Ultra Short Term Wind Forecasting using Facebook Prophet

Major Project Report

Submitted in Partial Fulfillment of The Requirement For Semester IV of

MASTER OF TECHNOLOGY

IN

ELECTRICAL ENGINEERING

(Electrical Power Systems)

By

Mehta Kushal Darshakbhai

(20MEEE07)

Department of Electrical Engineering

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382 481

May 2022

Undertaking for the originality of the work

I **Mehta Kushal Darshakbhai** roll no. **(20MEEE07)**, give undertaking that the major project entitled **"Ultra Short Term Wind Forecasting using Facebook prophet"** submitted by me, towards the partial fulfillment of the requirement for the semester IV of Master of Technology in Electrical Power Systems of Nirma University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student
Date:
Place: Ahmedabad

Endorsed by:

Project Supervisor
Dr S.C. Vora
Professor
Department of Electrical Engineering
Institute of Technology
Nirma University
Ahmedabad

Certificate

This is to certify that the Major Project Report (Part-II) entitled **"Ultra Short term wind forecasting using statistical tools"** submitted by **Mehta Kushal Darshakbhai (Roll No: 20MEEE07)**, towards the partial fulfillment of the requirements for the semester IV of Master of Technology (Electrical Engineering) in the field of Electrical Power Systems of Nirma University is the record of work carried out by him under our supervision and guidance. The work submitted has in our opinion reached a level required for being accepted for examination. The results embodied in this major project work to the best of our knowledge have not been submitted to any other University or Institution for award of any degree or diploma.

Date:

Dr. S. C. Vora
Guide ,
Professor & Head of department
Department of Electrical Engineering
Institute of Technology
Nirma University
Ahmedabad

Dr. R .N .Patel
Additional Director
Institute of Technology
Nirma University
Ahmedabad

Acknowledgement

I feel fortunate by submitting this report as part of a major project (Part II),as a part of masters curriculum at Institute of Technology,Nirma University.

It has been extreme privileged to be associated with my guide, **Dr S C Vora** (Professor & head of the department,department of electrical engineering) for the major project work.I have been greatly benefited by his knowledge ,constant support and guidance.

I would also like to show gratitude to **Dr Ollaf Kolle**,Professor,Max Planck Institute of Biogeochemistry,Germany for granting access to the database used in the project.

I am sincerely thankful to **Dr R.N.Patel**,Director,Institute of Technology,Nirma University for providing wonderful resources for the project purposes.

- Mehta Kushal Darshakbhai

20MEEE07

Abstract

In the domain of Wind forecasting,Ultra Short term wind forecasting is one of the challenging and important part in the power generation as well as in the energy management system.The need of the ultra short wind forecasting in which, the wind is forecasted in the 10 minutes time stamp for the sake of the pitch and yaw control in the wind turbine to extract the maximum efficiency in the power generation.Here,the forecasting model named facebook prophet is implemented due to the its robustness in capturing the seasonality of the highly volatile wind speed.Performance evaluation metric like root mean square error(RMSE) is calculated for the model evaluation.

List of Figures

Contents

Chapter 1

Introduction

Renewable power generation like Wind power generation,Solar power generation and hydro power plant have become dominant part in electricity generation all over the wide.Among them,wind power generation is fastest growing technology according to the data released by Global Wind Energy Council(GWEC).This data also corroborates that an installed wind power generation capacity has reached 591GW by 2018,almost doubled compared with 6 year ago and almost 6 times compared ten years ago.[1]

Due to wind's highly volatile and stochastic nature ,it significantly challenges[2] the power quality in the grid,which enhances issues like fluctuations in voltage as well as frequency and harmonics[3],which is subsequently responsible for instability of the power grid system.[4] Also accurate wind forecasting enhances coordination between supply and demand done by grid system planner and operators,which is known as demand response program[5].Accurate wind power forecasting leads to micro-grid operation coordination for generation as well as load side that improves competitiveness of new energy generation in the power market.It also enhances the efficiency of wind energy utilization and power system despatching.[6]

The reliable forecasting of electricity generation through wind source is challenging issue due to its non regular pattern and stochastic behaviour[7].It also causes feature extraction an harder for the good forecasting results,which is not suitable for reliable as well as secured power system operation by advance scheduling.[8]

Wind speed forecasting will be categorized as following:[9]

1. Ultra-short-term forecasting:
   Ultra-short-term forecasting is to predict wind speed from a few minutes to 30 min in advance. It mainly uses control wind turbines and evaluates power quality.

2. Short-term forecasting:
   Short-term forecasting is to predict wind speed from 30 min to 72 min in advance.It is mainly used for reasonable dispatch of power grid and maintenance of power quality

3. Medium-term forecasting:
   Medium-term forecasting refers to the forecast of wind speed for several days, weeks and months in advance

4. Long-term forecasting:
   The long-term forecast is to forecast the future wind speed several years in advance. Medium-term or long-term wind speed forecasting is mostly used to evaluate wind resources in order to facilitate the planning and construction of wind farms.

Because of the randomness and uncertainty of wind speed, ultra-short-term wind speed forecasting has great theoretical and practical value[10]
To ensure minimum reverse capacity,it is require to submit power generation capacity in the interval of 5-15 minutes to meet the power requirement for power supply and demand in a spot electricity market.This exercise make sure the reliability and stability of the electrical power system[6].In this exercise ,it is also necessary to be done with precise economic operation as the forecasting in terms of ultra short term and short term is very complicated operation.The impact of forecasting is very high for the optimal security reverse and wind farm controlling,which participates in spot electricity market.It also strengthens to understand dynamic behaviour ,optimize performance ,control frequency as well as voltage and health monitoring of wind farm.[6][11]

1.1 Global Wind Power Generation

There is a wide technological revolution in wind power generation in onshore and offshore premises.Entire world has kept momentum to recognize the Paris agreement goal to reduce global warming by 1.5°,through wind energy. The Compound annual growth rate (CAGR) for the next five years under current policies is forecast as 6.6%. Gloabl Wind Energy Council (GWEC) Market Intelligence expects that 557 GW of new capacity will be added in the next five years – that equates to more than 110 GW of new installations each year until 2026.[12]The onshore wind market added 72.5 GW worldwide.

Wind power generation is surely poised to play major role in worldwide power transition. According to United Nations Climate Change Conference (COP26),the year on year growth(YoY) growth is 12%.[12]

The compound annual growth rate (CAGR) for onshore wind generation is 6.1 %,with approximate annual installation of 93.3 GW,expected to built 466 GW by 2022-2026.

It is believed that the global offshore generation will grow form 21.1.GW from 2021 to 31.4 GW in 2026,which will lead global installation market share to 24.4 % by 2026.

The offshore wind market enjoyed its best ever year in 2021, with 21.1 GW commissioned. That represents three times more than the previous year. New offshore installations represented 22.5% of all new installations last year, helping bring the world's total offshore capacity to 57 GW, which is 7% of global installations.

Wind energy is not growing nearly fast or widely enough to realise a secure and resilient global energy transition. At current rates of installation, GWEC Market Intelligence forecasts that by 2030 we will have less than two-thirds of the wind energy capacity required for a 1.5°C and net zero pathway, effectively condemning us to miss our climate goals. Wind power generation will boost industrial growth,just and socially responsible, while resting on a clear and viable economic proposition. [12]

1.2 Wind Power Generation in INDIA

The growth of wind energy in India is enormous and proves to be an option to mitigate the challenges to meet electricity demands, environmental pollution, greenhouse gas emission and depleting fossil fuel etc. India has the second largest wind market in Asia after China and fourth amongst the global cumulative installed countries of the world after USA and Germany.[13]

India has gain monmentum in wind power genearation with fourth highest wind installed capacity of 39.25 GW as on 2021.

For the better operation and regualtion goverment of India has ministry name ministry of New Renewable Energy (MNRE) The goverment of India has taken aggressive moves to evade problem of importing coal,green house emission,lack of fossil fuel resources,from which India is leading toward green and clean power generation.

The goverment is encouraging wind power sector though private investment and giving various kind of financial as well as fiscal incentives e.g accelated depreciation benifit,custom duty reduction for several components. Along with these,the goverment has taken initiative to boost power generation through wind energy as followed under Generation Based Incentive(GBI):

1. Technical support including wind resource assessment and identification of potential sites through the National Institute of Wind Energy, Chennai.

2. National Institute of Wind energy,Chennai provides technical assistance like identification of potential site and resource assessment of wind energy.

3. In order to facilitate inter-state sale of wind power, the inter-state transmission charges and losses have been waived off for wind and solar projects to be commissioned by March, 2022.

4. To boost initiative,the charges levied on inter state sale of wind power and losses have been relinquished for renewable energy.

5. Issued Guidelines for Tariff Based Competitive Bidding Process for Procurement of Power from Grid Connected Wind Power Projects with an objective to provide a

framework for procurement of wind power through a transparent process of bidding including standardization of the process and defining of roles and responsibilities of various stakeholders. These Guidelines aim to enable the Distribution Licensees to procure wind power at competitive rates in a cost effective manner.

6. The government has made some guideline to procure wind power at reasonable rate through transparent process of bidding.

1.3 Challenges

The volatile nature and instability of wind speed leads to the ineffective operation in power system domain and this causes higher operation and maintenance cost of the grid.An accurate wind speed forecasting gives an optimal wind energy penetration in the grid system and helps participants in the wind power market reduce economic losses.

Wind speed forecasting have many challenges as described below:[14]

- One of the first challenges in wind forecasting is to effective utilization of the environmental factors affecting wind speed.
- For the forecasting, an adequate and accurate data set is vital that might be challenging task for the particular geographical place.
- In particular ultra short term forecasting ,to capture seasonality of the data is very tough task and several researches and algorithms are being developed by the researcher to counter this issue.
- wind is consist of linear and non linearity characteristics.It is very complex task to analyze these both for the purpose of the forecasting.

1.4 Literature Review

Wind speed forecasting can be categorized in four categories depending upon the time span:[9]

1. Ultra short term wind forecasting
2. Short term forecasting
3. Medium term forecasting
4. Long term forecasting.

Also,forecasting models can be divided into four types depend upon the methodologies:

1. Physical models:
   In this model,it include numerical weather prediction (NWP) and weather researcher forecasting(WRF).These models considers various meteorological variable for the wind speed forecasting.It has shown good results in medium term and long term wind speed forecasting.[15] The advantage of physical methods is that the need for historical wind farm data is little.The disadvantage of physical methods lies in the need to establish accurate geophysical models. Rough model will reduce the accuracy of prediction. At the same time, the calculation of the NWP model is complex, which needs the help of many computers. Therefore, it is difficult to predict the ultra-short-term wind speed based on physical methods.

2. Artificial intelligence based models:
   In the era of artificial intelligence,there has been many models which are inferior in forecasting of wind speed/power forecasting domain[16].These models like support vector machine(SVM),extreme machine learning(ELM),kalman filter have shown their efficiency to capture the non linear trends in the wind speed.Deep neural technologies like long short term memory(LSTM),gated current unit(GRU) are also widely used in forecasting domain[17].

3. Hybrid models:
   This type of models are made by combining different forecasting models.The combinations of the models are based upon the characteristics and patterns of the wind

speed. [18]. Generally ,it employs a linear model for the prediction of the linear component and a non-linear model for the non-linear component in time series.[18]

4. Statistical models:

   This method is basically used for the short term wind forecasting as it enhances performance in low time stamp periods.[19].These models includes autoregressive moving average(ARMA),autoregressive integrated moving average,seasonal auto regressive integrated moving average (SARIMA) model. The input data of statistical methods are historical data of wind speed. This method constructs a statistical model by extracting the mapping relation between historical input data and the output to predict future wind speed. The main processes include model identification, parameter estimation, and model validation. Statistical methods usually do not consider the complex physical process of wind speed generation. Compared with physical methods, the statistical method has the characteristic of simpler calculation. The linear relationship of the wind speed series analysis can be easily captured. Time series method also has many shortcomings that cannot be ignored, such as poor prediction accuracy of low-order model and complex calculation in parameter estimation of high-order model. Furthermore, wind speed is not a linear time series, which is the biggest disadvantage of a time series model.

At present, ultra-short-term wind speed forecasting methods mainly include physical methods and statistical methods. Meteorological information is usually provided by numerical weather prediction (NWP) models . The typical statistical methods include time series analysis, intelligent prediction method, deep learning method, and combination forecasting method. Statistical method is an important part of the mainstream method of ultra-short-term wind speed forecasting. The following table summarises the advantages and drawbacks of the forecasting models: .

Tools	**Advantages**	**Limitations**
Artificial Neural Network	Can fit the non-linear patterns of data, and has good generalization ability and self-learning ability	High training samples Fall into local optimum easily
Hybrid Model	advantages of each single model & Improvement of prediction accuracy	Non-linear patterns of data cannot be further mined by linear combination
Time Series Model	Most reliable forecasting Accurate for short-term forecasts.	Requires large number of past input values Less accurate for long-term forecasts.

1.5 Objective and Motivation

From the literature survey ,it has been observed that Artificial intelligence and machine learning algorithms have played irreplaceable role in the wind forecasting.The reason behind its success lies upon its learning capability by trial and error method and it has increased its performance over time ,rather than just stereotype or traditional programming. In contrast ,there is little development in statistical tools for the forecasting.It should be noted that statistical tools are not just easy to build, but also have faster calculation feature.Another thing is,it has less probability for over fitting rather than any machine learning models.
The objective behind this thesis:

1. An accurate ultra short term forecasting for an optimum control pitch and yaw control
2. To capture sub daily seasonality of wind speed due its stochastic nature and to obtain possible highly accurate forecasting.

1.6 Organization of thesis

In chapter 1 ,it gives background,scenario and purpose behind this forecasting model.In chapter 2 ,there is a idea about how wind speed and power generation are related to each other.While in chapter 3 ,it gives basic idea about time series model.From chapter 4 includes proposed methodology e.g.Facebook prophet .Chapter 5 contains Deviation settlement mechanism ,while In chapter 6 ,it justifies result.In chapter 7,we have concluded the proposed methodology .In chapter 8 ,there is a future scope.

Chapter 2

Wind Speed-Power Generation Relationship

The wind turbine extracts wind energy through blades and converts it into mechanical energy by blade rotation to generate electricity and realize energy conversion.

If the wind speed at a certain moment is v, then the theoretical power that the wind turbine can capture per unit time at this time is,

$$W = 0.5.\rho.A.v^3.Cp \tag{2.1}$$

The equation 2.1 states Wind Speed -Power generation relationship.[20]
Where;
v is the wind speed;
ρ is the air density;
A is the swept area of the blade;
Cp is theoretical wind energy utilization coefficient.

The term Cp is called a power coefficient.

The power coefficient Cp is ratio between the power extracted by the wind turbine and energy available in a the wind stream It has a theoretical maximum of 0.593 (or 59.3).[21]
In this form, Equation 2.1 is called Betz's Law and the power coefficient is called the Betz limit.
In practice, the Betz limit is about 0.4–0.5.[21]

2.0.1 Power Curve

The power of wind turbine varies with wind speed and the power generation is dependent on the power coefficient.The power cannot be generated as extreme low wind speed due to insufficient torque to suffice friction.In same situation power also cannot be generated at extreme high wind speed value which can cause damage to the system.Figure 2.1 show the output power and wind speed relation for the power generation.[21]

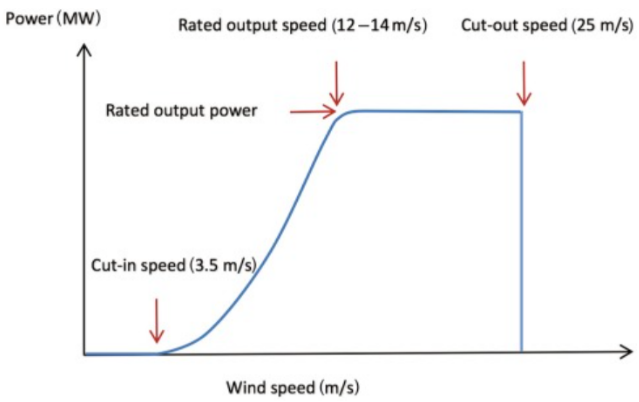

Figure 2.1: Power Curve

Various terminologies used in Figure 2.1 are explained as below:

Cut-in speed: It is the speed of the wind at which the generator starts producing power and a trickle of electricity is produced as a result.Normally cut-in speed is valued at around 3.5 m/s. At cut-in speed, the turbine generator can be used as a motor to overcome the inertia required to turn the blades.

Rated power: this is the advertised power for which a wind turbine is designed to produce under specific conditions.However, around certain speeds, known as rated wind speed (12–14 m/s), the power output reaches a limit that is called the rated power output

Rated speed: It is minimum speed of the wind at which the generator delivers advertised power.At this speed the activation of blade regulation happens.The range lies between 11–15 m/s.

Cut-out speed: At this speed ,protection mechanism is activated to stop power generation.The value of cut-out speed is around (25 m/s).

Chapter 3

Bayesian Structural Time Series Model

The bayesian time series model is consist of many time series forecasting model.It does consist of trend,seasonality,regression and error.This model is responsible for independent as well as additional impact of regressors in model.This model uses Markov chain monte carlo (MCMC) methods.An essential regressor will be selected based on overall frequency of numbers selected over the total iterations.[22]
It is worth to note that Bayes theorem is highly considered in the domain of data science and forecasting,It is also widely in use such as multi-step ahead prediction,general estimation and analysis of statics.[22]
Theoretically, in Bayes theorem, if X and Y are two events, then the probability of event X with the occurrence of event Y can be calculated using equation 3.1 This is the joint probability of two events and does not suggest symmetrical characteristics.[22]

$$P(X \mid Y) = \frac{P(X) \cdot P(Y \mid X)}{P(Y)}, \text{ for } P(Y) \neq 0 \tag{3.1}$$

In equation 3.1 Bayes theorem is defined with the following terms:
$P(X|Y)$: posterior probability;
$P(X)$: prior probability;
$P(Y|X)$: likelihood probability;
$P(Y)$: evidence.
If the value of the prior, likelihood, and evidence is known, the posterior probability can

be calculated mathematically.

The model uses a Bayesian framework based on prior distributions on the parameters to perform posterior inference including forecast uncertainty.Prophet performs either maximum a posterior probability (MAP) estimation or full Bayesian statistical inference with Markov Chain Monte Carlo (MCMC) sampling to fit the training data. To make a forecast in the test period, Prophet assumes that the average frequency and magnitude of trend changes in the future will be the same as that observed in the history

Chapter 4

Facebook Prophet

Facebook Prophet was developed by **Sean J Taylor** and **Ben Letham**.Prophet is a procedure for forecasting time series data based on an additive/multiplicative model.It can fit non-linear trends with sub daily,daily,weekly ,monthly or yearly seasonality. It can work robustly where the seasonality in data has robust effect.[23]
Here some important inferential advantages of using a generative model such as an ARIMA, this formulation provides a number of practical advantages:[23]

- Flexibility: It has feature to include multiple seasonality and also it can accommodate flexible trend points in data points according to the requirement.
- It has feature of being comparatively fast.
- Unlike ARIMA,it does not have long checklist of parameters for tuning.
- The forecasting model has easily interpret able parameters that can be changed by the analyst to impose assumptions on the forecast. It would be advantageous if practitioner has adequate knowledge about predictor and regresors.

Facebook Prophet is an open-source algorithm for generating time-series models that have multiple seasonality and trend points . At its core is the sum of three functions of time plus an error term: growth $g(t)$, seasonality $s(t)$ and error $e(t)$:

4.1 Seasonality

Seasonality in time-series data is a pattern that occurs at a regular interval.There is lot of data insight in seasonality for the forecasting model.It is necessary to make estimation of seasonality adequately.
Seasonality can be estimated by:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right)\right) \tag{4.1}$$

Where,P is regular period data.

It requires estimation of $2N$ parameters $\beta =\ [a_1, b_{1,\ldots}, a_N, b_N]^\top$

This can be achieved by seasonality matrix;

$$X(t) = \left[\cos\left(\frac{2\pi(1)t}{P}\right), \ldots, \sin\left(\frac{2\pi(N)t}{P}\right)\right] \tag{4.2}$$

The seasonal component can be defined by:

$$S(t) = X(t)\beta \tag{4.3}$$

Due to sub daily seasonality N has come to value 2,while in case of yearly and weekly seasonality it has ranged from 3 to 10.

Additive Seasonality

It's pretty rare for actual time series to have constant crest and trough values and instead, we typically see some kind of general trend like an increase or a decrease over time. In our sales price plot, for example, the median price tends to go up over time. If the amplitude of our seasonality tends to remain the same, then we have what's called an additive seasonality. Below is an example of an additive seasonality.
We can even think of our basic cosine model from earlier as an additive model with a constant trend! We can model additive time series using the following simple equation:

$$Y[t] = T[t] + S[t] + e[t] \tag{4.4}$$

where; $Y[t]$: time-series function
$T[t]$: Trend (general tendency to move up or down)
$S[t]$: Seasonality (cyclic pattern occurring at regular intervals)
$e[t]$: Residual

Multiplicative Seasonality

In this type, the amplitude of our seasonality becomes larger or smaller based on the trend.
We can model this with a similar equation as our additive model by just swapping the additions for multiplications.

$$Y[t] = T[t] * S[t] * e[t] \tag{4.5}$$

where; $Y[t]$: time-series function
$T[t]$: Trend (general tendency to move up or down)
$S[t]$: Seasonality (cyclic pattern occurring at regular intervals)
$e[t]$: Residual

4.2 Trend

There are two trend models supported by Prophet which have been described below:

4.2.1 Growth Mode

In time series data,there shall be small bends toward upwards or downwards in the trend.This trend is called growth mode.Modeling our data set with an unsuitable growth mode will lead the modelling to over fit or under fit.
Prophet supports two types of growth:

1. Logistic Growth Mode

2. Flat Growth Mode

4.2.2 Logistic Function

:

The logistic function generates an S shaped curve and the equation ,

$$\frac{L}{1 + e^{-k(x-x_0)}} \tag{4.6}$$

The above equation was developed by Pierre Francois Verhulst. With logistic growth,Prophet always requires a value that the data will never surpass,which is known as ceiling.In contrast,in case of the declining growth ,the term is known as floor.

4.2.3 Flat growth mode

In flat growth mode ,the trend line remains almost constant throughout the entire data set.The reason behind deviation in constant line might be due to the seasonality,additional regressor or the noise. In figure 4.29 it is clear is visible that the trend remains constant throughout all the years.

Hence,We will consider Flat growth mode for this forecasting model.

4.3 Error

The error term error $e(t)$ describes any peculiar changes which are not considered in this model.Here ,in this model it is assumed that these error are normally distributed in the data points.

4.4 Data Source and Data Preparation

The data of this experiment is from Max Planck Institute for Biogeochemistry,Jena(Germany).Jena is located in Northern Hemisphere.Due to its distance to coastal areas and position in the Saale valley, wind speeds tend to be very low; predominant direction is South West.

All the simulations are performed on the Jupyter Notebook using Anaconda Package for data science.The computer configuration has Intel i5 $8th$ generation,4 core with base clock speed of 1.8 Ghz speed. The data were collected in 10 minutes time interval and summary of the data is described in the following part.

Here,the outliers were removed as well as the missing data points were imputed by forward backward filling method . Total number of data points are 3,15,789. The forecasting is done for 80,789 data points.

4.4.1 Statistical Description

In Jena,Germany ,there are four different seasons as described below:

1. Spring: March to May
2. Summer: June to August
3. Autumn: September to November
4. Winter: December to February

Here ,Season wise statistical description is approached for analysis:

1. Spring

Figure 4.1: Weekly Averaged Wind speed in 2008

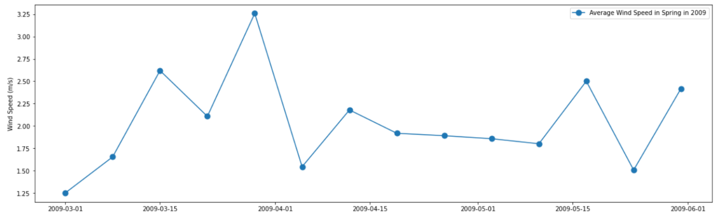

Figure 4.2: Weekly Averaged Wind speed in 2009

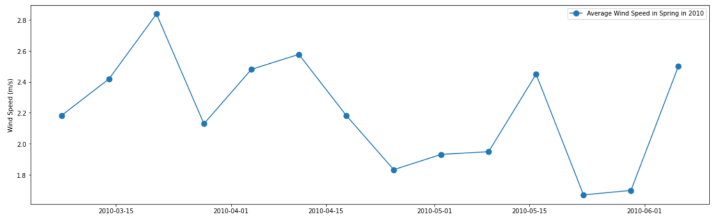

Figure 4.3: Weekly Averaged Wind speed in 2010

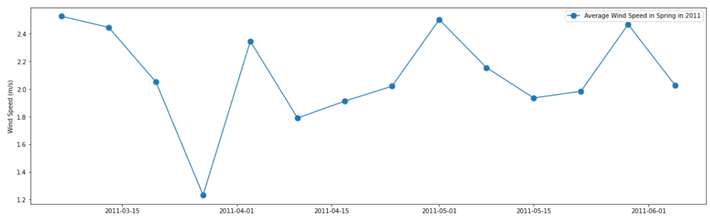

Figure 4.4: Weekly Averaged Wind speed in 2011

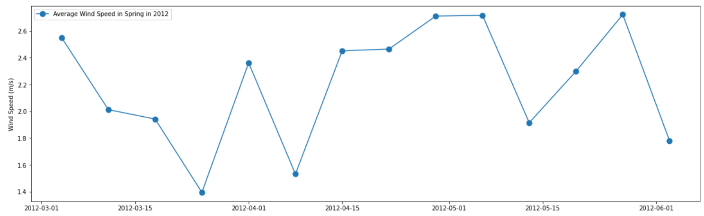

Figure 4.5: Weekly Averaged Wind speed in 2012

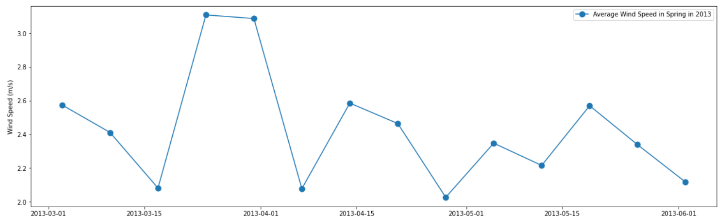

Figure 4.6: Weekly Averaged Wind speed in 2013

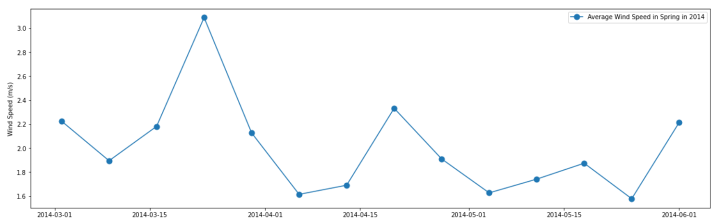

Figure 4.7: Weekly Averaged Wind speed in 2014

2. Summer

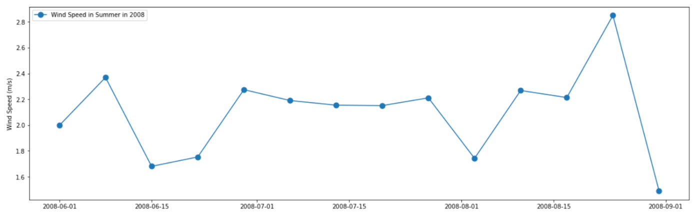

Figure 4.8: Weekly Averaged Wind speed in 2008

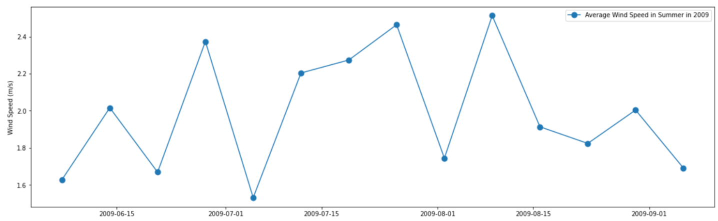

Figure 4.9: Weekly Averaged Wind speed in 2009

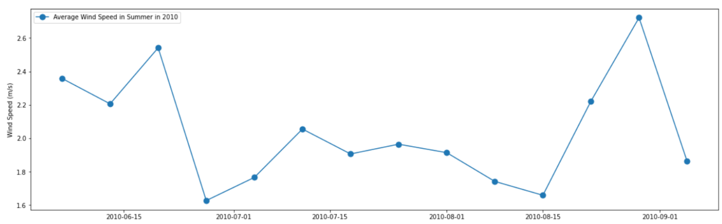

Figure 4.10: Weekly Averaged Wind speed in 2010

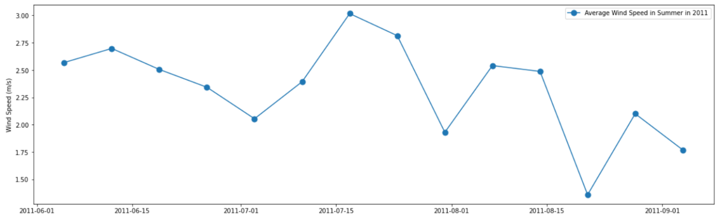

Figure 4.11: Weekly Averaged Wind speed in 2011

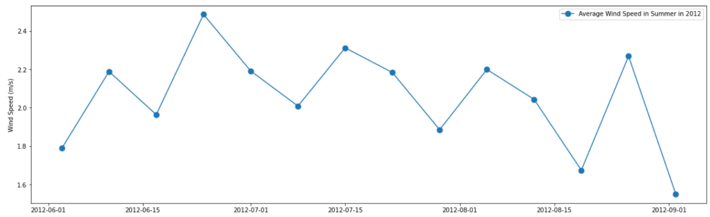

Figure 4.12: Weekly Averaged Wind speed in 2012

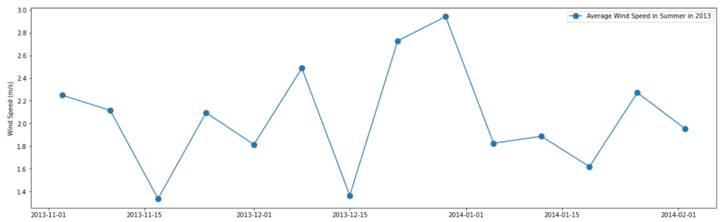

Figure 4.13: Weekly Averaged Wind speed in 2013

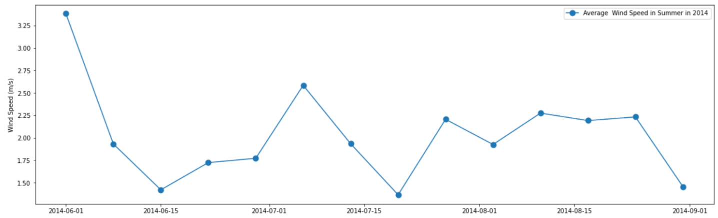

Figure 4.14: Weekly Averaged Wind speed in 2014

3. Autumn

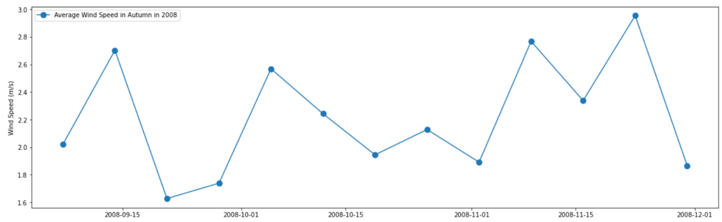

Figure 4.15: Weekly Averaged Wind speed in 2008

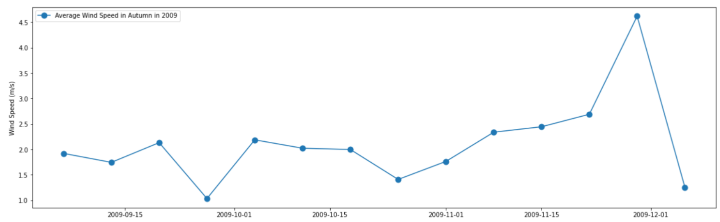

Figure 4.16: Weekly Averaged Wind speed in 2009

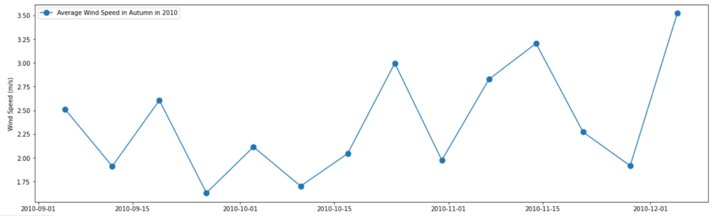

Figure 4.17: Weekly Averaged Wind speed in 2010

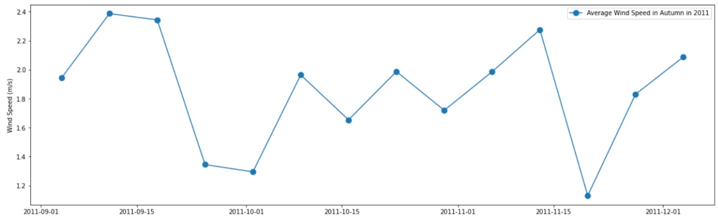

Figure 4.18: Weekly Averaged Wind speed in 2011

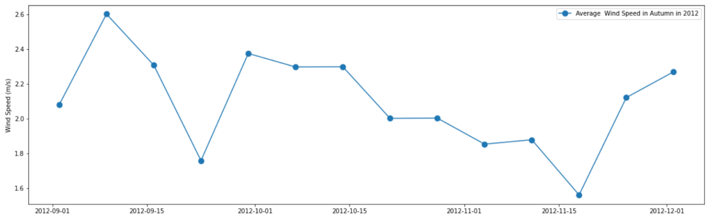

Figure 4.19: Weekly Averaged Wind speed in 2012

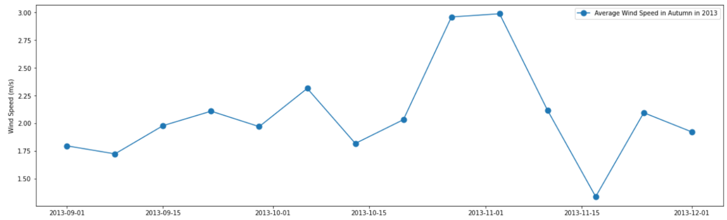

Figure 4.20: Weekly Averaged Wind speed in 2013

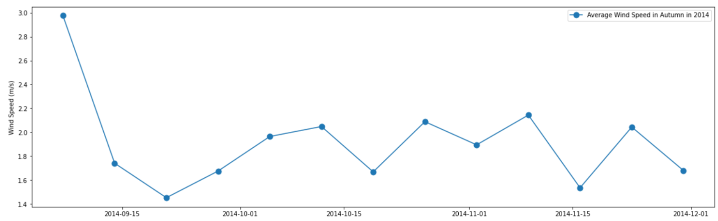

Figure 4.21: Weekly Averaged Wind speed in 2014

4. Winter

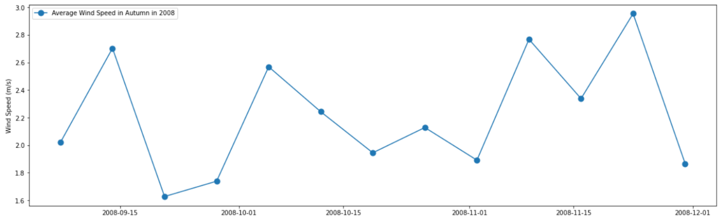

Figure 4.22: Weekly Averaged Wind speed in 2008

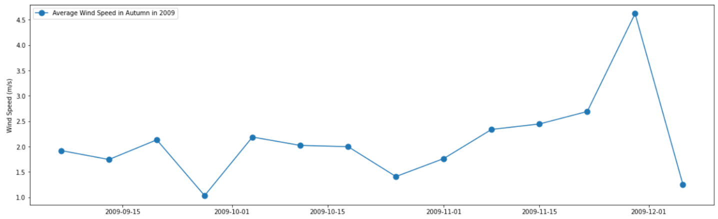

Figure 4.23: Weekly Averaged Wind speed in 2009

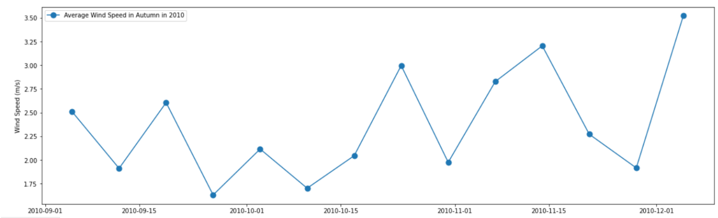

Figure 4.24: Weekly Averaged Wind speed in 2010

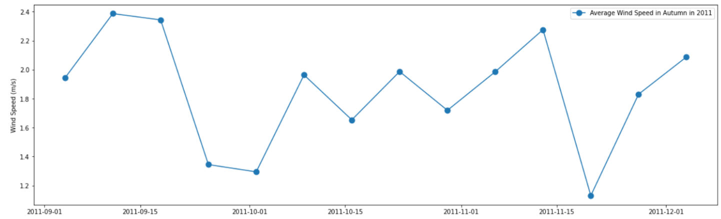

Figure 4.25: Weekly Averaged Wind speed in 2011

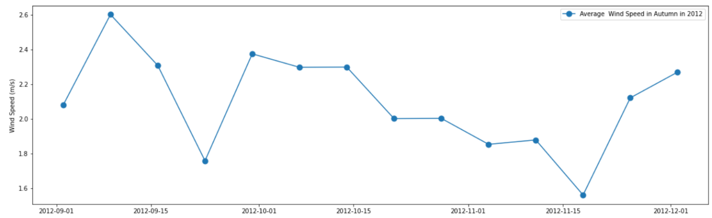

Figure 4.26: Weekly Averaged Wind speed in 2012

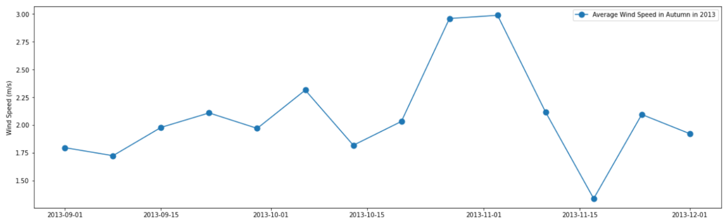

Figure 4.27: Weekly Averaged Wind speed in 2013

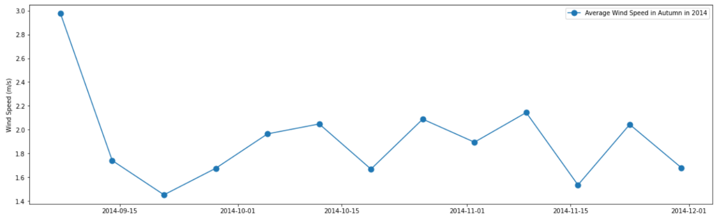

Figure 4.28: Weekly Averaged Wind speed in 2014

Statistical Summary:

Season	Year	Max Wind Speed (m/s)	Min Wind Speed (m/s)	Mean Wind Speed (m/s)
Spring	2008	12.05	0	2.40
	2009	10.15	0	2.08
	2010	10.04	0	2.18
	2011	11.78	0.14	2.09
	2012	14.09	0.12	2.20
	2013	12.32	0.12	2.42
	2014	12.21	0.11	2.00
Summer	2008	10.59	0.0	2.10
	2009	9.76	0.0	2.00
	2010	10.71	0.14	2.04
	2011	12.53	0.13	2.34
	2012	10.62	0.10	2.07
	2013	9.32	0.11	1.89
	2014	12.63	0.08	1.93
Autumn	2008	10	0.0	2.0
	2009	12.96	0.0	2.16
	2010	10.55	0.12	2.30
	2011	11.54	0.13	1.83
	2012	8.92	0.14	2.09
	2013	13.19	0.10	2.10
	2014	10.20	0.12	1.91
Winter	2009	9.71	0.0	2.04
	2010	10.91	0.12	2.18
	2011	11.54	0.13	2.29
	2012	12.12	0.13	2.29
	2013	9.64	0.10	2.04
	2014	12.86	0.12	2.36

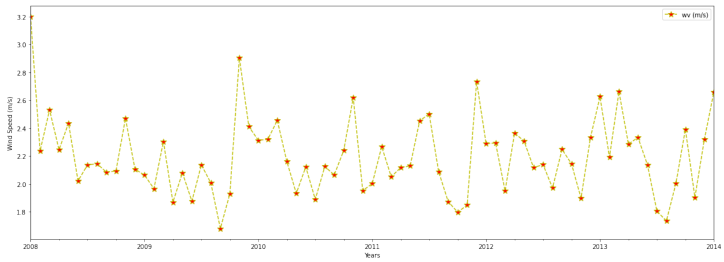

Figure 4.29: Mean Wind Speed from 2008 to 2014

The above figure 4.29 is overall mean wind speed value on monthly basis throughout these 6 years. The statistical mean value stands at 2.4 m/s.While maximum value is 14.2 m/s and minimum value (after removing negative entries) is 0 m/s.The data is stationary.One of the finest advantages of using Facebook prophet is found that ,there is no such necessary to make data stationary as in case of other forecasting models like ARIMA,it is necessary to make data stationary and requires lots of exercise.

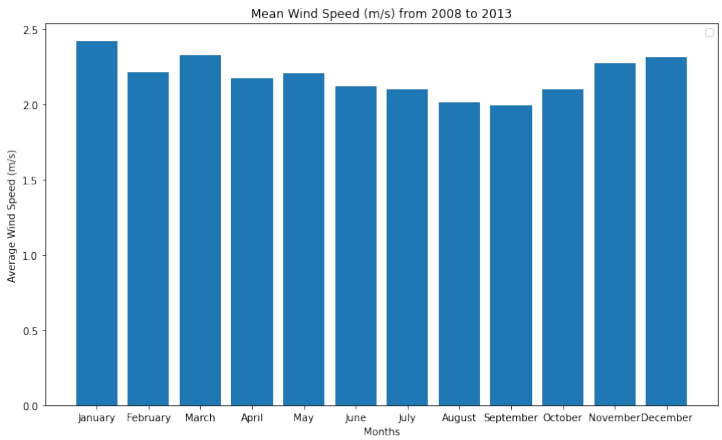

Figure 4.30: Monthly Mean Wind Speed from 2008 to 2014

Above figure 4.30 shows that there throughout all the years,there is no mean wind speed value change on monthly basis .

4.4.2 Feature Engineering

The output of power is dependent on many atmospheric variable e.g.temperature,humidity,pressure etc.For the better efficiency of the model it is required to select parameters precisely.

Because different meteorological factors have different influences on wind farms' power output, if the number of input meteorological variables are not enough, there may be no features with high correlation in the input . However, not all meteorological factors have important influence on wind power prediction. If we introduce many irrelevant vectors, we will increase the complexity of the algorithm and reduce the time efficiency and the prediction accuracy.

Pearson correlation coefficient is one of many statistical correlation coefficients, also known as product-difference correlation[24].Assuming there are two data sequences X and Y, the pearson correlation coefficient between two variables is calculated as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2(y_i - \overline{y})^2}} \tag{4.7}$$

The range of pearson correlation coefficient is $[1, 1]$. If the coefficient is positive, the two data are positively correlated; if the coefficient is negative, the two data are negatively correlated. The larger the absolute value of the correlation coefficient is, the stronger the correlation is. The closer the correlation coefficient is to 0, the weaker the correlation is. According to the absolute value, researchers defined the correlation degree as: 0.8 to 1.0 is very strong correlation, 0.6 to 0.8 is strong correlation, 0.4 to 0.6 is moderate correlation, 0.2 to 0.4 is weak correlation, and 0 to 0.2 is very weak correlation or no correlation.

4.4.3 Data Preparation

Through data sorting and visual processing, it can be found that there are missing and abnormal phenomena in the data due to sensor faults, data transmission errors, data input errors, unit operation adjustment and other problems[25]. For missing data and abnormal data, we use forward and backward filling method. There has been an incident that many values in wind speed are negative due to whatsoever reason.These values have been removed from the dataset. It is necessary to give date time column name :ds and predictor name i.e.wind speed(m/s) :y.

We have normalized the raw wind speed data in the data preprocessing stage using normalization; the processed data were normally distributed over the interval (0, 1).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.8}$$

Where;
X=Value of particular parameter,
X_{min}=minimum value in the dataset,
X_{max}=maximum value in the dataset

Here,We have considered following features for the forecasting modelling.

1. pressure in millibar
2. temperature in kelvin
3. dew point in degree Celsius
4. relative humidity in %
5. vapour pressure in millibar
6. vapour pressure deficit in millibar
7. specific humidity
8. maximum wind speed in m/s

4.5 Modelling uncertainty in Seasonality

In Facebook prophet ,the uncertainty in seasonality is modelled by MCMC chain.A Markov chain models a sequence of events ,dependent upon the previous state of event.Markov chain model is a stochastic model which has Markov property. Markov property is satisfied when current state of the process is enough to predict the future state of the process and the prediction should be as good as making prediction by knowing their history. It is a very easy process to model random process.[26]. A Markov process Xt is a stochastic process with the property that, given the value of Xt, the values of Xs for $s > t$ are not influenced by the values of Xu for $u < t$. In words, the probability of any particular future behavior of the process, when its current state is known exactly, is not altered by additional knowledge concerning its past behavior. A discrete-time Markov chain is a Markov process whose state space is a finite or countable set, and whose (time) index set is T = (0, 1, 2, . . .).

In formal terms, the Markov property is that[27]

$$\begin{aligned} &\Pr\{X_{n+1}=j \mid X_0=i_0,\ldots,X_{n-1}=i_{n-1},X_n=i\} \\ &=\Pr\{X_{n+1}=j \mid X_n=i\} \end{aligned} \tag{4.9}$$

for all time points n and all states i0, . . . , in–1, i, j.

Here we have taken 1000 MCMC sample which is by default value.

4.6 Modelling uncertainty in Trend

Prophet estimates uncertainty in trend through maximum a posteriori (MAP) estimation. Here,a maximum a posteriori probability (MAP) estimate is an estimate of an unknown quantity, that equals the mode of the posterior distribution. According to the maximum a posteriori probability (MAP) method [28], the unknown set of parameters is treated as a random vector and its posterior, for a given set of output observations y is expressed as

$$p(\boldsymbol{\theta} \mid \boldsymbol{y})=\frac{p(\boldsymbol{y} \mid \boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})} \tag{4.10}$$

where $p(\theta)$ is the associated prior density function.

Here,this model simulates 1,000 different future trend to estimate uncertainty,which is by default value.

4.7 Hyper Parameter Tuning

In machine learning ,it is necessary for a model to capture a all points to be considered as a good model.Incase it does nit capture all the points ,it is known as underfit model.In such case,model does not perform well on training as well as testing dataset.

When a model capture the data points beyond its true relation of actual and predicted value,this model is known as overfit model.In such scenario,the model performs well in training data points and moderate to low on testing data points.

In prophet sometimes it happens that the model goes for the overfitting,depending upon the types of dataset and parameter tuning.To avoid both these scenarios,it is necessary to regularize the model for better performance.

Here in prophet package, there are several packages available for the regularization.In case of the seasonality,the parameter is known as "prior scale".The prior probablity distribution is also known as prior,of the particular quantity is the probability distribution of values that we expect prior to learning some more additional information.

Another point for the consideration for tuning is,trend point detection,which is known as changepoint prior scale.The trend change point are the locations in our dataset where the trend components of the model suddenly changes its slope.This values lie between 0.001 to 0.5.
In real life,for the parameter tuning Grid Search method is use,in which there is a library called itertools.It gives various forecasting model and corresponding metric parameter with its lowest value.But here ,due to constraints in dataset it could not be useful here and parameters are chosen on trial and error basis.

4.8 Cross Validation

In machine learning it is sacrosanct to keep the training data and testing data separate.It is said not to train a model and test its performance on the same data.Setting data aside for testing purposes has a downside though that data has some vital information that we would want to include in the training.

For the purpose of the cross validation of the dataset, the full dataset is split into three parts;

1. Training data 60 %
2. Validation data 20 %
3. Testing data 20 %

After the this datset is split into these three sets,the model is trained on the training dataset and performance of this model is evaluated on the validation dataset.Here,a new set of hyperparmater is selected for this particular algorithm and the model is get traine again on the training dataset and the re-evaluation is done on the validation dataset.This repeative procedure is done for the numbers of the hyperparameter combinations.

The set of the hyperparameters with the good performance on the validation dataset is selected for the model ,the trained datsets and validation datasets are combined to train the model and this model is evaluated on the test dataset.This evaluation is considered as the final performance of the model.

4.8.1 Forward Chaining Cross Validation

In this proposed methodology,the cross validation technique used name is **Forward Chain Cross Validation** also known as **Rolling Origin Cross Validation**.As the type of this dataset is time-series,there shall be no random reshuffling of data to begin but a test set must be set aside.It should be keep in mind that the test dataset must be the final portion of the dataset.

Terminologies used to obtain Cross Validation:

1. **Initial**: It is the minimum quantity of the dataset required for the beginning of the training operation.

2. **Horizon**: It is the length of the time for the forecasting is required.

3. **Period**:It is the amount of time between each fold.It could be greater,lesser or equal to the horizon time.

In machine learning and statistical learning,only training and testing dataset could lead the model to the biased modelling ,but by cross validation the forecasted value would not have any biased value.Hence,the model shall provide a more accurate representation ,what we can expect our model to forecast new and unseen data.

Chapter 5

Deviation Settlement Mechanism Regulations 2022

Due to increasing penetration of renewable power generation like wind power generation and solar power generation,it is vital to consider forecasting scheduling and deviation settlement mechanism(DSM).DSM is regulation board for the grid operation in which it achieves stability by imposing incentive / penalties.The central electricity regulation commission (CERC) has made regulations for the forecasting domain.These kind of regulations help in regulating the quantum of electricity being injected into the grid and provide a mechanism to penalise plants for deviation from the scheduled generation beyond a permissible limit.

It is mandatory for every power producers to send their forecasted power generation to their respective load despatch centres for the better grid stability.Grid operators imposes penalties in case of over/under power generation.

Error In forecasting may be caused by either by weather parameters or due to changes in DSM across the states.The permissible forecasting error changes states by states in India.Mostly all across India ,it is $\pm$ 15 %,but in some states like Haryana ,Madhya Pradesh and Tamil Nadu ,it is $\pm$ 10 %.[29]

In this forecasting model,the RMSE is 0.09%.Taking general scenario by considering regulation mechanism in India,we can say that the output can be considered for ultra short term wind forecasting.

Chapter 6

Result and discussion

There are several accepted performance metrics are used to gauge the forecasting output.By definition,root mean square error(RMSE) is the standard deviation for the forecasting error.[30]

Here,the criterion for this model performance is root mean sqaure error (RMSE). RMSE can be defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{6.1}$$

The equation 6.1 states performance of the forecasting model.
where,
$\hat{y}_i$ is predicted value while,
y_i is real value.

The performance of the model is shown in below figure 6.1.It can be observed that ,most of the time the forecasting model is justified .In a word,except for some partial time this model has best fitting performance on the given dataset.

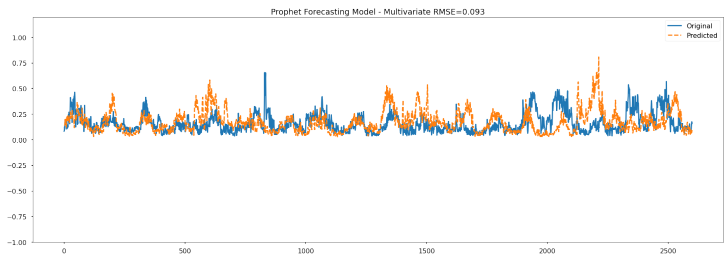

Figure 6.1: Forecasting Output

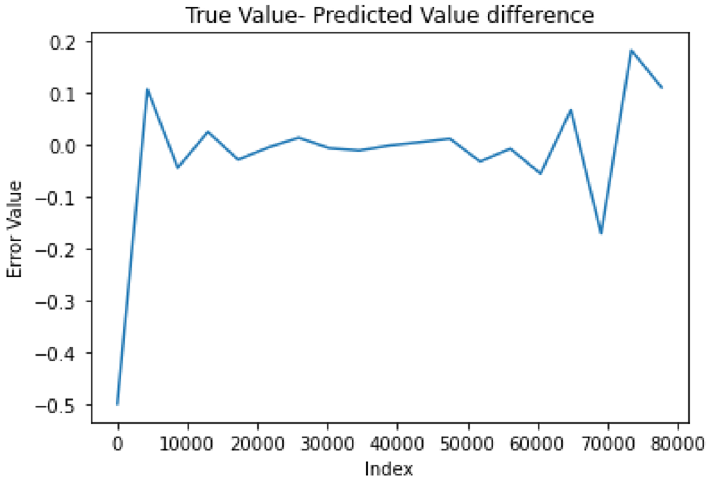

Figure 6.2: Error difference

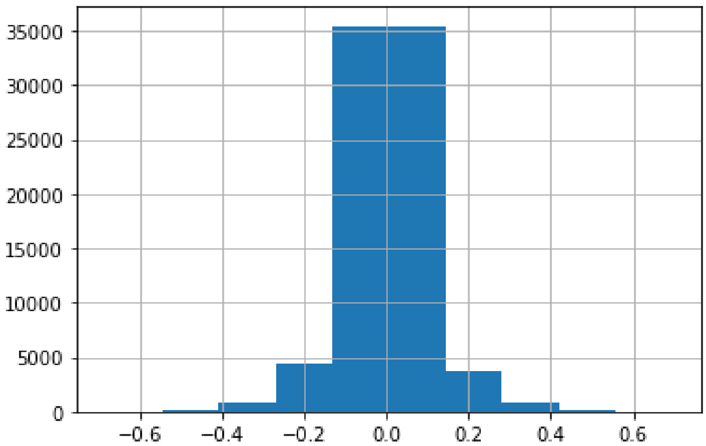

Figure 6.3: Error Histogram

The ultra short term wind forecasting errors are shown in fig 6.1.The overall RMSE stands at 0.09. The ultra short term wind forecasting error distribution is shown in 6.3.It is also observed from the fig 6.3 that most of errors lie in the range of [-2,2] and small portion goes beyond that limit. i.e over forecast and under forecast has almost same range. From this it is possible to make estimation of reserve in wind energy[31].Fig It is visible from 6.2 that the reserve estimation shall be considerable using this forecasting model. An variations in errors are quite stochastic that causes time series resembling random walk in a short period and that make an algorithm difficult to track its patterns.

It is clearly visible that error distribution is quite recommendable.The distribution is concentrated as well as non biased performance for ultra short term wind forecasting.

Chapter 7

Conclusion

In wind forecasting,it is necessity to pay heed for the multiple aspect affecting highly volatile wind speed such as altitude,topography,temperature and pressure.These components make forecasting inaccurate due to its uncertain nature.
In this thesis ,we can conclude the following things:

1. In this result ,we can say that the performance of this forecasting model is quite satisfactory.The error caused here is due to absence of the variation of the parameters due to its stochastic nature,which leads to the model hard to capture the patterns.

2. Here,a new approach leads to the improving accuracy in wind speed for the power generation for the safe and stable grid operation.

3. his thesis puts forward an ultra- short-term forecasting model of wind power output based on many meteorological factors.We integrated lots of meteorological parameters and carried out correlation analysis based on correlation coefficient and redundancy analysis to improve the calculation speed and solution accuracy.Finally, we verified the effectiveness of the proposed model by experiments.

Chapter 8

Future Scope

Based on this study, there are still some challenges that should be addressed in further research. More extensive explorations and investigations in the wind system should be conducted. The research directions to be considered are as follows:

The physical model based on NWP model has advantages in long-term prediction. Therefore, the combination of physical model and the statistical model is not only a research direction, but also a very promising direction

For statistical models, the instability and dependence on data are an obvious limitation that needs to be addressed. Due to the multiple wind speed fluctuation patterns,the selection of training data has a significant influence on forecasting performance.

Although in this methodology,It has given one of the finest results for the forecasting,the challenges the are to capture the seasonality in the ultra short term wind forecasting.Most existing forecasting methods usually suffer from insufficient precision and high complexity.[6]

In this work,the wake effect is excluded.The wind speed forecasting for a particular location shall be inadequate if the intra-farm wakes and inter farm wakes are excluded.This would impact an overstimation in wind speed forecasting results.This would lead to great discrepancies in power output estimation due to cubic relation power output and wind speed by considering equation 2.1.Even a small gap in forecast will generate three times error in generated output.[32]

References

[1] Z. Zhen, G. Qiu, S. Mei, F. Wang, X. Zhang, R. Yin, Y. Li, G. J. Osório, M. Shafie-khah, and J. P. Catalão, "An ultra-short-term wind speed forecasting model based on time scale recognition and dynamic adaptive modeling," *International Journal of Electrical Power and Energy Systems*, vol. 135, 2 2022.

[2] Y. Pan, F. Liu, L. Chen, J. Wang, F. Qiu, C. Shen, and S. Mei, "Towards the robust small-signal stability region of power systems under perturbations such as uncertain and volatile wind generation," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1790–1799, 2017.

[3] Y.-K. Wu, J.-J. Zeng, G.-L. Lu, S.-W. Chau, and Y.-C. Chiang, "Development of an equivalent wind farm model for frequency regulation," *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 2360–2374, 2020.

[4] H. Chen, J. Zhang, Y. Tao, and F. Tan, "Asymmetric garch type models for asymmetric volatility characteristics analysis and wind power forecasting," *Protection and Control of Modern Power Systems*, vol. 4, no. 1, pp. 1–11, 2019.

[5] F. Wang, H. Xu, T. Xu, K. Li, M. Shafie-Khah, and J. P. Catalão, "The values of market-based demand response on improving power system reliability under extreme circumstances," *Applied energy*, vol. 193, pp. 220–231, 2017.

[6] M. A. Hossain, R. K. Chakrabortty, S. Elsawah, and M. J. Ryan, "Very short-term forecasting of wind power generation using hybrid deep learning model," *Journal of Cleaner Production*, vol. 296, 5 2021.

[7] P. Du, J. Wang, Z. Guo, and W. Yang, "Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting," *Energy Conversion and Management*, vol. 150, pp. 90–107, 2017.

[8] Y. Lin, M. Yang, C. Wan, J. Wang, and Y. Song, “A multi-model combination approach for probabilistic wind power forecasting,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 226–237, 2018.

[9] Z. Tian, “A state-of-the-art review on wind power deterministic prediction,” 10 2021.

[10] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, “A review of wind speed and wind power forecasting with deep neural networks,” *Applied Energy*, vol. 304, 12 2021.

[11] M. A. Mohandes and S. Rehman, “Wind speed extrapolation using machine learning methods and lidar measurements,” *IEEE Access*, vol. 6, pp. 77634–77642, 2018.

[12] “Annual-wind-report-2022$_{s}creen_{f}inal_{a}pril,''$

[13] G. Heal, “The economics of renewable energy,” 2009.

[14] X. Pan, L. Wang, Z. Wang, and C. Huang, “Short-term wind speed forecasting based on spatial-temporal graph transformer networks,” *Energy*, vol. 253, p. 124095, 2022.

[15] J. Hu, J. Heng, J. Wen, and W. Zhao, “Deterministic and probabilistic wind speed forecasting with de-noising-reconstruction strategy and quantile regression based algorithm,” *Renewable Energy*, vol. 162, pp. 1208–1226, 2020.

[16] J. Wang, S. Wang, and W. Yang, “A novel non-linear combination system for short-term wind speed forecast,” *Renewable Energy*, vol. 143, pp. 1172–1192, 2019.

[17] M. Neshat, M. M. Nezhad, E. Abbasnejad, S. Mirjalili, L. B. Tjernberg, D. Astiaso Garcia, B. Alexander, and M. Wagner, “A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the lillgrund offshore wind farm,” *Energy Conversion and Management*, vol. 236, p. 114002, 2021.

[18] J. Shi, J. Guo, and S. Zheng, “Evaluation of hybrid forecasting approaches for wind speed and power generation time series,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3471–3480, 2012.

[19] J. Jung and R. P. Broadwater, “Current status and future advances for wind speed and power forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, 2014.

[20] P. Kosky, R. Balmer, W. Keat, and G. Wise, "Chapter 16 - green energy engineering," in *Exploring Engineering (Third Edition)* (P. Kosky, R. Balmer, W. Keat, and G. Wise, eds.), pp. 339–356, Boston: Academic Press, third edition ed., 2013.

[21] S. P. Neill and M. R. Hashemi, "Chapter 4 - offshore wind," in *Fundamentals of Ocean Renewable Energy* (S. P. Neill and M. R. Hashemi, eds.), E-Business Solutions, pp. 83–106, Academic Press, 2018.

[22] A. Shrestha, B. Ghimire, and F. Gonzalez-Longatt, "A bayesian model to forecast the time series kinetic energy data for a power system," *Energies*, vol. 14, no. 11, p. 3299, 2021.

[23] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

[24] J. Jiading, W. Feng, T. Rui, Z. Lingling, and X. Xin, "Ts_xgb: Ultra-short-term wind power forecasting method based on fusion of time-spatial data and xgboost algorithm," *Procedia Computer Science*, vol. 199, pp. 1103–1111, 2022.

[25] Y. Zhang, B. Zhou, X. Cai, W. Guo, X. Ding, and X. Yuan, "Missing value imputation in multivariate time series with end-to-end generative adversarial networks," *Information Sciences*, vol. 551, pp. 67–82, 2021.

[26] A. Biswal, M. D. Borah, and Z. Hussain, "Chapter eleven - music recommender system using restricted boltzmann machine with implicit feedback," in *Hardware Accelerator Systems for Artificial Intelligence and Machine Learning* (S. Kim and G. C. Deka, eds.), vol. 122 of *Advances in Computers*, pp. 367–402, Elsevier, 2021.

[27] M. Pinsky and S. Karlin, *An introduction to stochastic modeling*. Academic press, 2010.

[28] S. Theodoridis, "Chapter 12 - bayesian learning: Inference and the em algorithm," in *Machine Learning* (S. Theodoridis, ed.), pp. 585–638, Oxford: Academic Press, 2015.

[29] "Central electricity regulatory commission new delhi."

[30] "Multi-step ahead ultra-short-term wind power forecasting based on time series analysis."

[31] J. Dowell, G. Hawker, K. Bell, and S. Gill, "A review of probabilistic methods for defining reserve requirements," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2016.

[32] R. M. A. Feroz, A. Javed, A. H. Syed, S. A. A. Kazmi, and E. Uddin, "Wind speed and power forecasting of a utility-scale wind farm with inter-farm wake interference and seasonal variation," *Sustainable Energy Technologies and Assessments*, vol. 42, p. 100882, 2020.