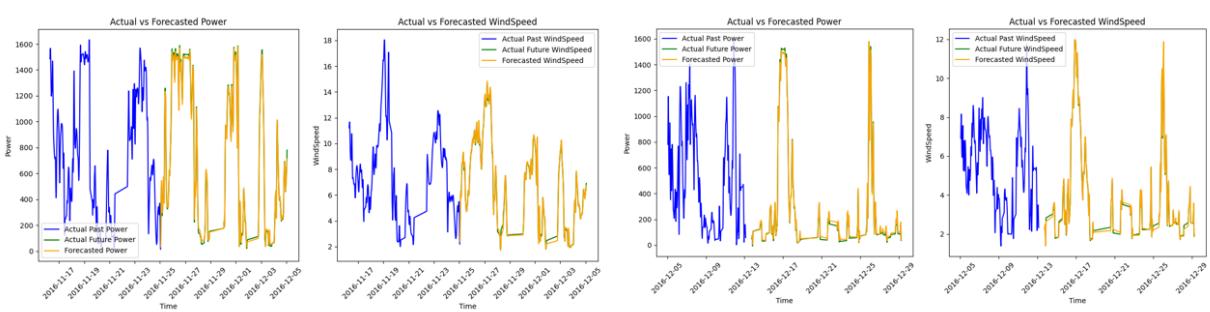
Oct:



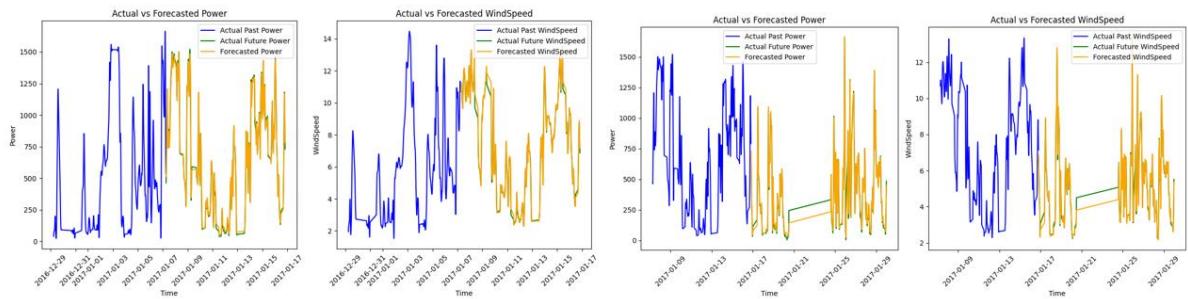
Nov:



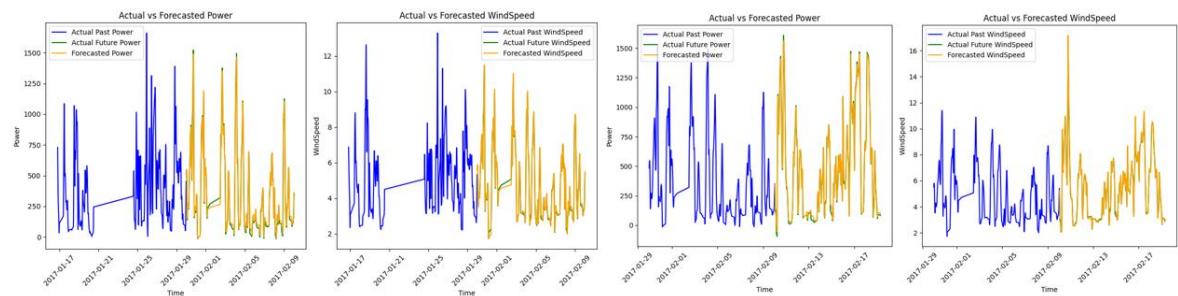
Dec:



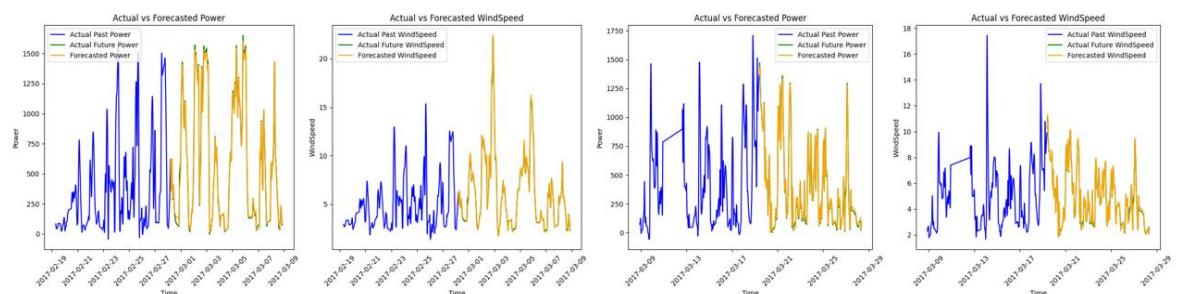
Jan:



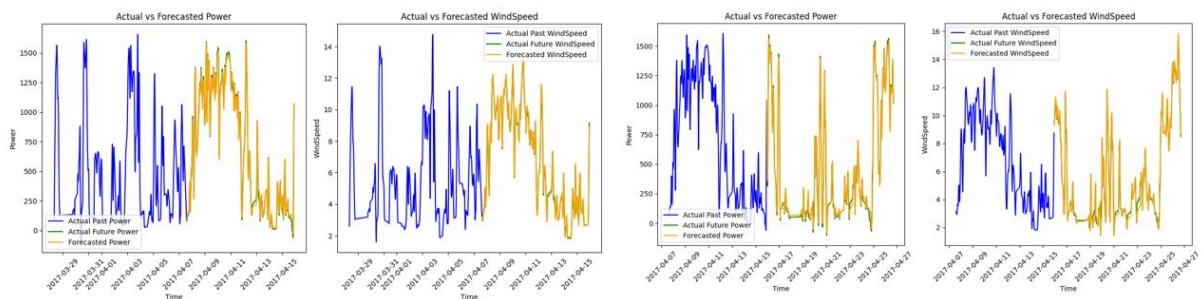
Feb:



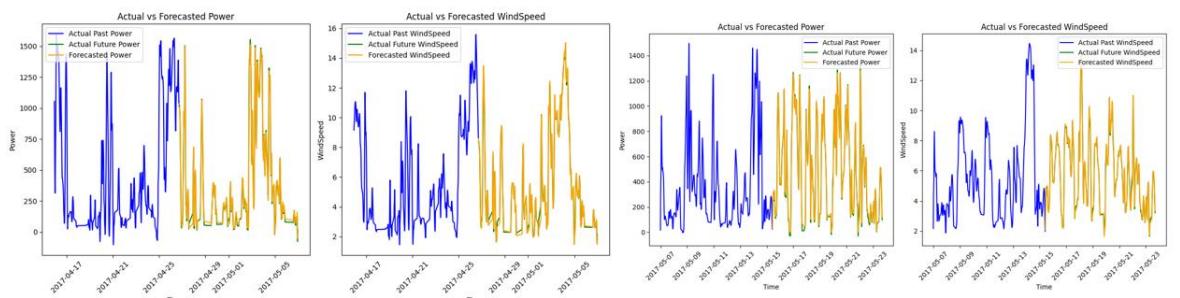
March:



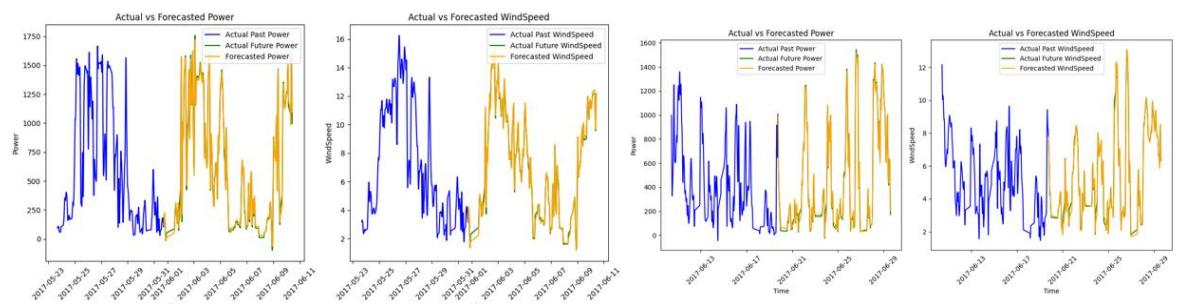
April:



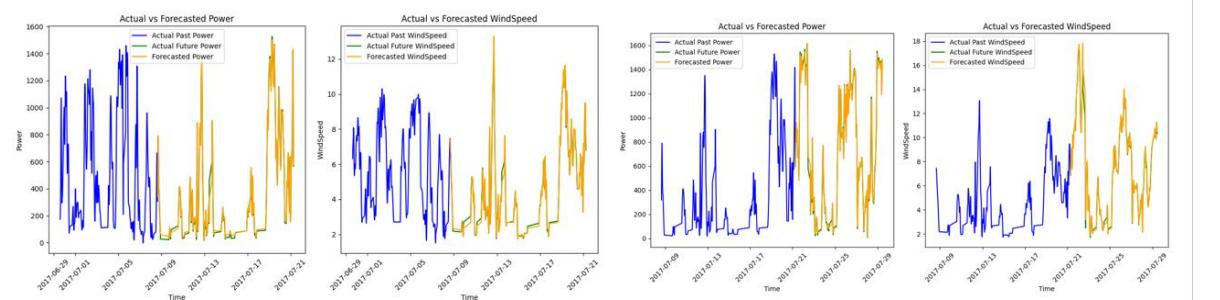
May:



June:



July:



Aug:

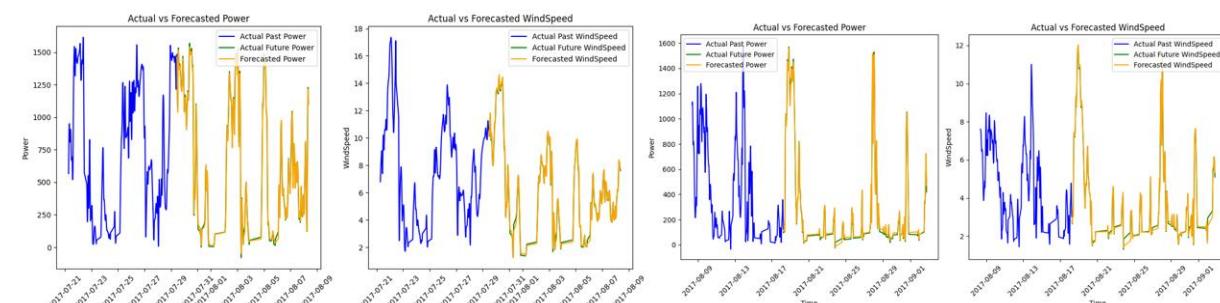


Fig 5.82: Weekly Prediction Plots using ED Model for 60Min Forecasting Window for Dataest-3

All the metric results are tabulated below:

ENCODER DECODER ARCHITECTURE RESULTS for 60 min							
Duration		Power			Wind Speed		
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Sept	Week 2	54.74	12.56	0.9	0.28	3.59	0.89
	Week 4	45.49	10.93	0.87	0.26	2.85	0.91
Oct	Week 1	49.94	13.23	0.88	0.28	3.77	0.88
	Week 2	43.72	10.14	0.87	0.23	2.9	0.89
	Week 3	40.45	19.46	0.89	0.25	4.27	0.89
Nov	Week 1	34.86	20.56	0.87	0.22	4.21	0.9
	Week 3	39.17	7.89	0.87	0.24	2.54	0.89
Dec	Week 1	44.56	8.14	0.86	0.22	2.86	0.9
	Week 3	32.88	14.66	0.87	0.18	3.53	0.89
Jan	Week 1	45.61	7.28	0.9	0.24	2.62	0.88
	Week 3	52.19	17.08	0.87	0.31	4.04	0.88
Feb	Week 1	41.24	24.14	0.89	0.24	3.68	0.89
	Week 3	44.92	11.39	0.89	0.27	3.23	0.9
March	Week1	46.27	18.76	0.9	0.32	4.45	0.9
	Week3	40.41	16.97	0.89	0.23	3.86	0.89
April	Week1	51.27	11.42	0.9	0.33	4.03	0.89
	Week2	44.96	9.41	0.89	0.23	2.95	0.91
	Week3	48.72	38.25	0.89	0.31	4.27	0.89
May	Week1	45.58	16.69	0.88	0.3	4.49	0.88
	Week3	48.97	23.26	0.91	0.26	3.58	0.91
June	Week1	59.58	14.18	0.88	0.28	3.31	0.89
	Week3	39.42	14.49	0.89	0.24	3.96	0.89
July	Week1	46.17	18.78	0.89	0.2	3.2	0.9
	Week3	40.5	8.7	0.86	0.26	2.78	0.91
Aug	Week1	43.08	15.62	0.9	0.21	3.19	0.91
	Week3	32.8	14.11	0.91	0.17	3.26	0.91

Fig 5.83: Performance metrics of ED model for 6 step ahead (60 min ahead) forecasting

5.6 Calculation of Potential Wind Power and Average power loss:

As dictated in equation 5.3, potential power at each timestep will be calculated. Lets summarize the important parameters used for Kuwait turbine dataset:

Here are technical details of wind park is tabulated

Characteristic	Details
Model Name	Siemens-Gamesa G97
Rated power	2 MW
Hub height	78.98 m
Rotor diameter (D)	97 m
Swept area (A)	7,390 m ²
Cut-in wind speed	3 m/s
Rated wind speed	11 m/s
Cut-off wind speed	25 m/s

Table 5.6: Technical characteristics of Wind turbine for dataset-3

Methodology to capture loss in summer and winter region:

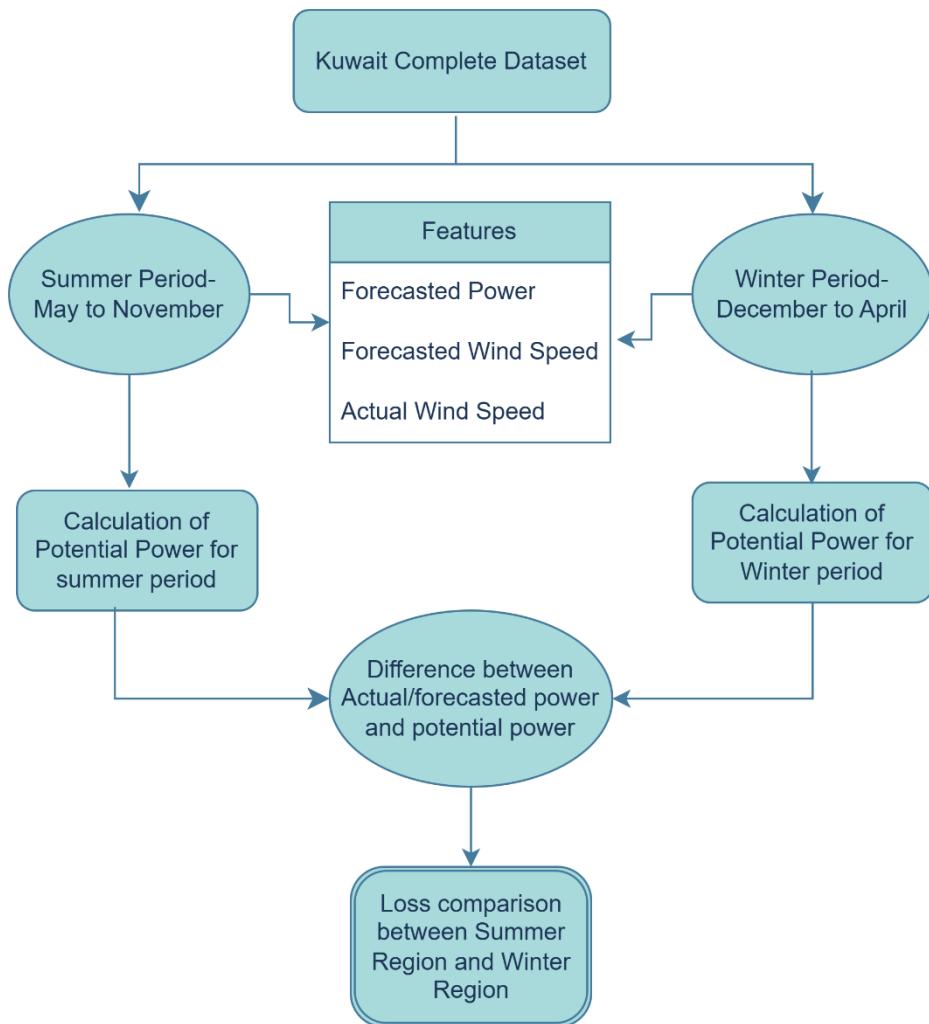


Fig 5.84 : Methodology to understand loss in summer and winter region

Data has been collected for summer period and winter period respectively and potential power is calculated.

Loss between potential power generation and one step ahead of forecasted power

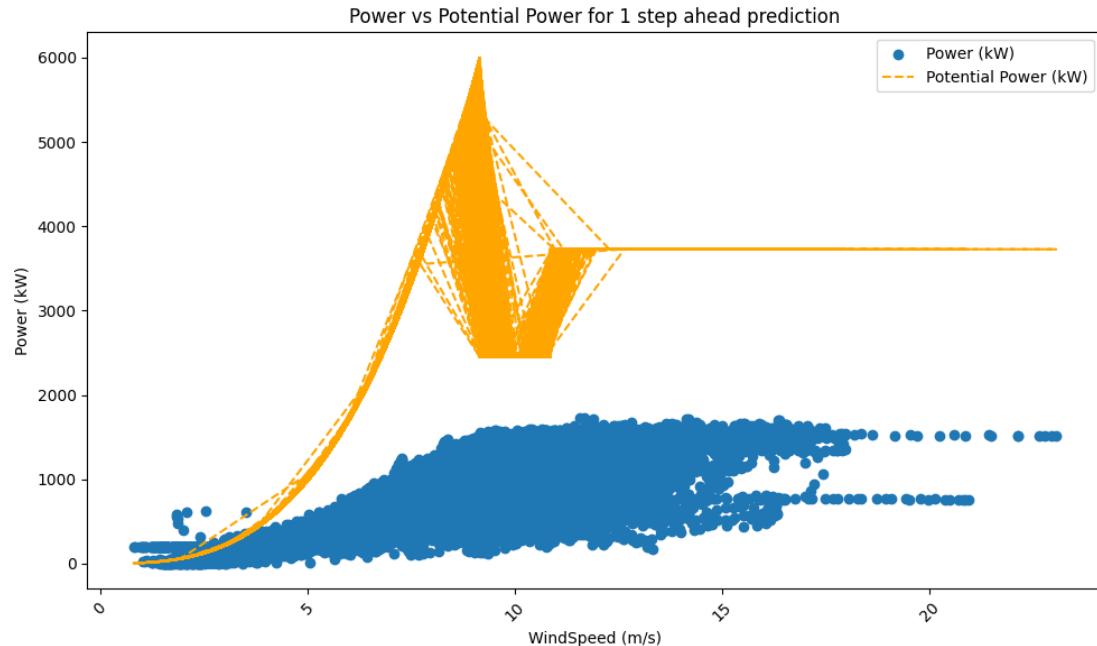


Fig 5.85: Potential Power Vs Forecasted power generated in summer period for 1 step ahead forecasting

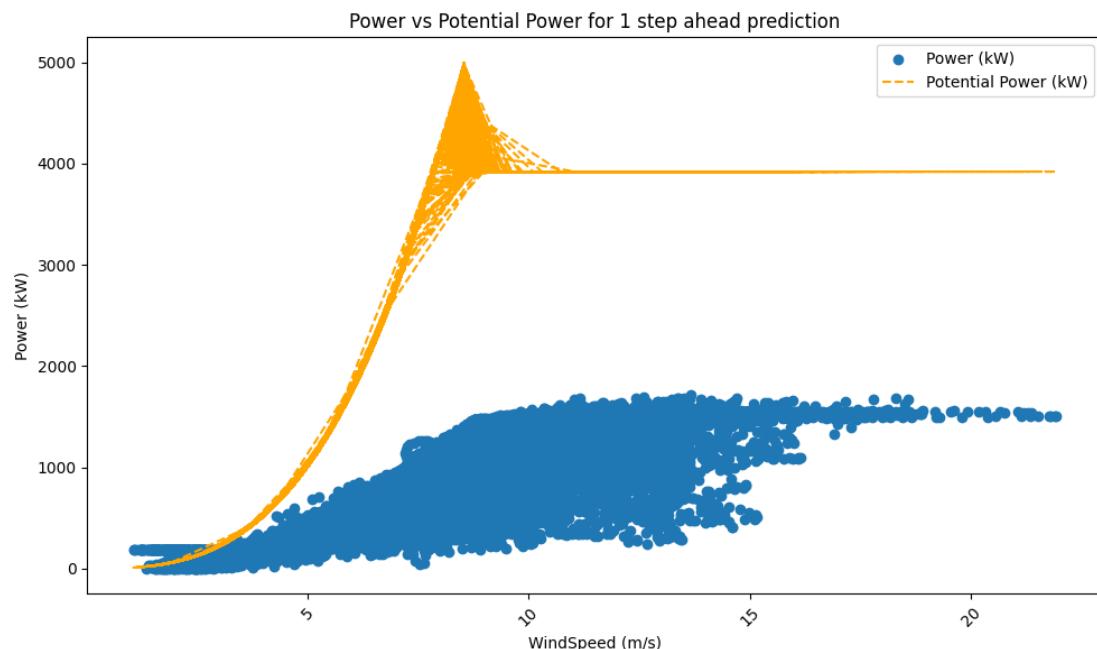


Fig 5.86: Potential Power Vs Forecasted power generated in Winter period for 1 step ahead forecasting

Loss between potential power generation and three step ahead of forecasted power:

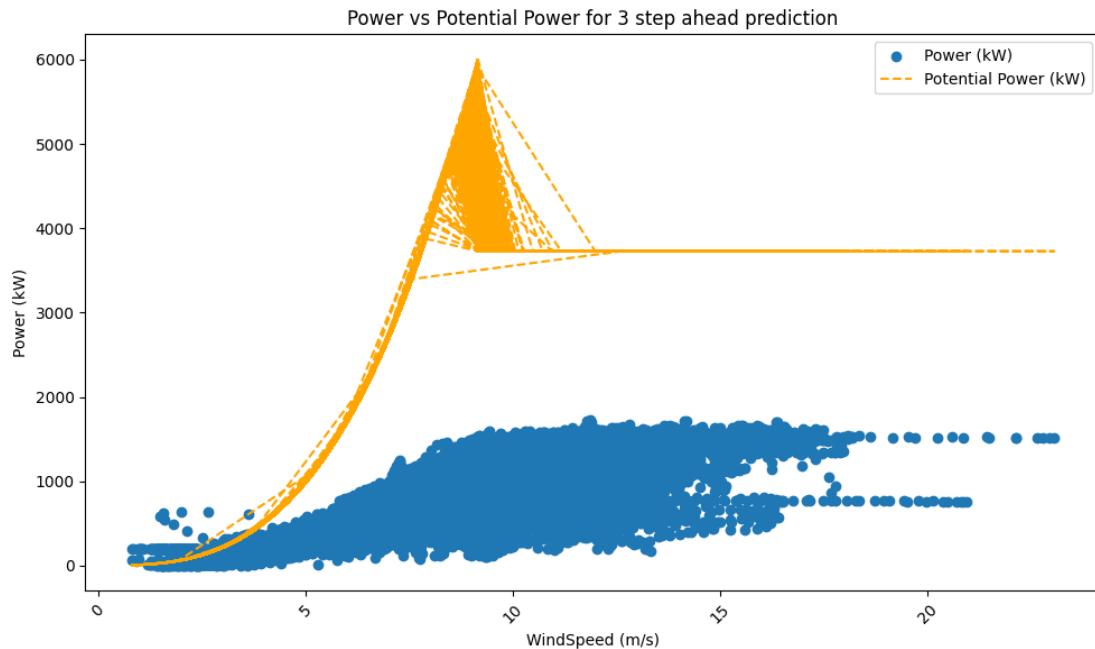


Fig 5.87: Potential Power Vs Forecasted power generated in Summer period for 3 step ahead forecasting

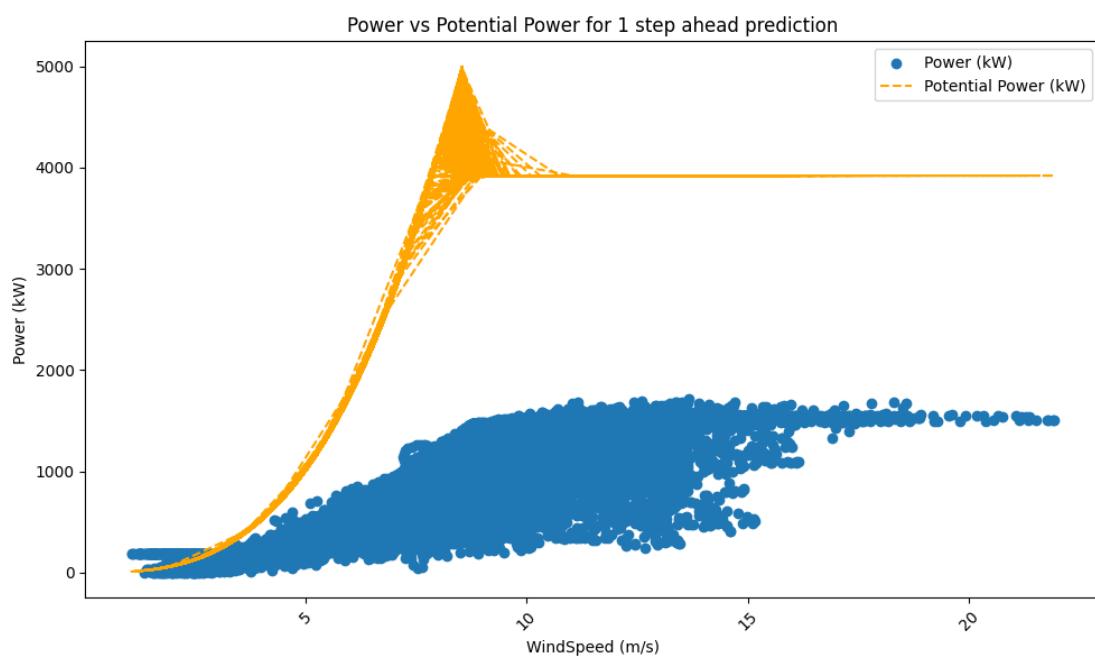


Fig 5.88: Potential Power Vs Forecasted power generated in Winter period for 3 step ahead forecasting

Loss between potential power generation and six step ahead of forecasted power:

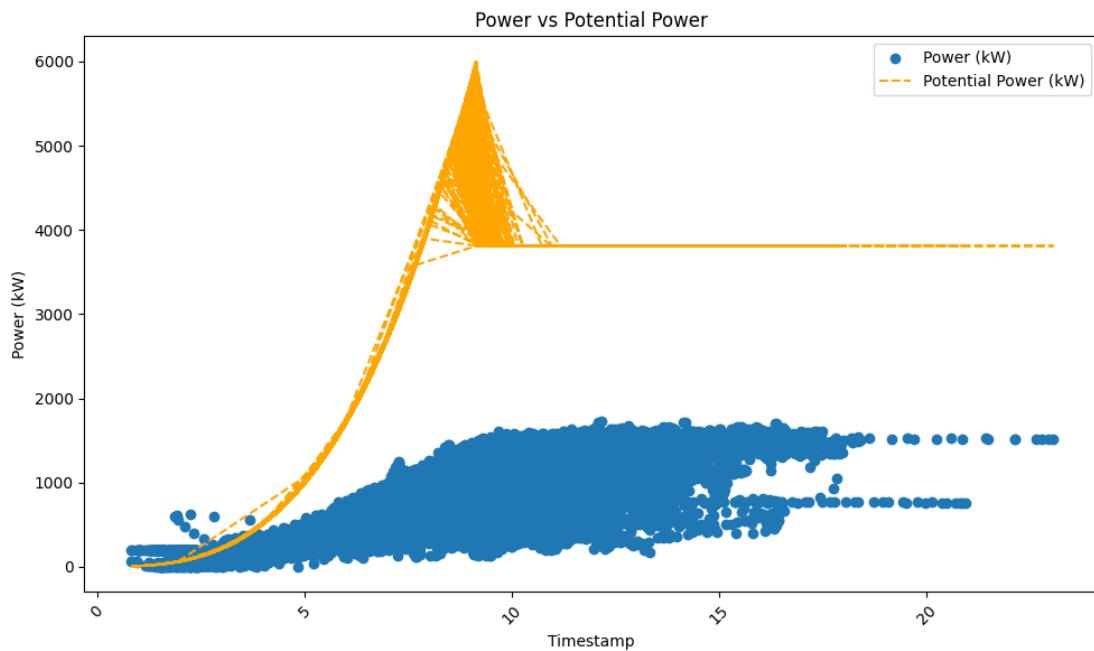


Fig 5.89: Potential Power Vs Forecasted power generated in Summer period for 6 step ahead forecasting

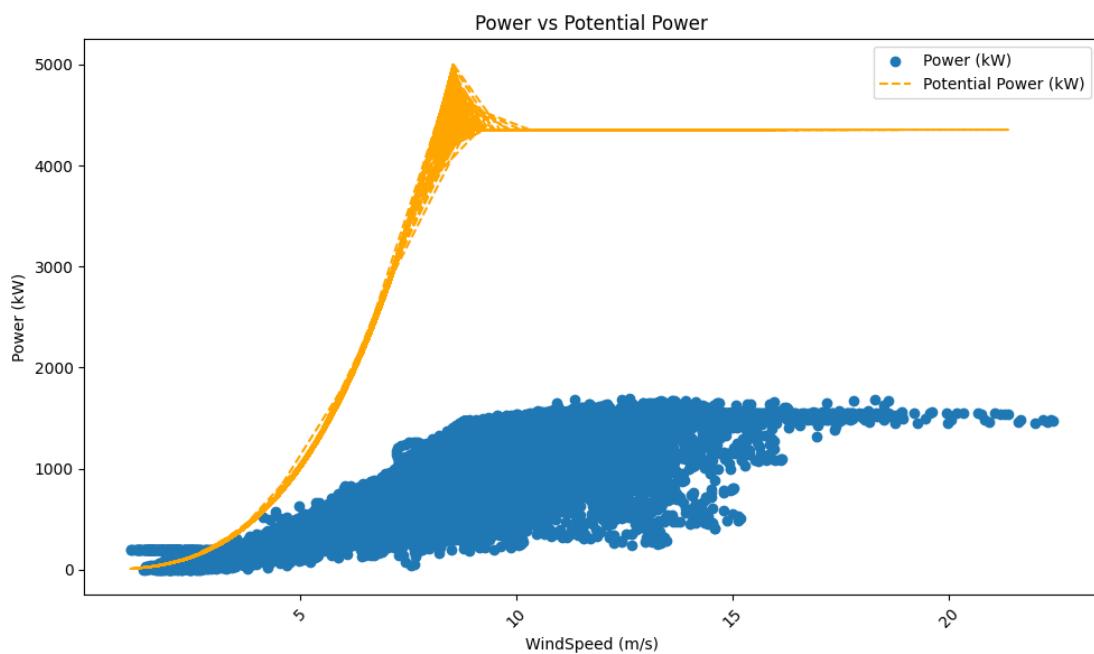


Fig 5.90: Potential Power Vs Forecasted power generated in Winter period for 6 step ahead forecasting

DATASET-3 AVERAGE LOSS			
Forecasting steps	Summer Period loss	Winter Period loss	% Loss difference
1 step	64.62	62.77	1.85
3 step	64.62	63.58	1.04
6 step	66.29	65.78	0.51

Table 5.7: Average loss calculations for dataset-3

This power curve plot shows the relationship between wind speed (x-axis) and power output (y-axis) for a wind turbine. Here are some observations:

Power Output vs. Wind Speed Relationship:

At low wind speeds the power output is minimal, as expected, because turbines need a certain threshold of wind speed (known as the cut-in speed) to start generating power.

As wind speed increases, power output rises sharply and reaches its peak. This upward slope corresponds to the typical increase in power with wind speed until it reaches the rated wind speed.

Potential Power Curve:

The orange line represents the "Potential Power" of the turbine, likely showing the theoretical maximum power the turbine can generate at each wind speed. This curve rises steeply until it plateaus, indicating the rated power output.

After reaching the rated speed, the potential power curve levels off, indicating that the turbine is designed to cap power output beyond a certain wind speed to avoid mechanical or safety risks.

Real Power Output vs. Potential Power:

The blue scatter points represent the actual power data, which mostly lie below the potential power curve, showing that real-world performance often falls below the theoretical maximum.

The deviation between actual and potential power could be due to factors like wind turbulence, suboptimal operational conditions, or turbine limitations.

Noise and Outliers:

There are some points where the actual power seems to fall off sharply at relatively higher wind speeds, which may indicate operational stops or limitations such as cut-out speeds where the turbine shuts down to prevent damage.

5.6 Cross Dataset Experiments:

In phase 3, two datasets have been explored which are mostly affected by temperature change.

Cold region dataset (Ireland) and Hot region dataset (Kuwait) are trained separately and their trained weights are used to check performance of other dataset, i.e, trained weights of cold region data is used to check performance of hot region test data.

Here are the results for both the dataset for all steps of time intervals:

CROSS DATA EXPERIMENT ON HOT CLIMATE DATASET USING WEIGHTS OF COLD CLIMATE DATASET																									
Duration		10 Min Results						30 Min Results						60 Min Results											
		Power			Wind Speed			Power			Wind Speed			Power			Wind Speed								
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Sept	Week 2	202.39	39.99	0.71	1.25	15.4	0.69	233.15	48.21	0.65	1.47	18.54	0.66	302.22	70.46	0.58	1.89	25.06	0.6						
	Week 4	165.32	33.82	0.7	1.19	12.93	0.71	208.24	45.2	0.68	1.44	16.39	0.67	297.83	67.82	0.58	2	23.38	0.6						
Oct	Week 1	197.6	51.76	0.7	1.26	14.22	0.67	231.18	59.17	0.65	1.48	18.3	0.62	306.42	92.29	0.56	1.92	26.01	0.57						
	Week 2	162.36	34.13	0.7	1.05	13.06	0.69	203.6	42.76	0.66	1.32	16.65	0.61	289.61	60.11	0.59	1.84	23.77	0.54						
Nov	Week 1	140.84	78.17	0.7	0.92	16.17	0.7	180.67	98.49	0.66	1.2	21.64	0.66	271.89	70.23	0.59	1.19	33.07	0.62						
	Week 3	140.33	27.25	0.69	0.96	11.34	0.69	174.82	38.19	0.65	1.21	14.85	0.65	246.35	58.38	0.58	1.52	32.06	0.58						
Dec	Week 1	154.49	30.7	0.71	0.95	12.25	0.69	191.89	40.21	0.66	1.18	15.9	0.65	267.41	56.99	0.58	1.67	22.96	0.58						
	Week 3	109.66	43.11	0.71	0.69	12.8	0.72	28.14	61	0.65	0.88	16.91	0.67	188.53	85.6	0.59	1.28	23.23	0.59						
Jan	Week 1	156.52	31.32	0.7	0.89	11.87	0.72	200.85	34.52	0.59	1.18	13.51	0.64	216.95	58.79	0.61	1.34	20.82	0.63						
	Week 3	198.78	60.54	0.69	1.35	18.76	0.69	252.34	78.26	0.62	1.7	23.96	0.64	352.8	99.43	0.55	2.33	33.43	0.56						
Feb	Week 1	139.57	63.77	0.71	0.92	14.08	0.68	178.49	97.23	0.66	1.2	18.91	0.64	239.2	136.82	0.6	1.63	26.6	0.57						
	Week 3	157.18	70.74	0.72	1.21	17.24	0.73	190.61	43.72	0.7	1.33	17.46	0.67	251.86	62.27	0.62	1.76	23.78	0.6						
March	Week1	136.54	36.93	0.71	1.08	11.71	0.72	165.78	41.39	0.67	1.26	14.82	0.67	217.43	71.45	0.6	1.65	20.68	0.59						
	Week3	144.16	55.14	0.7	0.96	15.63	0.71	179.86	85.84	0.64	1.21	20.26	0.66	241.47	85.84	0.58	1.62	28	0.56						
April	Week1	185.21	38.55	0.73	1.33	15.97	0.72	229.75	50.63	0.68	1.66	20.85	0.69	305.41	73.03	0.66	2.19	29.04	0.63						
	Week2	149.83	29.36	0.7	0.96	12.05	0.75	179.19	39.98	0.66	1.19	15.6	0.72	244.13	60.5	0.58	1.66	22.13	0.65						
May	Week1	159.6	98.31	0.69	1.15	17.02	0.69	190.92	112	0.65	1.4	21.33	0.65	246.29	146.76	0.58	1.79	27.88	0.59						
	Week3	182.45	41.77	0.65	1.1	15.87	0.63	210.66	72.63	0.68	1.43	19.79	0.69	295.44	116.24	0.62	2.01	29.06	0.62						
June	Week1	171.35	58.87	0.73	1.14	15.55	0.73	271.48	50.74	0.63	1.59	17.16	0.61	356.41	73.49	0.57	1.99	23.28	0.6						
	Week3	238.39	40.88	0.66	1.35	14.17	0.62	180.48	70.14	0.65	1.15	18.45	0.67	272.94	108.39	0.58	1.73	29.32	0.62						
July	Week1	173.59	70.27	0.72	0.93	14.25	0.72	206.29	96.06	0.67	1.17	19.11	0.65	282.3	132.9	0.6	1.69	27.69	0.57						
	Week3	156.96	31.83	0.62	1.03	11.75	0.69	177.8	55.88	0.62	1.21	19.28	0.62	245.27	83.75	0.57	1.66	27.57	0.56						
Aug	Week1	152.99	43.85	0.71	0.75	13.3	0.7	192.27	62.12	0.67	1.2	17.07	0.66	286.15	106.74	0.58	1.77	26.56	0.58						
	Week3	110.29	38.1	0.71	0.69	12.33	0.71	146.78	53	0.66	0.92	17.44	0.65	242.93	90.86	0.62	1.56	29.51	0.62						

CROSS DATA EXPERIMENT ON COLD CLIMATE DATASET USING WEIGHTS OF HOT CLIMATE DATASET																									
Duration		10 Min Results						30 Min Results						60 Min Results											
		Power			Wind Speed			Power			Wind Speed			Power			Wind Speed								
Month	Week	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
April	Week 2	1.2	0.89	1	0.04	0.55	1	5.99	5.26	1	0.06	0.94	1	5.81	4.83	1	0.01	0.19	1						
	Week 4	7.5	0.86	1	0.1	0.57	1	8.89	1.02	1	0.05	0.32	1	17.01	1.95	0.99	0.04	0.29	0.99						
May	Week 1	0.24	0.1	1	0.01	0.06	0.99	0.71	0.27	1	0.01	0.1	1	0.86	0.33	1	0.01	0.07	0.99						
	Week 2	0.22	0.1	1	0.01	0.13	0.99	0.67	0.3	1	0.01	0.13	0.99	0.89	0.39	1	0.01	0.02	0.99						
June	Week 1	0.73	0.57	0.99	0.02	0.39	0.99	4.82	4.07	1	0.05	0.78	0.98	5	4.26	1	0.01	0.17	0.99						
	Week 3	0.45	0.26	1	0.02	0.21	0.99	3.41	2.16	1	0.03	0.46	0.99	3.48	2.24	1	0.01	0.13	1						
July	Week 1	0.1	0.03	0.99	0.001	0.05	1	0.6	0.24	0.99	0.01	0.09	1	0.71	0.27	1	0.01	0.09	1						
	Week 3	0.22	0.08	0.99	0.01	0.13	1	1.59	0.82	0.99	0.01	0.19	1	2.1	1.11	0.99	0.02	0.06	1						
Aug	Week 1	0.34	0.2	1	0.01	0.15	1	1.77	1.04	1	0.02	0.26	1	2.32	1.37	1	0.01	0.07	1						
	Week 3	0.28	0.13	1	0.001	0.03	1	1.68	0.76	1	0.02	0.21	1	2.05	0.94	0.99	0.01	0.09	1						
Sept	Week 3 & 4	0.55	0.26	1	0.01	0.15	1	5.6	4.6	0.99	0.05	0.61	1	2.69	1.29	0.99	0.01	0.06	1						
	Week 5	2.08	2.02	1	0.04	0.63	1	7.28	7.45	1	0.07	1.06	1	2.27	1.02	1	0.01	0.09	1						
Oct	Week 2	0.68	0.24	1	0.01	0.15	1	1.74	0.79	1	0.02	0.2	1	3.35	1.9	1	0.01	0.12	1						
	Week 4	0.75	0.18	1	0.01	0.05	0.99	1.69	0.4	1	0.01	0.11	1	5.59	1.15	0.99	0.09	0.95	0.99						
Nov	Week1	0.54	0.31	1	0.02	0.29	0.99	2.17	1.29	1	0.02	0.24	0.99	2.92	1.79	0.99	0.01	0.1	1						
	Week3	1.01	0.22	0.99	0.02	0.16	1</																		

Average Performance Results for Cross dataset experiments:

Feature Name	COLD WEATHER DATA AVG ACCURACY RESULTS									CROSS DATASET RESULTS USING HOT REGIONAL DATA WEIGHTS								
	Average Performance of 1 step ahead forecasting			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting			Average Performance of 1 step ahead forecasting			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting		
	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Wind Speed	0.008	0.066	0.99	0.01	0.09	0.99	0.05	0.5	0.99	0.02	0.205	0.99	0.025	0.31	0.99	0.016	0.15	0.99
Power Generation	0.441	0.84	0.99	0.52	0.18	0.99	1.97	0.55	0.99	0.91	0.34	0.99	2.78	1.55	0.99	3.49	1.41	0.99

Feature Name	HOT WEATHER DATA AVG ACCURACY RESULTS									CROSS DATASET RESULTS USING COLD REGIONAL DATA WEIGHTS								
	Average Performance of 1 step ahead forecasting			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting			Average Performance of 1 step ahead forecasting			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting		
	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Wind Speed	0.25	3.59	0.88	0.25	3.56	0.89	0.25	3.51	0.89	1.04	14.35	0.69	1.28	18.38	0.65	1.74	26.42	0.59
Power Generation	44.9	11.67	0.88	44.06	15.04	0.99	44.51	15.31	0.88	160.44	48.86	0.69	190.7	63.27	0.65	267.4	87.42	0.59

Fig 5.92: Average accuracy results for cross data experiments

Analysis:

1. For cold regions, the temperature ranged from -5°C to 20°C and for hot climatic regions, the temperature ranged from 28°C to 48°C.
2. The model is trained by considering the past observations of temperature as one of the input features for the specific windmill location.
3. Due to extreme conditions in cross-dataset experiments, accuracy degrades. As model training is based on a particular environment's past observations, these results are as expected.
4. Here are the comments that need to be enlisted.

Comments:

5. The thesis provides the solution for a particular windmill which comes under certain environmental conditions.
6. As, here, power generation depends on ambient temperature change, cross dataset accuracy degrades.
7. Though model architecture is unique, methodology and training performed vary as per environmental conditions and windmill locations.
8. Results above highlight the importance of windmill's location-wise training.
9. Customer oriented methodology and approach is built as per wind mill location.

Justifications on Performance of Model Results:

1. Initially, traditional models have been explored and due to seasonality in the data, SARIMAX model have been applied. Multiple inputs are needed to feed to the model for training, and thus, SARIMAX is preferred over SARIMA.
2. Average RMSE observed for SARIMAX is above 200 and MAPE is greater than 200%. Results obtained from traditional model were not appropriate.
3. To focus more on long term memory, stacked LSTM model architecture is used. For single step and single variable input feature, average RMSE was in between 150-200 but accuracy degrades after going through multi step and multi variable input features.
4. To overcome this, sequential model architectures have been explored and implemented with noise reduction techniques. Wavelet transform was found to be the best noise reduction technique.
5. With the defined methodology, when ED+WT model is trained, significant amount of improvement over accuracy is observed.
6. Further this model is used to check the performance on extreme cold and extreme hot regions.
7. Below is the average comparative result analysis done for all models on dataset-1.

Models	RMSE	MAPE	MDA(%)
SARIMAX	296.1	422.42	83.3
Stacked LSTM	532.9	738.47	47.76
Encoder Decoder	192.5	117.35	49.76
Encoder Decoder + Wavelet Transform	2.38	0.29	99

Table 5.8: Comparative Study of all Models

As per the results Encoder-Decoder architecture along with wavelet transform as denoising method is found to be the best model to go with.

CHAPTER-6

CONCLUSION AND FUTURE WORK

Dataset-1 Conclusion:

The study has been done and implemented successfully in Part A and Part B of the methodology. Forecasting results have been generated for 10 min, 30 min, and 60 min ahead of the forecasting window.

The results are discussed, and it is concluded that,

- Due to non linear relation between features SARIMA model results are not accurate.
- The stacked LSTM architecture is suitable only for univariate and single-step-ahead forecasting windows.
- For multistep and multifeatured forecasting, the Encoder Decoder architecture provides around 79% improvement.
- Further, it has been studied that data is noisy and need to be denoised. Various denoising techniques like exponential smoothing, fft, and wavelet transform are implemented, and it is found that wavelet transform gives more than 90% accuracy than any other denoising technique.
- The avg improvement at each step is tabulated below:

	Feature Name	% AVG Improvement For 10 Min			% AVG Improvement For 30 Min			% AVG Improvement For 60 Min		
	Metric	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
ACTIVEPOWER	SE	82.21	91.86	51.3	80.13	92.17	41.96	68.19	87.12	36.25
	DE	38.45	27.49	24.22	41.35	63.71	23.68	42.94	10.14	24.37
	FFT	99.2	99.51	55.97	96.91	98.31	46.9	99.12	99.56	27.62
	WT	99.41	99.68	108.68	99.41	99.62	108.42	98.9	99.43	107.16

	Feature Name	% AVG Improvement For 10 Min			% AVG Improvement For 30 Min			% AVG Improvement For 60 Min		
	Metric	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
WIND SPEED	SE	82.37	82.35	44.97	81.52	82.29	38.69	71.64	72.27	32.97
	DE	78.11	78.58	57.63	82.42	82.73	53.98	76.66	76.31	47.7
	FFT	77.1	75.95	86.48	77.78	77.81	87.6	24.72	25.95	27.49
	WT	99.46	99.41	107.38	99.14	98.97	108.84	98.95	99.02	105.5

Fig 6.1: Average Improvement Over ED architecture using Smoothing Techniques

From the tabular results, it is noted that wavelet transform, as a denoising technique, along with encoder-decoder architecture, is the most efficient one.

Dataset-2 Conclusion:

- The study has been done in colder climate region where temperature varies from -5°C to 20°C.
- As per previous results, the encoder-decoder model has been implemented. Noisiness in the data is observed and accordingly wavelet transform has been implemented.

- 1 step, 3 step and 6 step ahead of forecasting is done.
- Model gives average RMSE of 0.01-2, MAPE of 0.01% to 5% and MDA is above 90%.
- Overall model accuracy shows the robustness of the model.
- Average power generation loss in winter period is more than that in summer period and varies in the range of 3% to 5%.

The overall average accuracy has been tabulated below:

Feature Name	Average Performance of 1 step ahead			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting		
	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Wind Speed	0.008	0.066	0.99	0.01	0.09	0.99	0.05	0.5	0.99
Power Generation	0.441	0.84	0.99	0.52	0.18	0.99	1.97	0.55	0.99

Fig 6.2: Average Accuracy using WT+ED model

Dataset-3 Conclusion:

- The study has been done in hot and dusty climatic regions where temperature varies from 20°C to 48°C.
- The region of interest is the observations exceeding 40°C temperature value.
- As per previous results, the encoder-decoder model has been implemented. Noisiness in the data is observed and as per methodology, it is noted that wind gusts are impacting the turbine region and thus, denoising of the data is not considered to keep natural inherent property of wind.
- The results obtained were highlighted by implementing encoder decoder architecture for 1,3 and 6 steps ahead of power/windspeed forecasting.

- 1 step, 3 step and 6 step ahead of forecasting is done.
- Model gives average RMSE of 40-45, MAPE of 1% to 15% and MDA is above 85%.
- Average power generation loss in summer region is more than that in winter period and varies in the range of 0.5% to 2%

The overall average accuracy has been tabulated below:

Feature Name	Average Performance of 1 step ahead			Average Performance of 3 step ahead forecasting			Average Performance of 6 step ahead forecasting		
	RMSE	MAPE	MDA	RMSE	MAPE	MDA	RMSE	MAPE	MDA
Wind Speed	0.25	3.59	0.88	0.25	3.56	0.89	0.25	3.51	0.89
Power Generation	44.9	11.67	0.88	44.06	15.04	0.99	44.51	15.31	0.88

Fig 6.3: Average Accuracy using ED model for dataset-3

FUTURE WORK:

- Analysis of wind turbines in mountain regions.
- Wind turbine power potential analysis in Himalaya regions
- Generating results on more realistic data.

Modifications Performed:

1. SARIMAX model is experimented just to validate its importance – Section 5.1.1, page No 71
2. Problem formulation for SARIMAX model is added. Section 3.4 Page No-46
3. Cross Dataset experiment is done and analysis of the results along with comments are noted. Page-172
4. Justification on results of model results and their comparative analysis is done. Section 5, Page No-173

Feedback of Aug Report:

Feedback on Chapters

- **Introduction:** The introduction is written adequately. Include numbered subsections. For example, "Motivation" can be section 1.1. Also, "Extended Introduction" is an inappropriate title. It needs to be changed to some topic name. -**DONE**
- **Literature Survey:** An adequate literature survey has been done. At the end, "Proposed Method" can be rephrased to "2.8 Key findings from Literature Survey".- **DONE**
- **Problem Statement:** Equations should be numbered in a manner similar to Figures. Also, mention the problem statement in one place instead of mentioning a problem statement and then an extended problem statement.- **DONE**
- **Methodology:** Denoising the data is an important part of the work. Hence, include some discussion on denoising data and wind gusts in methodology as well.- **DONE**
- **Results:** nan
- **Conclusion and Future Work:** nan

Feedback on Presentation

No presentation has been submitted.- **Presentation Submitted for Sept Month**

Overall Feedback

Table of contents has errors. These need to be corrected. The text contains some grammatical errors. These also need to be corrected. As a part of phase 3 work, studying wind turbines in cold climates has been added. Also, in accordance with the feedback received during Phase 2 viva, analysis on abrupt changes in wind speed in the nature of wind gusts has been added in the context of denoising the data.

Feedback of Sept Report:

Dear Kaivalya,

I hope this email finds you well. Your report has been reviewed. Please find below the feedback on your report:

Some new literature has been cited pertaining to wind speed forecasting in cold climates, and the latter has been added to the problem statement. Also, by nature, wind speed can change abruptly. To accommodate such abrupt changes, wind gusts have been considered, and a workflow for denoising the data taking wind gusts into consideration has been formulated. Results obtained after applying the new methodology are expected in future reports. A general comment is that the changes made since the last report submission should be highlighted.

Comment:

1. Results for colder climatic regions and hot climatic regions have been generated.
2. As final report is getting submitted, highlighting the modified parts is avoided.

REFERENCES

- [1] Ministry of New and Renewable Wind Energy: <https://mnre.gov.in/>
- [2] Our World in Data : <https://ourworldindata.org/renewable-energy>
- [3] Danish Ministry of Climate, Energy and Building. Energy policy report, tech. rep., 2012.
- [4] G. Giebel et al. The state-of-the-art in short-term prediction of wind power—a literature overview. 2nd edition,” tech. rep., Technical University of Denmark, 2011.
- [5] Europe’s onshore and offshore wind energy potential. An assessment of environmental and economic constraints, tech. rep., 2009.
- [6] **Naixiao Wang** et al. (2023): Short-term Wind Power Probabilistic Forecasting Based on Attention Mechanism and Bidirectional LSTM.
- [7] **Aoife M. Foley** (2011): Current methods and advances in the forecasting of wind power generation.
- [8] **George Sideratos** , Nikos D. Hatziaargyriou (2020): An Advanced Statistical method for Wind Power Forecasting.
- [9] **S. wang** et. al, (2022): Short-term wind power prediction based on multidimensional data cleaning and feature reconfiguration.
- [10] **David Barbosa** et.al. (2017): Different models for forecasting wind power generation: Case study.
- [11] **G. Peter Zhang** (2001): Time-Series forecasting using hybrid ARIMA and NN model.
- [12] **Jianzahu wang** et al. (2019): A novel hybrid model for short-term wind power forecasting.

- [13] **Sulagna Mahata** et al. (2023): A Statistical Analysis Model of Wind Power Generation Forecasting for Western region of India.
- [14] **H Demolli** et al. (2020): Wind power forecasting based on daily wind speed data using machine learning algorithms.
- [15] **Farah Shahid** et. al, (2021): A novel genetic LSTM model for wind power forecast.
- [16] **Shikha Singh** et. al, (2007): Wind power estimation using ANN.
- [17] **Mohammed A.A., Al-qaness**, et al. (2022): Boosted ANFIS model using augmented marine predator algorithm with mutation operators for wind power forecasting.
- [18] **Truc Thi Kim Nguyen** et.al (2022): Robust Short-Term Wind Power Forecasting using a Multivariate Input and Hybrid Architecture.
- [19] Avinash Achar et.al. (2022) Seasonal Encoder Decoder Architecture For Forecasting.
- [20] **Dzmitry Bahdanau, KyungHyun Cho**, et al. (2016): Neural machine translation by jointly learning to align and translate.
- [21] **Minh-Thang Luong** et.al,(2015): Effective approaches to attention based neural machine translation.
- [22] **Ashish Vaswani** et al. (2017): Attention is all you need
- [23] **Irfan Prathama** et. al. (2022): Time series data forecasting and approximation by smoothing techniques.
- [24] **Liwen Quin et. al.** (2022): Multi time scale load forecasting based on wavelet threshold denoising and prophet.
- [25] ROSE Criteria https://en.wikipedia.org/wiki/Signal-to-noise_ratio
- [26] **Nguyen Thi Hoai Thu1** et al.,(2023): A Hybrid Model Of Decomposition, Extended Kalman Filter And Autoregressive - Long Short - Term Memory Network For Hourly Day Ahead Wind Speed Forecasting Vol. 27, No 9, Page 3063-3071

- [27] **Juan Pedro Llerena Caña** et al., (2021): Forecasting Nonlinear Systems with LSTM: Analysis and Comparison with EKF
- [28] **Radian Belu** et al. (2023): Effects of complex wind regimes and meteorological parameters on wind turbine performances.
- [29] **Si Han Li (2023)** Impact of climate change on wind energy across North America under climate change scenario RCP8.5
- [30] **Jon Kasper (2011)** Wind Turbine Production losses in Cold Climate.
- [31] GB/T 18710-2002 Methodology of wind energy resource assessment for wind farm China Standard Press, 2002, GB/T187102002
- [32] **Quin Lin et.al (2013)** Denoising of Wind Speed Data by Wavelet Thresholding.
- [33] **Carlsson, Viktor (2009)** . Measuring routines of ice accretion for Wind Turbine applications. Umeå: Umeå Universitet
- [34] https://en.wikipedia.org/wiki/Beaufort_scale
- [35] **Rindeskär, Erik (2010)**. Modelling of icing for wind farms in cold climate.
- [36] <https://climate.mit.edu/ask-mit/it-true-wind-turbines-dont-work-winter>
- [37] **BIR, G. & MIGLIORE, P. (2004)**. Preliminary Structural Design of Composite Blades for Two- and Three-Blade Rotors. In: NREL (ed.) *WER4.3303*. Golden, Colorado, USA
- [38] **TAMMELIN et. al (2000)** Energy Production in Cold Climate, Meteorological Publication No 41. In: INSTITUTE, F. M. (ed.). Helsinki. Finland.
- [39] **FU, P. & FARZANEH, M. 2010**. A CFD Approach for Modeling the Rime-Ice Accretion Process on a Horizontal-Axis Wind Turbine. *J. Wind Eng. Ind. Aerodyn.*, 98, 181–188
- [40] **BOSE, N. 1992**. Icing on a small horizontal-axis wind turbine-Part 1: Glazing profiles. *Journal of Wing Engineering and Industrial Aerodynamics*, 45, 75-85.
- [41] **TAMMELIN et. al (1998)** Wind Energy Production In Cold Climate. In: COMMISSION, T. E. (ed.) *WECO*. Copenhagen, Denmark

- [42] **HAALAND. et.al (2011)** Estimating Production Loss due to Icing on Wind Turbines. MSc, University of Troms
- [43] Wind Energy Article-Anti-icing and De-icing Technologies for Wind Turbines
- [44] **Mohammad Al-Khayat et. al. (2021)** Performance analysis of a 10-MW wind farm in a hot and dusty desert environment. Part 2: Combined dust and high-temperature effects on the operation of wind turbines.
- [45] [https://www.ge.com/news/reports/just-deserts-wind-turbine-can-handle-sandstorms -desert-sun](https://www.ge.com/news/reports/just-deserts-wind-turbine-can-handle-sandstorms-desert-sun)
- [46] **S.S.Chandel et.al. (2014)** Wind power potential assessment of 12 locations in western Himalayan region of India
- [47] **Bhargava Kumar et.al. (2019)** Theoretical Comparison of SARIMAX and DeepAR models for Time Series Forecasting