ATMOSPHERIC EFFECTS ON SHORT TERM WIND POWER FORECASTING

A Thesis Submitted to
the Graduate School of Engineering and Science of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Energy Engineering

by
Yüksel Kalay

July 2021
İZMİR

ACKNOWLEDGMENTS

I would like to express my special thanks of gratitude to my advisor Assist. Prof. Dr. Ferhat BİNGÖL for his inspiring scientific point of view and continuous motivation throughout my research.

I would also like to thank my mother Meral KALAY and my father Mustafa KALAY for their endless supports and self-sacrifices up until now.

ABSTRACT

ATMOSPHERIC EFFECTS ON SHORT TERM WIND POWER FORECASTING

Wind power all over the world are being popularizing unlike decrease in conventional sources due to environmental issues. However, power acquired from wind is not stable during day and night, which means that intermittent due to nature of the source. Forecasting in wind power plant is very challenging compared to forecasting of production of conventional power plant. Although there are many robust and site-specific models in order to forecast wind power accurately, decrease of deviation in wind power forecasting by using statistical, physical and hybrid models is still open to new approaches. In this study, four different forecast models based numerical weather prediction (NWP) models for three different wind farms which have different atmospheric conditions are examined to improve wind farm-based power forecasting. For this purpose, wind power forecasting of the providers was categorized based on atmospheric effects, which are site temperature and turbulence. Results have been compared with real time power production from wind turbine supervisory control and data acquisition (SCADA) system. Afterwards, new method based on selecting best provider for specific condition was developed by considering atmospheric effects on power forecasting. It should be noted that the method is an engineering approach, not a new forecast model. In many cases, newly developed method has succeeded to outperform in comparison to results belonging to forecast providers. Hourly and daily wind power forecasting that have significant role in electricity market has been improved for selected wind farms by the help of an engineering approach used in this study. Same method is also implementable to another wind farm if required inputs exist.

Keywords: *Wind Power Forecasting, Atmospheric Effects, Temperature, Turbulence*

ÖZET

KISA VADELİ RÜZGAR ENERJİSİ ÜRETİM TAHMİNLERİNDE ATMOSFERİK ETKİLER

Rüzgar enerjisinden elektrik üretimi, fosil yakıtlardan elektrik üretimine kıyasla giderek tüm dünyada popüler hale gelmeye başlamıştır. Ancak rüzgardan elektrik üretimi, fosil yakıtlardan elektrik üretimi gibi gece ve gündüz sürekli olarak devam edememektedir. Rüzgarın doğası gereği üretim aralıklı olmaktadır ve bu durum rüzgar enerjisinden elektrik üretimini tahminini oldukça zorlaştırmaktadır. Literatürde rüzgar enerjisi tahmini için güçlü ve sahaya özgü birçok istatistiksel, fiziksel ve hibrid modeller bulunmaktadır. Bu tahmin modellerinin iyileştirmesi üzerine halen çalışmalar devam etmekte olup tahminlerin iyileştirilmesi farklı mühendislik yaklaşımlarına açıktır. Bu çalışmada da farklı atmosferik koşullara sahip üç farklı rüzgar santrali için farklı nümerik hava modellerine dayalı dört farklı tahmin sağlayıcısından elde edilen sonuçlar sıcaklık ve türbülans gibi atmosferik etkiler dikkate alınarak değerlendirilmiştir. Tahmin sağlayıcısı sonuçları sıcaklık ve türbülans göz önüne alınarak gruplandırılmış, sonuçlar türbinlerin gerçek üretim değerleriyle karşılaştırılmıştır. Belirlenen koşullarda en iyi tahmin sağlayıcısı seçilerek mühendislik yaklaşımıyla yeni bir metot geliştirilmiştir. Yeni geliştirilen metotun yeni bir tahmin modeli olmadığı, yalnızca yeni ve farklı bir mühendislik yaklaşımı olduğu unutulmamalıdır. Yeni metot yardımıyla elektrik piyasasında önemli bir role sahip saatlik ve günlük rüzgardan elektrik enerjisi üretim tahminleri seçilen rüzgar santralleri için iyileştirilmiştir. Geliştirilen metot eğer gerekli girdiler mevcut olursa bir başka rüzgar santralinde de uygulanabilir.

Anahtar Kelimeler ve Deyimler: *Rüzgar Enerjisi Tahmini, Atmosferik Etkiler, Sıcaklık, Türbülans*

TABLE OF CONTENTS

LIST OF FIGURES

Figure	**Page**

LIST OF TABLES

Table **Page**

CHAPTER 1

INTRODUCTION

Decrease in conventional energy systems, such as fossil fuel plant, thanks to opponent on air pollution and global warming has been caused development of renewable and sustainable energy sources like wind power. Wind power is one of the renewable sources such as hydro, solar and bioenergy power and these kinds of energy sources have limitless and cost-free fuel unlike fossil fuels. The fuel of the wind power is the blowing wind, which is not continuous during day and night. It always varies in time and for this reason predicting the wind is not so simple[1]. Power production of single wind turbine can be related to the wind speed and direction as well as site complexity, air density and time of the day. Power production balancing is a critical key component for this kind of unstable power plant to take care of power supply and demand. Therefore, accurate wind speed forecasting is very important for reliability and efficiency of power production of the wind farm in operation[1]. In addition, accurate power forecasting of the wind farms is necessary for maintenance planning during less windy days and cost-effective operation in power trading market. There are various methods in literature in order to predict wind speed and wind power accurately. These can be categorized into three methods; statistical (time series based), physical (numeric weather models), and hybrid methods (combined models). Accuracy of the models changes based on time-scale and topography of wind farm site[2]. These can be also categorized based on time-horizon of the forecasting, which are very short term, short term, medium term, and long term. Very short-term forecasting (a few minutes to 30 minutes) is useful for turbine control purpose and real-time grid operation. Short term and medium-term (one hour to one day) forecasting can be also used for load dispatch planning and energy trading in electricity market while long-term forecasting (up to 1 month) can be used for maintenance planning and operation management[3]. In order to acquire maximum profit on electricity market, operator company of the wind farms are especially focused on short-term power forecasting.

The forecasting methods can forecast the wind speed and corresponding power production accurately. Nevertheless, there may be significant amount of deviation in wind farm power production if the wind farm is exposed to atmospheric icing and turbulence.

In this case, the forecasting the power production of the wind farm is really challenging. To overcome this problem, forecasting results acquired by different models can be categorized based on icing phenomena that can be related to site temperature, and site turbulence intensity that can be related to site wind speed. Thus, different models can be used or preferred depending upon atmospheric conditions. For instance, there are two forecast models from different forecast providers belonging to specific wind farm. One can provide acceptable results during winter times that have icing conditions, other can provide satisfactory results in lack of icing phenomenon. Depending upon this selective criterion related to atmospheric conditions on the forecast model, some improvement on power forecasting can be achieved. In this study, four different short-term forecast models based on numerical weather prediction (NWP) model belonging to three wind farms that have different topography and site conditions are included to investigate deviation in power forecasting due to atmospheric effects. By considering atmospheric conditions mentioned above, some improvement in comparison to the forecast models has been made. It should be noted that this methodology is an engineering approach, which is not newly developed forecasting model. In following sections, literature review regarding short term power forecasting and its applications, data and methodology included in this study, information about wind farm site and power forecasting providers have been shared in detail.

1.1. Literature Review

Power supply and demand must be balanced at any time in electricity grid. Unbalanced electricity grid could create deviation in power quality. Unlike conventional energy sources, wind is non-dispatchable source of energy, which means wind power production cannot be decreased or increased based on power supply and demand. Fluctuations in wind power can be observed with seasonal and daily as well as short-term variations. If the penetration of wind power production increases, fluctuations in electricity grid due to intermittent nature of the wind will increase[4]. Transmission system operator (TSO) are in charge of balancing power supply and demand in electricity grid. There are two main mechanisms in electricity market. First mechanism is day-ahead market, it is spot market where independent power producers (IPPs) submit quantity of electricity that will be produced for each hour of upcoming day. Electricity spot price for each hour of following day based on various bids is determined by the help of an auction system. Second mechanism is balancing power production. It is controlled by TSO and intra-day market is constituted to balance lack and surplus of power production, which are mainly caused by failures of power plants and deviation in short term power forecasting. By means of the intra-day market there is an additional option for the IPPs to prevent the imbalances in their portfolios and even upgrade their profits. TSO determines obligations regarding power supply and demand to balance grid operations. If any events by IPPs that damage balance on electricity grids are occurred, TSO imposes penalty to IPPs. It should be noted that penalties of negative and positive imbalances are not same amount of cost for IPPs. Therefore, short-term power forecasting, especially in wind power, is needed for reducing imbalance charges and penalties on day-ahead and intra-day market[5]. More accurate power forecasting is crucial because of the fact that small increase of power performance in electricity market is financially attractive to IPPs[6].

Statistical forecasting methods generally use Numerical Weather Prediction (NWP) data including wind speed, wind direction, temperature and wind power produced by the wind farms. These methods are specifically focused on linear and non-linear relationship among the variables such as wind speed, temperature, and power production. In order to create reliable statistical relationship, sufficient amount of historical data as input into the models are required. The historical data are imported to linear or non-linear model as training data[7]. After fine-tuning for model accuracy, desired variables can be predicted for specific time horizon, however if the time horizon increases, forecasting accuracy decreases correspondingly[8]. Autoregressive Integrated Moving Average (ARIMA) model is one of the linear univariate models, which means only one variable is included in the model unlike multivariate ones. ARIMA is the most well-known and useful model in order to predict the future based on historical time series. There are numerous applications regarding ARIMA model in the economics as well as engineering. The model has popularized in the 1970s, Box and Jenkins have used the model for time series analysis and forecasting. The model that also known as Box-Jenkins method can be used for any specific cases such as prediction of power generation, stock market demand etc[9]. Indeed, there are many ARIMA models depending upon purpose of the study in the literature such as ARIMA with exogenous variables (ARIMAX) for multivariate time series, seasonal ARIMA (SARIMA) for seasonal data modelling[10]. According to simplest non-seasonal ARIMA model, variables to be predicted are considered to be equal to linear functions of past variables and their random errors. It includes two main components, which are AR and MA part. AR part indicates order of autoregression, while MA part shows order of moving average. Differencing part is also identified as I, which transforms the non-stationary time series to stationary ones[11]. ARIMA structure is robust and easy to learn and put into practice due low computational time in short-term wind power forecasting however, these models do not provide sufficient forecasting performance if the time series are non-stationary, which means that mean, variance and covariance of the time series change over time[12]. Gallego et. al focused on AR model with local wind speed and direction measurements and demonstrated that importance of local measurement on power forecasting accuracy for Horns Rev Wind Farm in Denmark[13]. Duran et. al used AR model for different Spanish wind farms with wind speed as exogenous variable in addition to wind power variable and demonstrated that short-term wind power forecasting based on ARX model provides lower errors compared to conventional AR model[14].

Artificial neural networks (ANN) are one of the well-known forecasting models based upon biological structure of brain. These models can handle complex non-linear relationships between the input and output variables. Unlike ARIMA model, these are also compatible for working with noisy, non-stationary and incomplete datasets. There are many neural networks in the literature, however feedforward and feedback artificial neural networks are most popular models for time series forecasting[10]. The neural networks are generally comprised of three layers, which are named as bottom, intermediate and top. Input or predictor variables take part in bottom layer, while output or forecast variables include in the top layer. In addition, hidden neurons into ANN model can be included in the intermediate layer. Unlike feedback neural networks, there are no loops among the neurons in the feedforward model. In other words, the feedforward model has only one direction data movement. The ANN model can be implemented by the means of weights on the input variables, and these are calculated by using learning algorithms, which uses cost function like mean squared error (MSE). Backpropagation algorithm is one of the learning algorithms. It is based on backpropagation errors, which adjusts weights in the hidden layer by decreasing the error between observed and predicted values. If there is no hidden layer, simple neural network model resembles linear regression. After adding hidden layers, the model will be transformed to the non-linear form. There can be many hidden neurons in the hidden layer. Every neurons of the each layer provides inputs to the neurons in the next layer, that means inputs in the next layer are equivalent to the outputs in the previous layer[15]. Based on the literature, ANN models provide satisfactory results in short-power forecasting[2]. However, all of these models are developed and validated for site specific purpose only. Validated model for specific site may not provide accurate power forecasting for another site. For this reason, different form of ANN models, for instance back-propagation neural network (BPNN)[16] and long-short-term memory (LSTM)[17], can be used to increase accuracy of the forecasting that varies from site to site. Pelletier et. al developed multi-stage ANN with six inputs, which are wind speed, air density, turbulence intensity, wind shear, wind direction and yaw error and improved short-term wind power forecasting for 140 wind turbines in Nordic compared to other ANN models, which are parametric, non-parametric and discrete models[6]. Bilal et. al showed that using both temperature and wind speed parameter as input to ANN would improve the power forecasting compared to ANN that includes only wind speed parameter. The study also shows that additional atmospheric variables, such as air density, rainfall etc., can help to improve accuracy of ANN models[18]. Chang worked on BPNN, back-propagation is one of the learning algorithm for ANN model, including only wind power output parameter and improved short term power forecasting for a single wind turbine in Taiwan compared to conventional ANN model[19]. According to Singh et. al, there are many factors that affect wind turbine power output like terrain, air density, wind shear, wind speed and direction. A feed forward neural network model based on air density, wind speed and direction parameters was developed and the model has outperformed traditional ANN models. The study also demonstrates that wind direction has less impact on wind power forecasting in comparison to wind speed[20].

Physical forecasting methods are basically explained as downscaling of numerical weather prediction (NWP) data. This mechanism requires site boundary, roughness, digital elevation and obstacle map of the boundary and historical reanalysis data like wind speed, temperature, pressure. Depending upon these variables, atmosphere can be modeled by using complex mathematical equations, which are so time-consuming compared to statistical methods mentioned before[21]. Accurate wind speed forecasting can be implemented by means of numerical weather prediction models and then turbine power curve can be utilized to calculate power production of each turbine in a wind farm. However, in complex terrain, NWP model can cause some deviation in wind speed forecasting, which leads to errors in wind power forecasting. It should be noted that using manufacturing power curve instead of site-specific power curve to calculate power production leads to failure in power forecasting[22]. There are many site-specific NWP models in the literature such as regionally developed Weather Research and Forecasting (WRF) model, which gives satisfactory results in short-term power forecasting like statistical ANN models. Focken et al. developed physical model to forecast short-term wind power up to period of two days. Boundary layer of study was created by considering roughness, orography, and wake effect. Change on wind speed depending upon thermal effects at the atmosphere was also considered for wind speed forecasting at the hub height[23]. However, NWP models cannot consider site specific power production losses caused by environmental impacts such as icing, turbulence intensity. NWP models are generally developed for flat terrain like European sites. Even if NWP model in complex site, for instance Turkey, gives accurate wind speed forecasting, some power forecasting errors can be observed due to low site temperature and considerable turbulence. Therefore, power forecasting results acquired by NWP model should be improved and corrected by including statistical models such as ANN.

Hybrid methods consist of combination of different forecasting models, such as NWP model supported by statistical models or vice versa. Recently, many researchers and companies leading wind energy sector are oriented towards these kinds of hybrid forecasting methodologies in order to increase accuracy of short-term power forecasting because of the fact that any single forecasting model cannot overcome to provide accurate results for the specific site under today's conditions. It should be noted that combination of the forecasting models may not provide better results all the time. However, it should be preferred in order to minimize risk on power forecasting[22]. Lin et al. developed hybrid model, combining statistical methods, in order to reveal outliers of wind power forecasting. The model used wind speed, nacelle orientation, yaw error, blade pitch angle and nacelle temperature parameters from each turbine at an offshore wind farm in Scotland[24]. Lin and Liu also investigated importance of input variables on accuracy of wind power forecasting. According to the study, wind direction and air density parameters had less significance on wind power forecasting accuracy and blade pitch angle parameter had significant effect on the accuracy compared to wind speed and wind shear parameters[15]. De Giorgi et al. developed hybrid method, combining statistical and physical models, for a wind farm in Italy. Wind power, wind speed, pressure, temperature, and humidity variables were included to the model. Hybrid model provided better results compared to single ANN models. Based on the model, pressure and temperature variables provided by different NWP providers had significant importance on wind power forecasting[25].

Many hybrid methods on power forecasting have been improved by considering site-specific conditions. Most hybrid models that are developed site-specifically have been succeeded to provide accurate power forecasting, however it was not possible to model power production losses due to atmospheric phenomena. There are many root causes that affect wind power production losses, which are turbine and grid availability, power and temperature curtailment, maintenance and shut down of the turbines, atmospheric events like icing, turbulence, and extreme wind speed. It should be noted that atmospheric conditions have significant importance in relation to wind power forecasting however there are several parameters regarding atmospheric conditions that are stochastic and not controllable. According to study of Pieralli et al., significant amount of power production losses is caused by change in wind climate and turbine errors are responsible for only 6% of the power production losses. Based on the study, icing has considerable effect on power production losses among other variables. The study also says that changing in wind conditions has also greater effect, deviation in wind speed and wind direction affect wind power output and will cause power production losses. In addition to these two parameters, park position of wind turbines in a single wind farm creates power production losses[26]. Based on the study of Pieralli et al., temperature and wind speed variables should be considered in order for revealing power production losses and improving accuracy of wind power forecasting.

CHAPTER 2

WIND FARM SITE DESCRIPTION

Three operational wind farms in different locations of Turkey have been included in this study. These have been referred to in this study as Wind Farm A, B and C. Wind Farm A has started to produce electricity in June 2015 and its physical capacity is equal to 55.2 MWh. Wind Farm B has been installed in October 2015 and it has 33.9MWh physical capacity. Wind Farm C has produced electricity since September 2019 and its physical capacity is 91.1 MWh. All of the wind farms mentioned here had produced electricity with maximum physical capacity from their installation date to early 2020. After change on legislative regulations in 2020, their physical capacities have been restrict based on wind farm licenses. If the wind farm has exceeded its license capacity limit, operator would pay a penalty. In order to prevent these kinds of penalties, power curtailment methodology has been applied to the wind farms since 2020.

Table 2.1 Installed Capacity of Wind Farm A, B and C

Wind Farm	Physical Capacity (MWm)	Capacity Limit (MWe)
A	55.2	50.0
B	33.9	30.0
C	91.1	87.0

2.1. Wind Farm A

The wind farm A has been located in Mediterranean Region in middle of Turkey, approximately 90 km north from Mediterranean Sea. It consists of 16 Vestas V112 wind turbines at 84 m hub height, as shown in Figure 2.1 below. The wind farm site lies on open and bare land at an average elevation of approximately 1656 m. The general terrain can be described as simple based on the Ruggedness Index[27] (RIX) value of 4.2%, which means very low percentage fraction of slopes more than 30%(17º). The details belonging to the wind farm can be seen below in Table 2.2. Wind turbine information in the Wind Farm A has been provided by the investor company. Base elevation and air density at the turbine points were calculated to acquire additional environmental information regarding the site. 1 arc Shuttle Radar Topography Mission (SRTM) digital elevation model[28] was used to calculate RIX value and find base elevation of the turbines. Air density was also calculated by the help of engineering method, which enables us to make the calculation by using only temperature and elevation data. Formula of the engineering method can be seen below. p_0, L, t_o, g, Ma and R indicate standard sea level pressure, vertical temperature gradient, standard sea level temperature, gravitation constant, molarity of air, gas constant, respectively[29].

$$p(z) = p_0[(1 - Lz/t_o)]^{\frac{gMa}{RL}} \tag{2.1}$$

Power production of Wind Farm A has been struggling due to turbulence and icing phenomenon. The site is included in the study to improve short-term wind power forecasting considering these two phenomena.

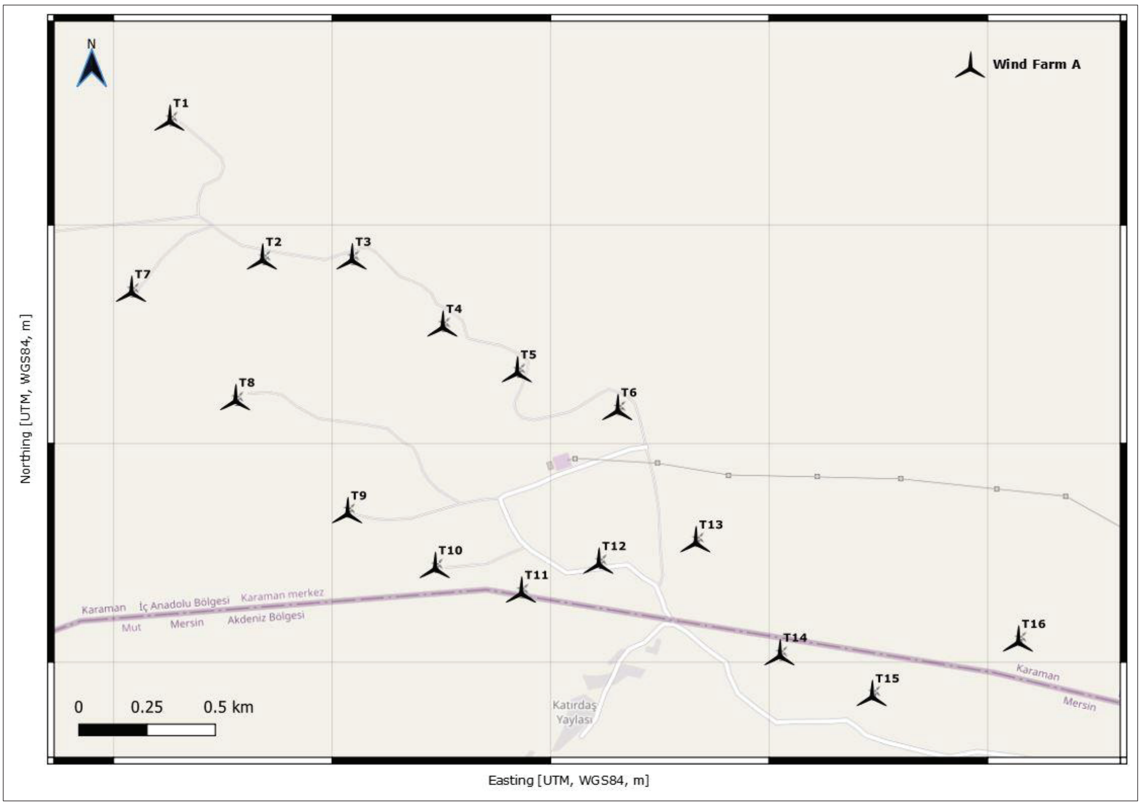

Figure 2.1 Wind Farm A Layout

Table 2.2 Wind Farm A General Information

Turbine	Manufacturer	Rated Power (MW)	Rotor Diameter (m)	Hub Height (m)	Base Elevation (m)	Air density (kg/m^3)
T1	Vestas	3.45	112	84	1640	1.025

cont. on next page

cont. of table 2.2

T2	Vestas	3.45	112	84	1644	1.011
T3	Vestas	3.45	112	84	1637	1.026
T4	Vestas	3.45	112	84	1659	1.013
T5	Vestas	3.45	112	84	1643	1.025
T6	Vestas	3.45	112	84	1659	1.010
T7	Vestas	3.45	112	84	1616	1.026
T8	Vestas	3.45	112	84	1625	1.017
T9	Vestas	3.45	112	84	1653	1.014
T10	Vestas	3.45	112	84	1675	1.003
T11	Vestas	3.45	112	84	1652	1.019
T12	Vestas	3.45	112	84	1649	1.013
T13	Vestas	3.45	112	84	1655	1.004
T14	Vestas	3.45	112	84	1701	1.006
T15	Vestas	3.45	112	84	1700	1.009
T16	Vestas	3.45	112	84	1693	1.012

2.2. Wind Farm B

The wind farm B has been situated in Aegean Region in West of Turkey, approximately 110 km east from Aegean Sea. 10 wind turbines have been installed, which are Vestas V112 at the hub height of 84 m, in the wind farm. Its layout can be seen in the Figure 2.2 below. The wind farm site is located on a very complex terrain with variable topography at an average elevation of approximately 1076 m. Ground cover on the site is ranging from single bushes and trees to dense forestry. The whole site area is covered by vegetation with varying heights, density, and maturity. RIX value of the site is equal to 38.9%, which explains very high percentage fraction of slopes more than 30%(17º). Power production of Wind Farm B has been struggling due to turbulence and icing phenomenon caused by complexity and high altitude of the site. The site is included in the study to improve short-term wind power forecasting considering these two phenomena. The detailed information of the wind farm provided by investor company is shown in both Figure 2.2 and Table 2.3 below. As it has been mentioned earlier, additional environmental information regarding the site was calculated by means of engineering method and 1-arc SRTM digital elevation model.

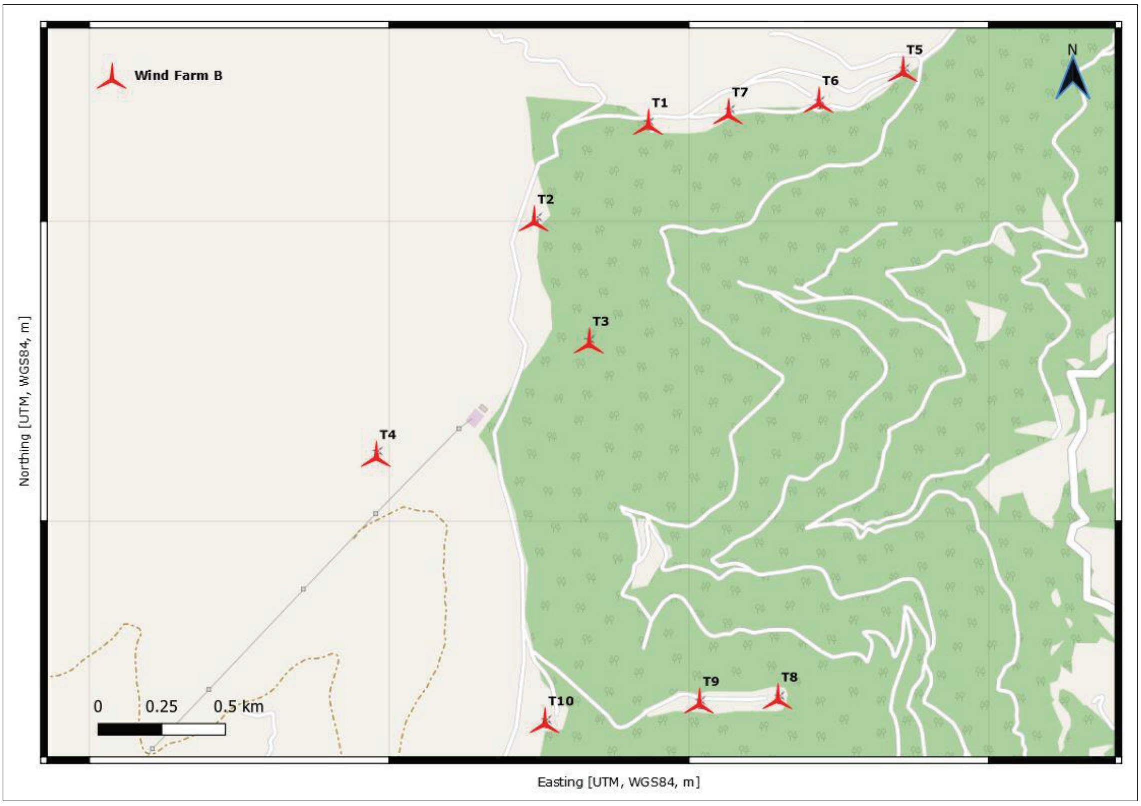

Figure 2.2 Wind Farm B Layout

Table 2.3 Wind Farm B General Information

Turbine	Manufacturer	Rated Power (MW)	Rotor Diameter (m)	Hub Height (m)	Base Elevation (m)	Air Density (kg/m^3)
T1	Vestas	3.45	112	84	1102	1.081
T2	Vestas	3.45	112	84	1135	1.072
T3	Vestas	3.45	112	84	1129	1.077
T4	Vestas	3.45	112	84	1060	1.084
T5	Vestas	3.45	112	84	1032	1.093
T6	Vestas	3.45	112	84	1074	1.087
T7	Vestas	3.3	112	84	1104	1.089
T8	Vestas	3.3	112	84	1013	1.093
T9	Vestas	3.3	112	84	1015	1.091
T10	Vestas	3.3	112	84	1100	1.080

2.3. Wind Farm C

The wind farm C has been located in Marmara Region in West of Turkey, approximately 3 km South from Marmara Sea. The wind farm has hybrid layout, which consists of Vestas V112 and Vestas V90. In total, there are 29 wind turbines at the different hub heights of 80 m and 84 m. The site lies on a wide area in the direction of east and west. Its layout can be seen in Figure 2.3 below. The wind farm site, at an average elevation of 527 m, is covered by forestry with varying heights, density, and maturity. RIX value of the site was calculated as 7.3%, which indicates low percentage fraction of slopes more than 30%(17º). Wind Farm C is situated on non-complex, and its location is relatively close to coastal area so, it has not been struggling due to turbulence and icing phenomenon unlike Wind Farm A and B. The site, under influence of lack of icing and turbulence, is included in the study to compare improvement of short-term wind power forecasting of Wind Farm A and B. General information of Wind Farm C provided by investor company are shared in below Table 2.4. Additional information regarding the base elevation and air density of the turbines can also be found in Table 2.4.

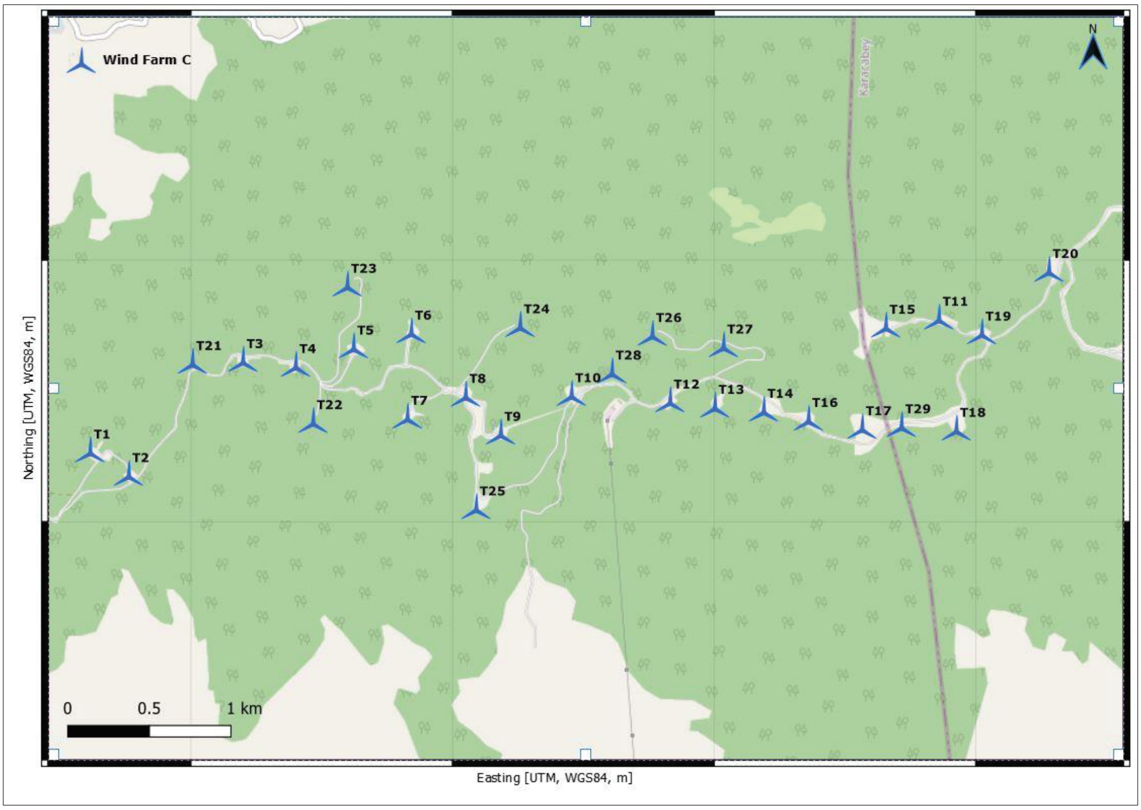

Figure 2.3 Wind Farm C Layout

Table 2.4 Wind Farm C General Information

Turbine	Manufacturer	Rated Power (MW)	Rotor Diameter (m)	Hub Height (m)	Base Elevation (m)	Air density (kg/m3)
T1	Vestas	3	90	80	466	1.167

cont. on next page

cont. of table 2.4

T2	Vestas	3	90	80	471	1.153
T3	Vestas	3	90	80	419	1.168
T4	Vestas	3	90	80	442	1.170
T5	Vestas	3	90	80	479	1.157
T6	Vestas	3	90	80	528	1.158
T7	Vestas	3	90	80	540	1.150
T8	Vestas	3	90	80	523	1.161
T9	Vestas	3	90	80	536	1.152
T10	Vestas	3	90	80	485	1.166
T11	Vestas	3	90	80	691	1.141
T12	Vestas	3	90	80	499	1.161
T13	Vestas	3	90	80	522	1.156
T14	Vestas	3	90	80	538	1.156
T15	Vestas	3	90	80	648	1.140
T16	Vestas	3	90	80	566	1.155
T17	Vestas	3	90	80	610	1.145
T18	Vestas	3	90	80	663	1.144
T19	Vestas	3	90	80	700	1.141
T20	Vestas	3	90	80	678	1.145
T21	Vestas	3.45	112	84	393	1.167
T22	Vestas	3.45	112	84	456	1.158
T23	Vestas	3.45	112	84	423	1.159
T24	Vestas	3.45	112	84	467	1.161
T25	Vestas	3.45	112	84	498	1.155
T26	Vestas	3.45	112	84	452	1.163
T27	Vestas	3.45	112	84	473	1.156
T28	Vestas	3.45	112	84	490	1.153
T29	Vestas	3.45	112	84	615	1.139

CHAPTER 3

WIND FARM SCADA

Approximately 25-30% of power production cost of a wind farm is originated from operation and maintenance of wind turbines. In order to decrease this kind of cost, monitoring and fault detection systems are quite necessary for the wind farms. For instance, vibration analysis, lubrication analysis and strain measurement should be followed up by the help of condition monitoring systems during lifetime of the wind turbines. Supervisory control and data acquisition system (SCADA) of the wind farm could help to monitor health condition of the wind farm by collecting a large quantity of measurements from the wind turbines. SCADA system includes many various parameters such as bearing and oil temperature, wind speed and direction, power output, pitch angle and rotor speed. By taking these parameters into account, several analyses such as power curve and power production loss analysis can be performed[30].

Operational wind farm data belonging to three wind farms, which named here as Wind Farm A, B and C, have been provided by the investor, which is a private company that leads to renewable energy sector. As it has been mentioned earlier, wind turbine SCADA data include various parameters regarding wind turbine operation. Parameters that were used in this study has been selected depending upon purpose of the study.

At the beginning of the study, turbine-based SCADA data between the years of 2019 and 2020 have been shared. The SCADA data that have 10-minute time steps, include 9 different channels regarding condition of the wind turbines. These can be categorized as; time stamp with 10 minutes interval; environmental conditions such as maximum, minimum, average and standard deviation of wind speed, average wind direction and ambient temperature, power production with availability time count as well. In addition to these parameters, some additional channels were created to implement this study. Reference power for comparing real power production on SCADA data, air density for calculating reference power and turbulence intensity for detecting unexpected losses in power production were calculated and added as new channels on the SCADA data of the wind turbines.

Data regarding wind speed conditions are measured by anemometer and wind vane on wind turbine nacelle, so there is some decrease in wind speed. To calculate turbulence intensity, standard deviation of wind speed is divided by average of the wind speed. In descriptive statistic, ratio of standard deviation to the mean is known as coefficient of variation (CV) or relative standard deviation (RSD), which is measure of dispersion or variability of data around its mean[31]. Time series of turbulence intensity can be created by using this formula mentioned.

Site specific air density can be calculated by means of engineering method[29]. This method needs only ambient temperature and base elevation of the turbine as input to calculate site-specific air density.

Reference power can be calculated by using air density of the site, average wind speed of the turbine and wind turbine manufacturer's power curve. The reference power might be helpful to observe possible issues on wind turbine power production. If the site air density, average wind speed and turbine power curve exist, the reference power can be acquired easefully by using look-up table of the turbine power curve that includes power production at various wind speed and air density. This method is one of the easiest ways to calculate reference power for comparing real power production of the wind turbines. However, it leads to high uncertainties because of the fact that measurement of the average wind speed is affected by wind turbine blades adversely. It should be also noted that look-up table of the turbine power curve provided by manufacturer will be different from site to site. For this reason, including site specific power curve depending upon SCADA data let the reference power calculation more accurate. In this study, reference power was calculated based on site specific power curve of each wind turbines. To make the calculation more reliable, one of the ready-to-use Python tools was used to acquire site specific power curve individually for all wind turbines.

Data channels on both collected by wind turbine SCADA system and calculated manually are included in this study. SCADA channel labels belonging to a wind turbine in one of the wind farms is shown Table 3.1 below.

Table 3.1 SCADA Data Channels

Abbreviation	Label	Unit	Type
PCTimeStamp	Date/Time	10 minutes	Scada
Amb_WindSpeed_Max	Max Wind Speed	m/s	Scada
Amb_WindSpeed_Min	Min Wind Speed	m/s	Scada

cont. on next page

cont. of table 3.1

Amb_WindSpeed_Avg	Average Wind Speed	m/s	Scada
Amb_WindSpeed_Std	Standard Deviation of Wind Speed	m/s	Scada
Amb_WindDir_Abs_Avg	Wind Direction	Degree	Scada
Amb_Temp_Avg	Ambient Temperature	Celsius	Scada
Grd_Prod_Pwr_Avg	Turbine Power Production/Consumption	kWh	Scada
HCnt_Avg_Tot	Time Count	seconds	Scada
Power	Turbine Power Production	kWh	Calculated
Reference	Turbine Reference Power	kWh	Calculated
TI	Turbulence Intensity	-	Calculated

10 minutes SCADA data, covering also calculated channels, have been converted as hourly in order to compare hourly forecasting of power production of the wind farms. Due to site-based power curtailment at the beginning of 2020, only SCADA data belonging to the year of 2019 could be analyzed in this study. Power consumption of the wind turbines has been neglected in this study, so the negative values in the SCADA data that indicate the power consumption have been changed as zero, which means no electricity or power production at the wind turbines. Reference power production has been calculated in order for detecting how the wind turbines are far from normal operation, which means that there are no power production losses during operation. To detect anomalies from normal operation, reference and SCADA power production were compared. Reference power production has been also compared with the forecasting of power production. Power production of the wind turbines has been analyzed based on ambient temperature and turbulence intensity in order to reveal possible deviation of forecasting of power production. Apart from the power production, metered data belonging to the wind farms have been investigated. There will be some difference between the power production and metered data because of electrical losses on cabling during electricity transferring from wind farm site to transformer station. Table 3.2 shows that total power and metered production of the wind farms in 2019. Power production data from SCADA and power forecasting data from forecast providers have been examined and investigated how power forecasting can be improved.

Table 3.2 Total Power Production of Wind Farm A, B and C*

	Total Power Production [MWh/a]	**Total Metered Production [MWh/a]**	**Electrical Loss [%]**
Wind Farm A	134012.7	132518.1	1.1%
Wind Farm B	90501.6	89094.6	1.6%
Wind Farm C	280118.6	276150.8	1.4%

* Total power and metered production are calculated by taking missing SCADA data into consideration.

CHAPTER 4

WIND FARM POWER FORECASTING

There are many companies related to wind power forecasting in both local and global. They have various methodology in order for providing wind farm power forecasting in microscale. However, these companies have common purpose for improving their numerical weather prediction (NWP) model and providing more accurate forecast by using reanalysis data, which is in mesoscale. During downscaling process from mesoscale to microscale, each company has unique process to calculate wind farm-based power forecasting. Turbine based power forecasting is also thinkable however, this methodology will increase deviation of power forecasting. Thus, the companies generally give preference to wind farm-based forecasting. In this study, four different forecasting providers have been included to compare power production of Wind Farm A, B and C and detect possible causes in deviation of power forecasting. Power forecasting data have been shared by the owner company of the wind farms and forecasting providers have been named as Provider 1, 2, 3 and 4, respectively.

Some forecast providers only calculate wind farm power production, some calculate also meteorological parameters in addition to power production belonging to the wind farms. Forecast provider 1 has hourly power production forecasting belonging to the wind farms. Apart from the power production, hourly relative humidity, pressure, temperature, wind speed and direction forecasting for Wind Farm A, B and C are provided by forecast provider 1. Forecast provider 3 just like previous one contains meteorological parameters, which are relative humidity, pressure, temperature, wind speed and direction. It also provides hourly power production for the wind farms. Forecast provider 2 and 4 have only hourly power production for the wind farms.

Detailed information regarding forecast providers can be found in Table 4.1 below. As stated previously, 10-min SCADA data have been converted to hourly values in order for comparing forecasting values and only SCADA data in 2019 have been included in this study because of the fact that dispatch quantity to the grid have been limited to the licensed capacity since January 2020 by the help of wind farm control units.

Control units of Wind Farm A, B and C enable power limited operation mode when it is necessary, which can be called as wind farm level curtailment.

Table 4.1 Forecasting Data Channels

Forecast Provider-1	**Forecast Provider-2**	**Forecast Provider-3**	**Forecast Provider-4**	**Time Scale**	**Explanation**
Provider-1 [MWh]	Provider-2 [MWh]	Provider-3 [MWh]	Provider-4 [MWh]	Hourly	Power Forecasting
Provider-1 WS	-	Provider-2 WS	-	Hourly	Wind Speed Forecasting
Provider-1 WD	-	Provider-2 WD	-	Hourly	Wind Direction Forecasting
Provider-1 T	-	Provider-2 T	-	Hourly	Temperature Forecasting
Provider-1 RH	-	Provider-2 RH	-	Hourly	Relative Humidity Forecasting
Provider-1 P	-	Provider-2 P	-	Hourly	Pressure Forecasting

Forecasting of total power production in 2019 provided by forecast provider 1 and provider 2 as well as forecast provider 3 and 4 can be seen below Table 4.2.

Table 4.2 Forecasting of Total Power Production of Wind Farm A, B and C*

	Wind Farm A [MWh/a]	**Wind Farm B [MWh/a]**	**Wind Farm C [MWh/a]**
Forecast Provider 1	117529.7	103430.1	283115.0
Forecast Provider 2	132353.2	98370.2	284940.9
Forecast Provider 3	134087.8	92453.3	282785.8
Forecast Provider 4	125286.4	96946.6	281721.0
SCADA Power Production	134012.7	90501.6	280118.6

* Forecasting of total power production are calculated based on concurrent SCADA data, excluding missing period.

Power production forecasting of the wind farms are volatile due to nature of the wind and unstable environmental circumstances. Although each forecast providers use numerical weather prediction (NWP) method to forecast power production of next hours, days and months, these have different models of NWP that are able to model the atmosphere in different conditions. Thus, some forecast providers have high deviation from actual power production, others have less deviation in power production forecasting. In order to investigate reasons of the deviation from actual power production, some meteorological parameters such as temperature, wind speed variables are needed.

4.1. Calculating Reference Power Production

Reference power production of the turbines were calculated based on calibrated wind speed and reference power curve. Reference power curves were acquired by the help of Python tool developed by international expert group IEA Wind Task 19[32]. Based on IEC 61400-12-1 Power Curve Measurements[33], wind measurement data from a met mast are required for detailed power production calculations, because of the fact that wind measurements of nacelle anemometer are disturbed by wind turbine blades. However, it is not possible to install the met mast for validating power curve of each wind turbines of a wind farm. In that case, wind measurements on nacelle can be calibrated by using site air density and air pressure. According to ISO 2533 Standard Atmosphere[34], calibration of nacelle wind speed can be applied as follows[32];

$$w_{site} = w_{std} \times \left(\frac{\rho_{std}}{\rho_{site}}\right)^{\frac{1}{3}} = w_{std} \times \left(\frac{\frac{P_{std}}{T_{std}}}{\frac{P_{site}}{T_{site}}}\right)^{\frac{1}{3}} \tag{5.1}$$

$$w_{\text{site}} = w_{\text{std}} \times \left(\frac{T_{std}}{T_{site}}\left(1 - 2.25577 \times 10^{-5)} \times h\right)^{5.25588}\right)^{\frac{1}{3}} \tag{5.2}$$

w_{site}, indicates the calibrated wind speed
w_{std}, indicates the nacelle wind speed
T_{site}, is the nacelle temperature
P_{std}, represents the standard air pressure at sea level (101325 Pa)
T_{std}, represents the standard temperature of 15 °C (288.15 K)
h, is site elevation in meters.

SCADA data generally does not include air pressure measurement so, static air pressure based on site elevation above sea level can be calculated or air pressure of high-resolution weather model can be used but measured air pressure is preferred one. After calibration of nacelle wind speeds, power production data should be filtered based on the rule, which is disregarding the power production data if nacelle temperature less than 3°C. The purpose of this rule is to eliminate power production in icing operation due to frosted turbine blades. In addition to that, normal operation is identified in the tool as $P_{min} > 0.005 \times P_{rated}$ and $P_{mean} > 0.01 \times P_{rated}$. According to filtered power production in normal operation and calibrated wind speed bin, power curve of each the turbines are calculated. SCADA data format should be edited as readable input by the Python tool and some additional information, which include rated power of the power curve, temperature limit of reference power, power level filter, stop limit multiplier, turbine elevation, bin size, maximum and minimum of wind speed should be inserted to .ini config file handled by Python. These essential parameters that are included in the config file as seen Table 4.3, Table 4.4 and Table 4.5. By means of the tool, reference power production was calculated in a short time for each turbine of Wind Farm A, B and C.

Table 4.3 Data Structure of Example Config

Data Structure Config		**Unit**
Timestamp index	0	-
Wind speed index	1	-
Wind direction index	2	-
Temperature index	3	-
Power index	4	-
Rated power	3300	kWh
Site elevation	1642	m

The index values above represent the column of the SCADA data. Rated power is equal to maximum power output from the selected wind turbine. Site elevation above sea level indicates altitude of the selected wind turbine.

Table 4.4 Wind Speed Binning in Example Config

Binning Config		**Unit**
Minimum wind speed	3	m/s
Maximum wind speed	25	m/s
Wind speed bin size	0.5	m/s

Parameters in both binning and filtering config are needed for reference power curve output and reference power in normal operation. If the $P_{min} > 0.005 \times P_{rated}$ and $P_{mean} > 0.01 \times P_{rated}$ then power production will be in normal operation mode. Reference temperature value are needed for disregarding power production in icing condition. Details can be found in Table 4.5.

Table 4.5 Data Filtering in Example Config

Filtering Config		**Unit**
Stop limit multiplier	0.005	-
Power level filter	0.01	-
Reference temperature	3	°C

An exemplary reference power curve calculation of 3.3 MW turbine by the Python tool could not share in Table 4.6 without permission of turbine manufacturer because, it is site specific power curve, which means that strictly confidential. Corresponding power texts represent the turbine power with corresponding wind speeds. The tool used in this study can be modified easefully depending upon type of the wind turbine at different wind farm site.

Table 4.6 An Example of Reference Power Curve Calculated by Python Tool

Wind Speed [m/s]	**Power [kWh]**
< 3	Corresponding Power*
3	Corresponding Power*
3.5	Corresponding Power*
4	Corresponding Power*
4.5	Corresponding Power*
5	Corresponding Power*
5.5	Corresponding Power*
6	Corresponding Power*
6.5	Corresponding Power*
7	Corresponding Power*
7.5	Corresponding Power*
8	Corresponding Power*
8.5	Corresponding Power*
9	Corresponding Power*
9.5	Corresponding Power*
10	Corresponding Power*
10.5	Corresponding Power*
11	Corresponding Power*
11.5	Corresponding Power*
12	Corresponding Power*
12.5	Corresponding Power*
13	Corresponding Power*
13.5	Corresponding Power*
14	Corresponding Power*
14.5	Corresponding Power*
15	Corresponding Power*
15.5	Corresponding Power*
16	Corresponding Power*
16.5	Corresponding Power*
17	Corresponding Power*
17.5	Corresponding Power*
18	Corresponding Power*

cont. on next page

cont. of table 4.6

18.5	Corresponding Power*
19	Corresponding Power*
19.5	Corresponding Power*
20	Corresponding Power*
20.5	Corresponding Power*
21	Corresponding Power*
21.5	Corresponding Power*
22	Corresponding Power*
22.5	Corresponding Power*
23	Corresponding Power*
23.5	Corresponding Power*
24	Corresponding Power*
24.5	Corresponding Power*

*site-specific power values are confidential.

After reference power production calculation, 10-min SCADA data have been converted to hourly values in order for creating concurrent period with forecasting data. Some negative values on grid power production channel due to internal consumption of the wind turbines have been replaced as zero. This is because forecast providers do not take wind farm consumption into account during process of calculating power production forecasting, this kind of consumption cannot be related to power production of the turbines.

4.2. Turbine Based Data to Wind Farm Scale

For Wind Farm A, reference power curve and reference power production values of 16 wind turbines were calculated based on 10-min period. After disregarding the negative values in grid power production, SCADA data of Wind Farm A were converted from 10-min period to hourly period. To acquire wind farm-based data, grid power production and reference power production of 16 wind turbines in total were summed and divided by a thousand for conversion of kWh to MWh. Temperature, wind speed, turbulence intensity and time count data of each turbines were simply averaged, and wind direction data of each turbines were also averaged by vectoral. Thus, power production of grid and reference, temperature, speed and direction of wind and turbulence intensity data were created at wind farm scale between dates of 01.01.2019 and 01.01.2020.

Same methodology for both 10 wind turbines of Wind Farm B and 29 wind turbines of Wind Farm C were applied to create hourly wind farm-based SCADA data. An exemplary data from Wind Farm A is shown in Table 4.7 below.

Table 4.7 An Example of Wind Farm Based Scada Data*

Date/Time	Turbine Power Production [MWh]	Turbine Reference Power [MWh]	Availability – Time Count	Average Ambient Temper ature	Average Wind Speed	Average Wind Direction	Average Turbulence Intensity
3/28/2019 0:00	4.72	7.26	3599.90	8.11	6.32	10.05	0.10
3/28/2019 1:00	5.00	6.85	3600.00	7.63	6.26	9.25	0.09
3/28/2019 2:00	3.02	3.75	3600.00	8.37	5.30	28.79	0.09
3/28/2019 3:00	4.08	5.24	3600.00	8.15	5.86	31.17	0.07
3/28/2019 4:00	7.12	8.09	3599.90	7.12	6.69	23.99	0.07
3/28/2019 5:00	11.11	13.07	3600.00	6.34	7.89	26.21	0.06
3/28/2019 6:00	10.34	11.74	3600.00	5.98	7.60	25.41	0.08
3/28/2019 7:00	12.39	13.98	3600.00	5.77	8.05	26.95	0.08
3/28/2019 8:00	18.94	21.54	3600.00	5.83	9.70	27.10	0.09
3/28/2019 9:00	18.68	21.10	3600.00	6.18	9.53	25.55	0.12
3/28/2019 10:00	11.50	12.14	3600.00	7.12	7.60	26.21	0.12
3/28/2019 11:00	7.07	7.84	3600.00	7.62	6.62	25.46	0.10
3/28/2019 12:00	12.16	13.30	3600.00	7.11	7.91	27.14	0.10

* Turbine-based SCADA data were converted to wind farm-based SCADA data in order for comparing wind farm-based power forecasting.

In addition to these data above, hourly power forecasting data channels were added as seen in

Table 4.8 to the wind farm-based SCADA data. Thus, comparison between the SCADA data and forecasting data were applicable to detect deviation in forecasting of power production. Hourly wind farm-based power production should be compared with forecasting of power production however, metered production data can also be included for this comparison. In this study, the wind farm-based power production was compared with forecasting power production as any forecast providers have not utilized this kind of metered data input during process of power forecasting. It should be noted that metered production data include electrical losses of the wind farm. Difference between the metered and SCADA production shows that electrical losses due to cabling in the wind farm. Electrical efficiency of Wind Farm A, B and C were calculated depending upon 1-year concurrent period of metered and SCADA production data. Results can be seen in Table 3.2 in detail.

Table 4.8 An Example of Wind Farm Based Forecasting Data

Date/Time	Forecast Provider 1 [MWh]	Forecast Provider 2 [MWh]	Forecast Provider 3 [MWh]	Forecast Provider 4 [MWh]
3/28/2019 0:00	11.79	11.81	10.96	10.34
3/28/2019 1:00	8.58	10.71	8.94	8.48
3/28/2019 2:00	5.93	9.64	8.22	6.84
3/28/2019 3:00	4.32	9.15	7.78	5.77
3/28/2019 4:00	4.06	9.10	7.81	6.59
3/28/2019 5:00	4.40	9.32	8.93	7.67
3/28/2019 6:00	4.51	9.90	10.72	8.44
3/28/2019 7:00	5.58	10.40	12.17	9.55
3/28/2019 8:00	5.74	10.64	12.43	9.71
3/28/2019 9:00	6.07	10.15	11.86	9.48
3/28/2019 10:00	5.47	9.77	10.52	9.06
3/28/2019 11:00	4.71	9.91	8.73	8.33
3/28/2019 12:00	4.13	10.60	7.97	8.26

CHAPTER 5

DATA AND METHODOLOGY

Scada and power production forecasting data have been shared by the owner company of the wind farms. Generally, power production forecasting has relatively high deviation from actual power production of the wind farms, so it is considered in this study how to decrease the deviation by including four different power forecasting data from forecast providers. Atmospheric effects on power production forecasting have been investigated by means of temperature and turbulence intensity variables.

5.1. Wind Farm Power Production Based on Temperature

Wind Farm A and B has been situated at higher altitude compared to elevation of Wind Farm C so, icing events during wind farm operation has been expected, especially in winter times. By taking the altitude of Wind Farm C into account, its icing loss is expected to be less than Wind Farm A and B however, it cannot be considered that there is no icing loss in the wind farm. If the site temperature is low such as close to 0°C or below 0°C, icing phenomena on turbine blades can be occurred however, there is no exact criteria of temperature limit for icing formation, and this criterion is also depending upon relative humidity that can change from site to site. If the relative humidity at the site is not high enough, icing phenomena cannot be occurred even if the temperature is below 0°C. Icing formation will affect normal operation of the wind farm adversely. Ageing on the turbines due to extreme ice loads and decrease in power production can be observed. For that reason, it is really challenging to provide accurate forecasting of power production. To reveal temperature effect on forecasting of power production, each forecasting data provided by the forecast providers are grouped by temperature of the wind farm site and some iterative selection on forecasting of power production can be applied in order for getting more accurate forecasting for the wind farms.

In order for selecting best forecasting power production in the range of specific temperature, SCADA and forecasting power production should be grouped by the temperature bin such as 1°C, 2°C or 3°C. However, ambient temperature data from

SCADA cannot be used for this kind of process. Ambient temperature from SCADA is real-time measurement so, temperature forecasting data are needed for grouping of forecasting power production. In this case, temperature forecasting data from one of the forecast providers can be included for grouping process. Unfortunately, meteorological parameters like temperature have been provided by the forecast provider 1 and 3 only. There seems to be two choices in total however, temperature forecasting data should be selected and preferred based on relationship with real-time ambient temperature data. To examine relationship between the ambient temperature from SCADA and forecasting temperature provided by forecast provider 1 and 3, correlation plots below are investigated depending upon which correlation is good enough. As seen in the below Figure 5.1 and Figure 5.2 for Wind Farm B as an exemplary of temperature correlation plots, temperature data of forecast provider 1 has good correlation, which is approximately 92% with the ambient temperature of Wind Farm B, in comparison to temperature data of forecast provider 3. Temperature data of forecast provider 1 has also good correlation with the ambient temperature of Wind Farm A and C, which is approximately 93% and corresponding correlation plots are shown in Figure 5.3 and Figure 5.4, respectively. This methodology includes high uncertainty due to nature of forecasting however, considering forecasting temperature of the provider 1 for grouping power production data will cause less uncertainty in comparison to forecasting temperature of the provider 3. As a result of this, the power production can be grouped by temperature bin based on forecasting temperature of provider 1. The temperature bin was selected as 1°C, 2°C and 3°C, respectively and best method has been determined as to move forward with temperature bin of 3°C due to the fact that this method has provided more accurate power forecasting close to SCADA power production. After grouping the forecasting of power production based on the temperature, forecasting of total production in each temperature bin are scaled by using SCADA power production. Afterwards, best forecast provider can be selected by using scaled production. To support these steps in selection of best forecast provider, the data regarding this process is illustrated in Figure 5.5, Figure 5.6, and Figure 5.7 below. For each wind farm included in this study, SCADA power production and forecasting power production data was grouped by the 3°C temperature bin and then scaled by SCADA power production. If the scaled power forecasting is how close to 1, the power forecasting will be close to SCADA production. Best forecast provider in terms of forecasting accuracy is determined by using scaled power production. In addition to Figure 5.5, Figure 5.6, and Figure 5.7 below, to explain the process in detail corresponding data related to the figures are shown right after the figures.

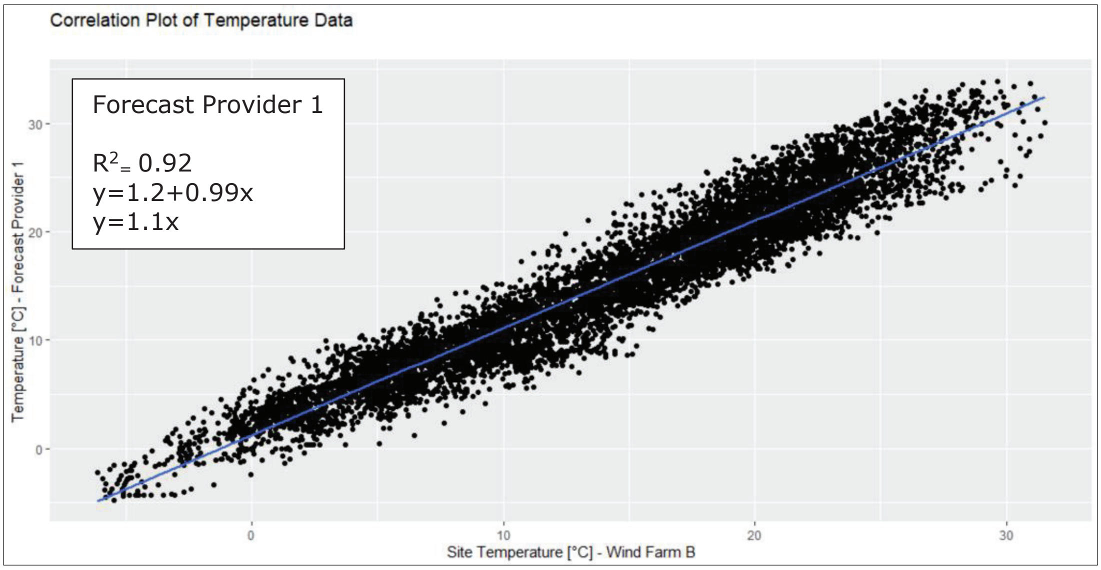

Figure 5.1 Correlation Plot btw Temperature of Forecast Provider 1 and Site Temperature of Wind Farm B

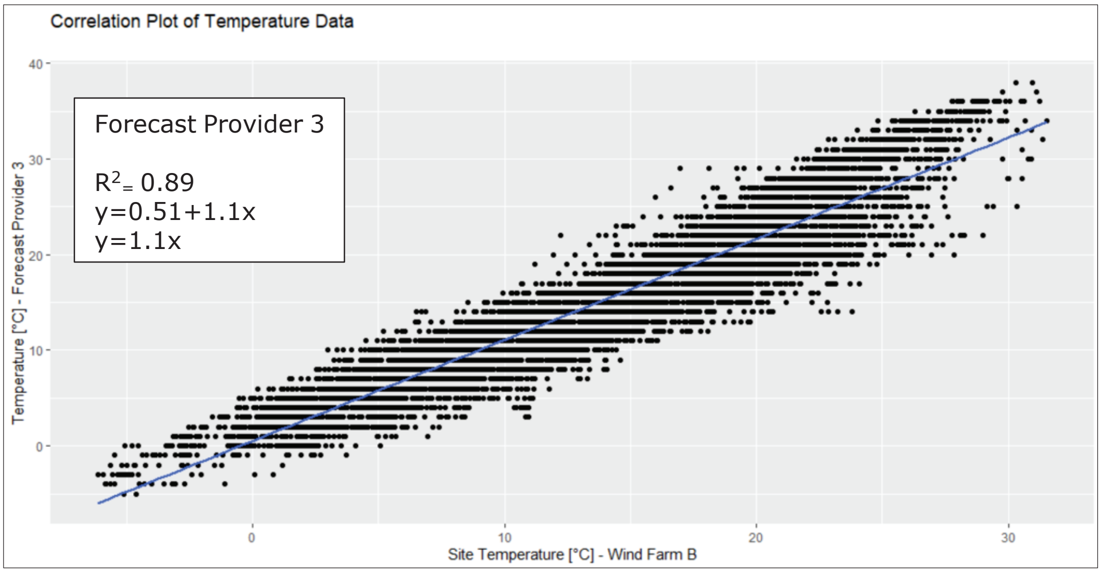

Figure 5.2 Correlation Plot btw Temperature of Forecast Provider 3 and Site Temperature of Wind Farm B

Obviously, temperature of forecast provider 3 is more scattered along the best line fit and due to this reason, it has less correlation with ambient site temperature. After investigation of correlation of temperature of forecast provider 1 for Wind Farm A and C, their forecasting of power production were grouped by temperature of forecast provider 1 just like process in Wind Farm B.

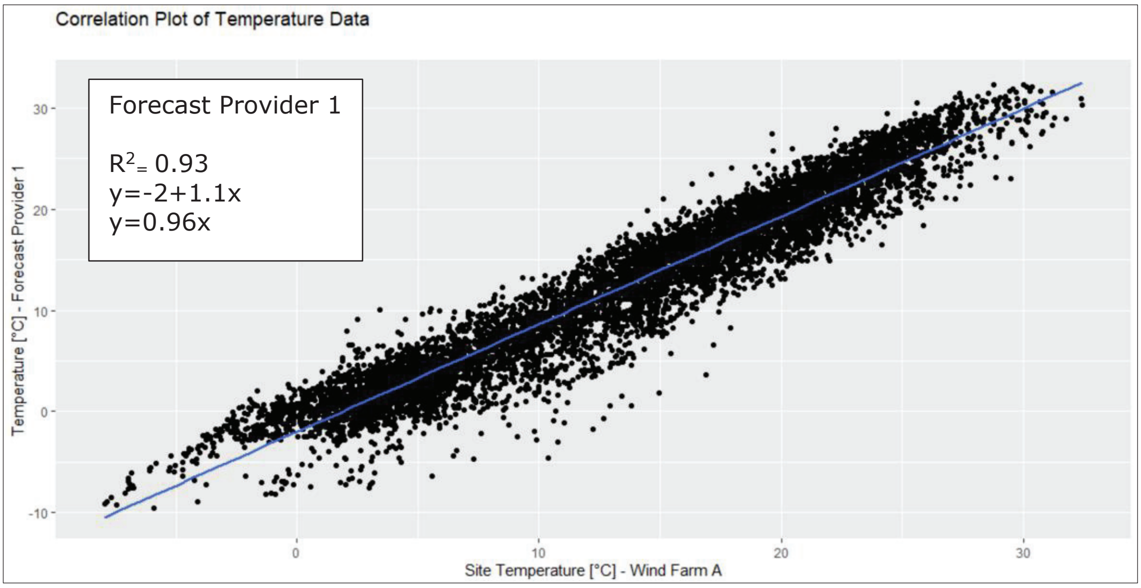

Figure 5.3 Correlation Plot btw Temperature of Forecast Provider 1 and Site Temperature of Wind Farm A

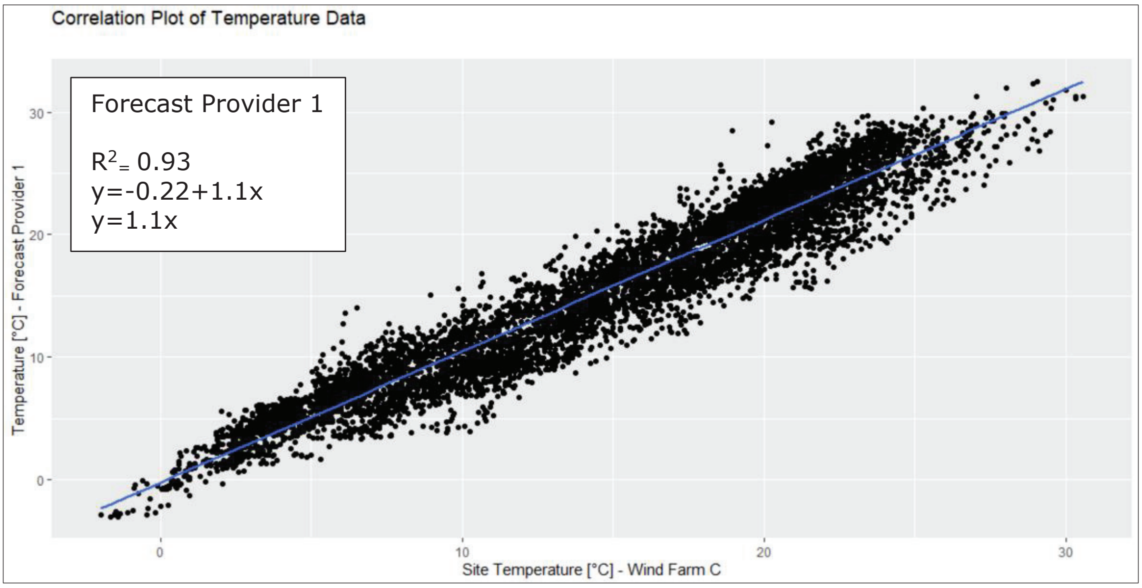

Figure 5.4 Correlation Plot btw Temperature of Forecast Provider 1 and Site Temperature of Wind Farm C

Table 5.1 Power Production of Wind Farm A Grouped by Temperature

Temperature Range [°C]	**Scada Production [MWh]**	**Forecast Provider 1 [MWh]**	**Forecast Provider 2 [MWh]**	**Forecast Provider 3 [MWh]**	**Forecast Provider 4 [MWh]**
(-10,-7]	825.77	464.05	588.09	699.63	396.92
(-7,-4]	1461.26	743.17	1026.25	986.00	800.44
(-4,-1]	4485.82	3129.38	4138.90	4354.23	3684.16
(-1,2]	19238.51	14467.48	21290.12	21510.91	18692.23
(2,5]	22479.33	18638.26	23406.03	24139.20	21290.36

cont. on next page

cont. of table 5.1

(5,8]	10872.89	9367.46	11307.27	11140.48	10371.40
(8,11]	7485.88	6720.88	7833.51	7572.60	7369.69
(11,14]	8350.19	7830.53	8129.38	8029.76	8070.14
(14,17]	15021.86	13909.28	13527.98	13638.12	13663.98
(17,20]	15431.97	14970.98	14299.71	14788.73	14289.61
(20,23]	12298.83	11971.24	11408.49	11655.29	11436.30
(23,26]	9469.63	9153.59	8906.41	9251.74	9018.79
(26,29]	5623.71	5364.68	5528.09	5494.38	5210.40
(29,32]	967.03	798.76	963.00	826.71	991.98

Based on the forecast provider 1, ambient temperature of Wind Farm A varies from minimum of -10°C to maximum of 32°C. Forecasting of power production and SCADA power production are aggregated by using temperature range, which is at intervals of 3°C. Thus, at the specific temperature intervals accurate power forecasting should be revealed.

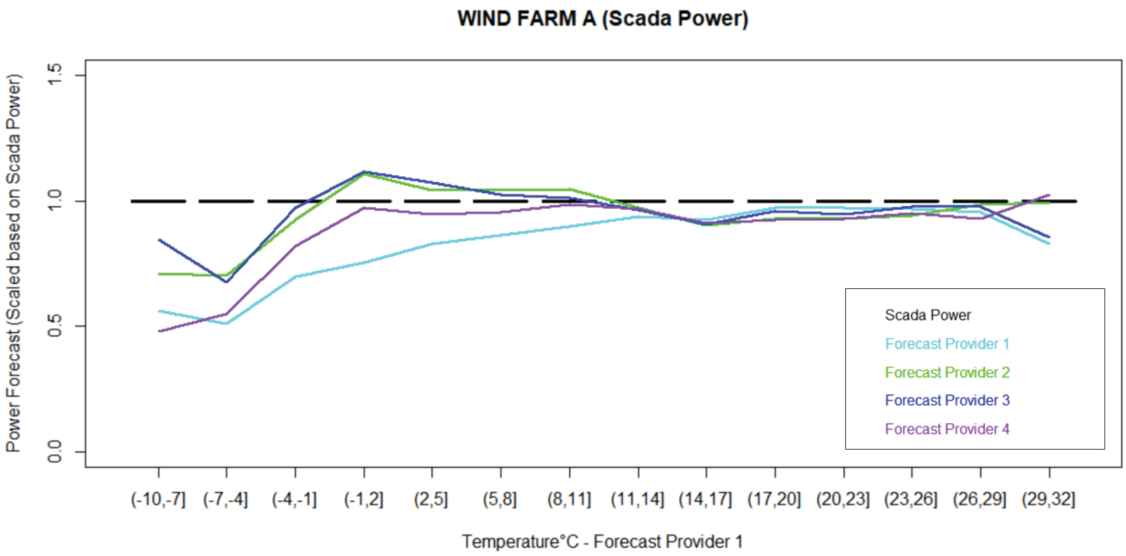

Figure 5.5 Scaled Power Production Forecasting of Wind Farm A by Temperature of Forecast Provider 1

Data in Table 5.1 can be scaled by using SCADA power production. Figure 5.5 above represents scaled forecasting of power production. As is seen from the figure, at low temperatures the forecast providers underestimate power production of Wind Farm A. Data of the figure above is shown seen in Table 5.2 and the forecast provider that provides more accurate power production forecast at the specific temperature intervals compared to other ones are determined. Based on this methodology, new time series of power production forecasting can be created. For instance, if the site temperature between -10°C and -7°C, forecast provider 3 should be preferred in order to choose most accurate power production forecasting.

Table 5.2 Scaled Power Production of Wind Farm A Grouped by Temperature

Temperature Range [°C]	Scaled Forecast Provider 1	Scaled Forecast Provider 2	Scaled Forecast Provider 3	Scaled Forecast Provider 4	Best Choice	Forecast Provider
(-10,-7]	0.562	0.712	0.847	0.481	0.847	3
(-7,-4]	0.509	0.702	0.675	0.548	0.702	2
(-4,-1]	0.698	0.923	0.971	0.821	0.971	3
(-1,2]	0.752	1.107	1.118	0.972	0.972	4
(2,5]	0.829	1.041	1.074	0.947	1.041	2
(5,8]	0.862	1.040	1.025	0.954	1.025	3
(8,11]	0.898	1.046	1.012	0.984	1.012	3
(11,14]	0.938	0.974	0.962	0.966	0.974	2
(14,17]	0.926	0.901	0.908	0.910	0.926	1
(17,20]	0.970	0.927	0.958	0.926	0.970	1
(20,23]	0.973	0.928	0.948	0.930	0.973	1
(23,26]	0.967	0.941	0.977	0.952	0.977	3
(26,29]	0.954	0.983	0.977	0.927	0.983	2
(29,32]	0.826	0.996	0.855	1.026	0.996	2

Based on the forecast provider 1, ambient temperature of Wind Farm B is between the range of minimum of -5°C and maximum of 34°C. Forecasting of power production and SCADA power production are aggregated by using 3°C temperature interval.

Table 5.3 Power Production of Wind Farm B Grouped by Temperature

Temperature Range [°C]	Scada Production [MWh]	Forecast Provider 1 [MWh]	Forecast Provider 2 [MWh]	Forecast Provider 3 [MWh]	Forecast Provider 4 [MWh]
(-5,-2]	628.00	869.86	511.56	783.48	477.51
(-2,1]	1293.52	1820.42	1282.22	1662.67	1491.28
(1,4]	6576.52	8373.61	7067.51	7729.21	7626.23
(4,7]	10107.37	10457.97	10990.50	10327.55	10448.13
(7,10]	13451.04	15259.01	15049.21	14606.00	14764.92
(10,13]	8993.08	10324.57	10501.56	10034.12	10180.40
(13,16]	7101.31	8185.93	8391.44	7212.90	7544.81
(16,19]	9775.32	11337.26	10405.42	9161.16	9650.07
(19,22]	10855.86	12842.17	11589.49	10229.81	11309.87
(22,25]	10436.84	12356.96	11057.15	10120.52	11369.59
(25,28]	6895.96	7522.49	7297.48	6666.49	7472.52
(28,31]	3456.33	3209.90	3311.56	3053.80	3654.34
(31,34]	930.39	869.97	915.14	865.62	956.91

Same methodology applied in Wind Farm A is also conducted for Wind Farm B. As is seen from the Figure 5.6, at low temperatures the forecast providers generally overestimate power production of Wind Farm B.

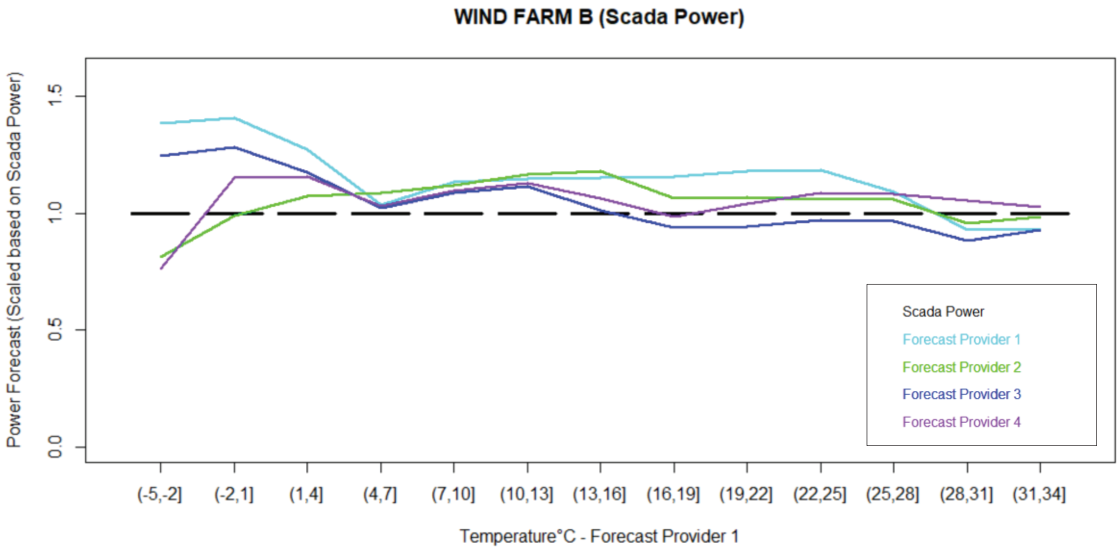

Figure 5.6 Scaled Power Production Forecasting of Wind Farm B by Temperature of Forecast Provider 1

New time series of power production forecasting of Wind Farm B is created depending upon the best choice column in Table 5.4.

Table 5.4 Scaled Power Production of Wind Farm B Grouped by Temperature

Temperature Range [°C]	**Scaled Forecast Provider 1**	**Scaled Forecast Provider 2**	**Scaled Forecast Provider 3**	**Scaled Forecast Provider 4**	**Best Choice**	**Forecast Provider**
(-5,-2]	1.385	0.815	1.248	0.760	0.815	2
(-2,1]	1.407	0.991	1.285	1.153	0.991	2
(1,4]	1.273	1.075	1.175	1.160	1.075	2
(4,7]	1.035	1.087	1.022	1.034	1.022	3
(7,10]	1.134	1.119	1.086	1.098	1.086	3
(10,13]	1.148	1.168	1.116	1.132	1.116	3
(13,16]	1.153	1.182	1.016	1.062	1.016	3
(16,19]	1.160	1.064	0.937	0.987	0.987	4
(19,22]	1.183	1.068	0.942	1.042	1.042	4
(22,25]	1.184	1.059	0.970	1.089	0.970	3
(25,28]	1.091	1.058	0.967	1.084	0.967	3
(28,31]	0.929	0.958	0.884	1.057	0.958	2
(31,34]	0.935	0.984	0.930	1.028	0.984	2

Based on the forecast provider 1, ambient temperature of Wind Farm C is between the range of minimum of -4°C and maximum of 32°C. Forecasting of power production and

SCADA power production are aggregated in the same manner by using 3°C temperature interval.

Table 5.5 Power Production of Wind Farm C Grouped by Temperature

Temperature Range [°C]	Scada Production [MWh]	Forecast Provider 1 [MWh]	Forecast Provider 2 [MWh]	Forecast Provider 3 [MWh]	Forecast Provider 4 [MWh]
(-4,-1]	653.28	491.45	644.38	607.91	676.48
(-1,2]	7261.95	6014.17	6529.72	6311.82	6534.07
(2,5]	26188.86	23462.38	25940.35	24103.85	25291.77
(5,8]	37157.51	35955.84	37983.11	36261.97	37548.14
(8,11]	32742.63	28815.33	31519.18	31319.62	30514.76
(11,14]	21635.29	20765.75	21796.09	22228.67	21231.40
(14,17]	23161.44	25700.01	26239.90	25048.92	25048.16
(17,20]	20649.61	22231.87	22937.90	21929.12	21824.91
(20,23]	49232.98	55319.60	51161.70	53708.36	51724.97
(23,26]	42322.37	44263.84	41652.82	42287.93	42273.24
(26,29]	18120.59	19124.00	17566.68	18036.60	18069.21
(29,32]	992.08	970.72	969.07	941.00	983.89

Same methodology applied in Wind Farm A and B is also conducted for Wind Farm C.

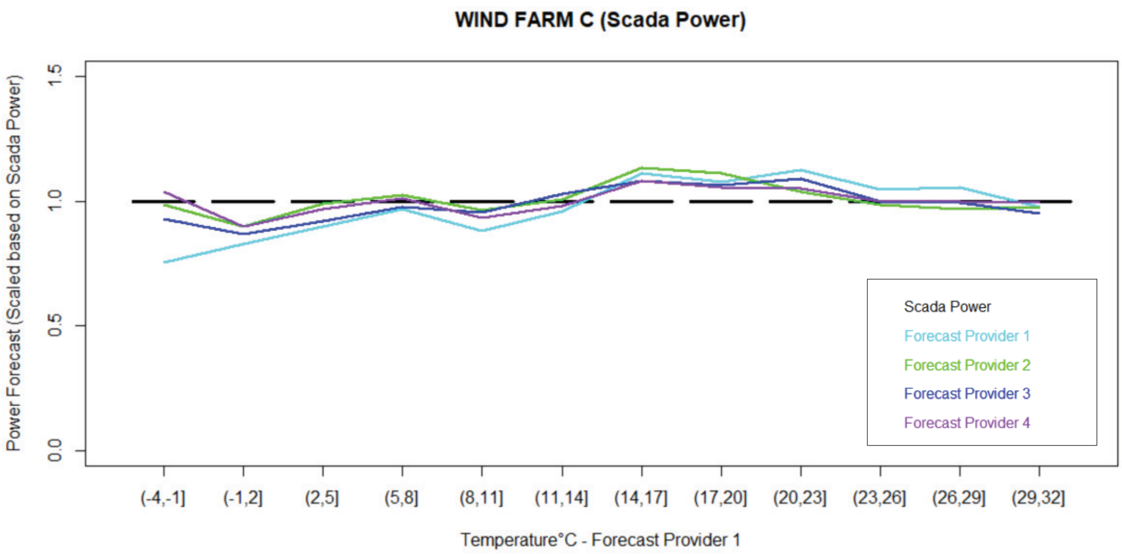

Figure 5.7 Scaled Power Production Forecasting of Wind Farm C by Temperature of Forecast Provider 1

New time series of power production forecasting for Wind Farm C is created depending upon the best choice column in Table 5.6. As is seen from the Figure 5.7 power production forecasting at low temperatures does not deviates much more from SCADA power production, which explains why scaled power production values are nearly close to one.

Unlike the significant deviation in the Wind Farm A and B, difference between SCADA and forecasting production seems acceptable based on Figure 5.7. It can be related to lack of icing events in the Wind Farm C site, this is because its altitude is considerably less than altitude of Wind Farm A and B.

Table 5.6 Scaled Power Production of Wind Farm C Grouped by Temperature

Temperature Range [°C]	Scaled Forecast Provider 1	Scaled Forecast Provider 2	Scaled Forecast Provider 3	Scaled Forecast Provider 4	Best Choice	Forecast Provider
(-4,-1]	0.752	0.986	0.931	1.036	0.986	2
(-1,2]	0.828	0.899	0.869	0.900	0.900	2
(2,5]	0.896	0.991	0.920	0.966	0.991	2
(5,8]	0.968	1.022	0.976	1.011	1.011	4
(8,11]	0.880	0.963	0.957	0.932	0.963	2
(11,14]	0.960	1.007	1.027	0.981	1.007	2
(14,17]	1.110	1.133	1.081	1.081	1.081	4
(17,20]	1.077	1.111	1.062	1.057	1.057	4
(20,23]	1.124	1.039	1.091	1.051	1.039	2
(23,26]	1.046	0.984	0.999	0.999	0.999	4
(26,29]	1.055	0.969	0.995	0.997	0.997	4
(29,32]	0.978	0.977	0.949	0.992	0.992	4

5.2. Wind Farm Power Production Based on Turbulence Intensity

Effect of turbulence intensity on power production forecasting can be investigated by following similar methodology mentioned in previous section regarding ambient site temperature. Power production forecasting and SCADA power production of the wind farms can be grouped by specific turbulence intensity values. However, forecasting of turbulence intensity has not been provided by forecast providers. Wind speed at the wind farm site can be associated with turbulence intensity. If the wind speed at the site is low, turbulence intensity is relatively high. Turbulence intensity is calculated simply as standard deviation value of wind speed divided by wind speed value.

To investigate effect of turbulence intensity, one of the turbines in Wind Farm A, B, and C should be selected by considering internal wake effects, in another saying selected turbines should be exposed to free wind speed at the site. The turbine wind speed and corresponding turbulence intensity can be drawn as plot to observe turbulence at the various wind speed. At the specific wind speed range, turbulence intensity can be grouped as very low, moderate, high, and heavy. It should be noted that the turbine wind speed is real-time measurement and forecasting of wind speed is also needed for grouping of power production based on specific turbulence intensity range. In this case, wind speed forecasting and site wind speed can be comparing by looking at the correlations and best wind speed forecasting is selected to represent site wind speed.

As mentioned before on previous chapters, some forecast providers, in addition to power production forecasting for the wind farms, also provide meteorological parameters such as temperature, pressure, relative humidity and wind speed. Wind speed forecasting from forecast provider 1 as well as forecast provider 3 and average of site wind speed from Wind Farm A, B and C can be compared in order to select representative wind speed forecasting data. Most representative wind speed forecasting data can be used in place of real-time wind speed measurement to identify turbulence intensity range at the site. Correlation between wind speed forecasting from forecast provider 1 and average of site wind speed from the wind farms are plotted in below.

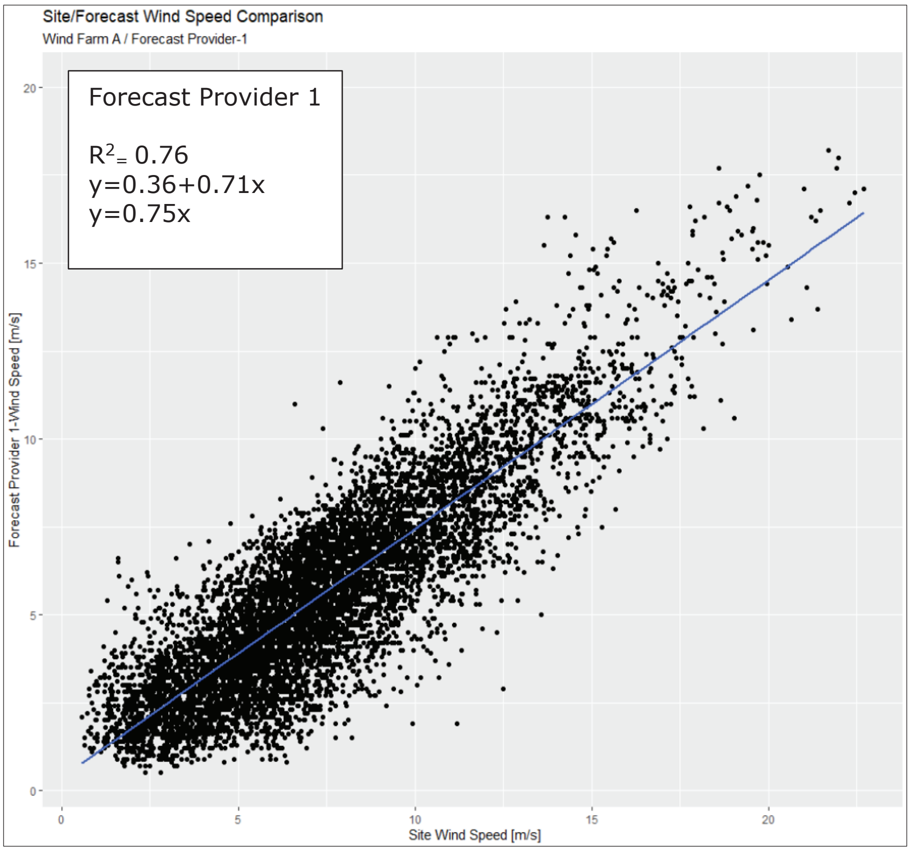

Figure 5.8 Correlation Between Average Wind Speed of Wind Farm A and Wind Speed Forecasting from Forecast Provider 1

Another correlation between the site wind speed of Wind Farm A and related wind speed forecasting from forecast provider 3 should be investigated. Based on the correlation, the best forecast provider, which has less deviation from the site wind speed compared to other providers, can be selected to represent site wind speed by using forecasting of wind speed. The wind speed correlation plot including forecast provider 3 is shown in below. According to Figure 5.9 , deficient correlation between the site wind speed and wind speed forecasting for the Wind Farm A is seen clearly. For this reason, in order to represent site wind speed at Wind Farm A, wind speed forecasting from forecast provider 1 should be preferred.

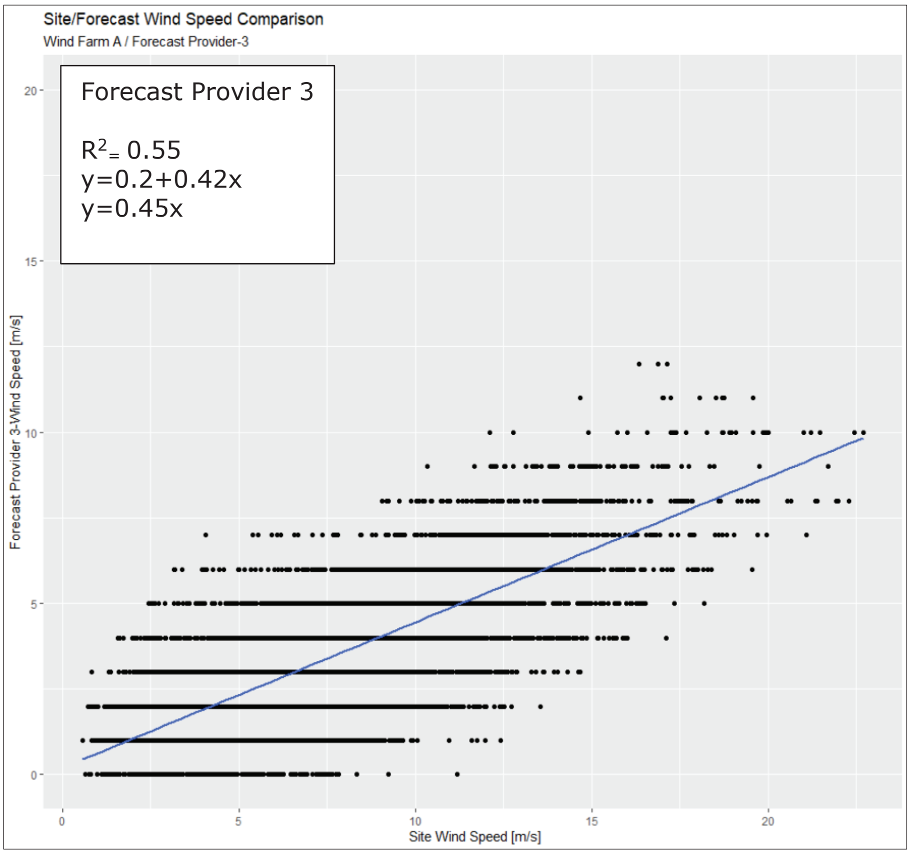

Figure 5.9 Correlation Between Average Wind Speed of Wind Farm A and Wind Speed Forecasting from Forecast Provider 3

For the Wind Farm B and C, correlation plots between average of site wind speed and related forecasting wind speed from forecast providers 1 and 3 were drawn, respectively. First two graphs show that correlation between site wind speed at Wind Farm B and wind speed forecasting of forecast providers 1 and 3. Based on the correlation, similar results like in Wind Farm A were observed and wind speed forecasting from forecast provider 1 was used to represent the site wind speed at Wind Farm B. Results can be seen from Figure 5.10 and Figure 5.11 in detail.

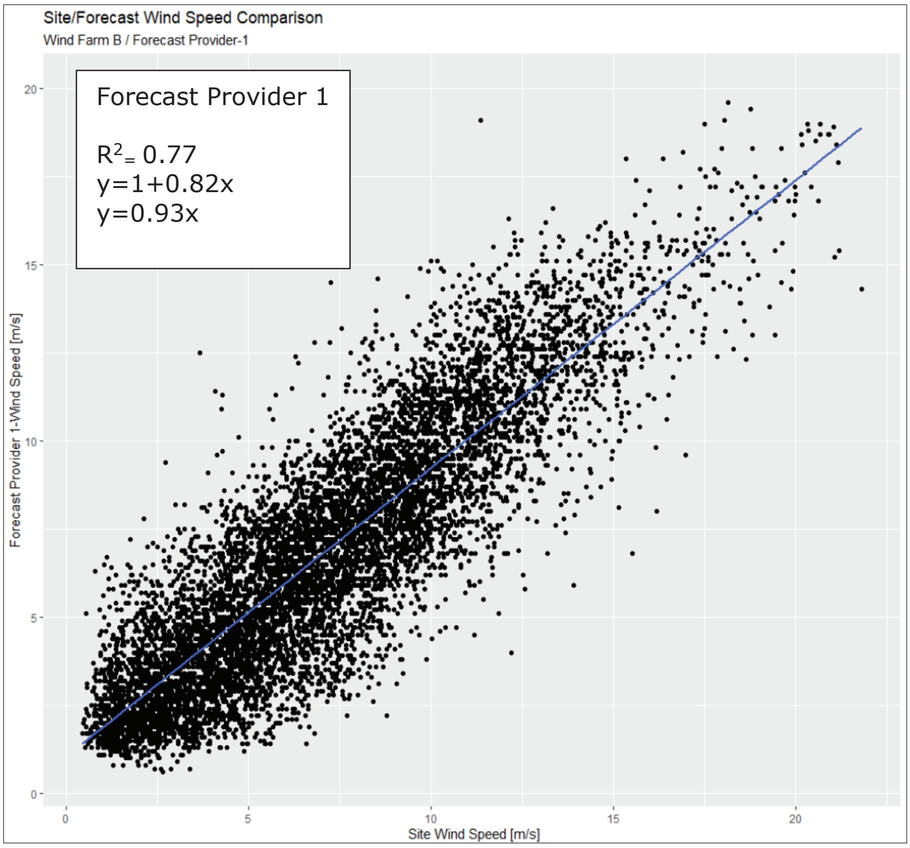

Figure 5.10 Correlation Between Average Wind Speed of Wind Farm B and Wind Speed Forecasting from Forecast Provider 1

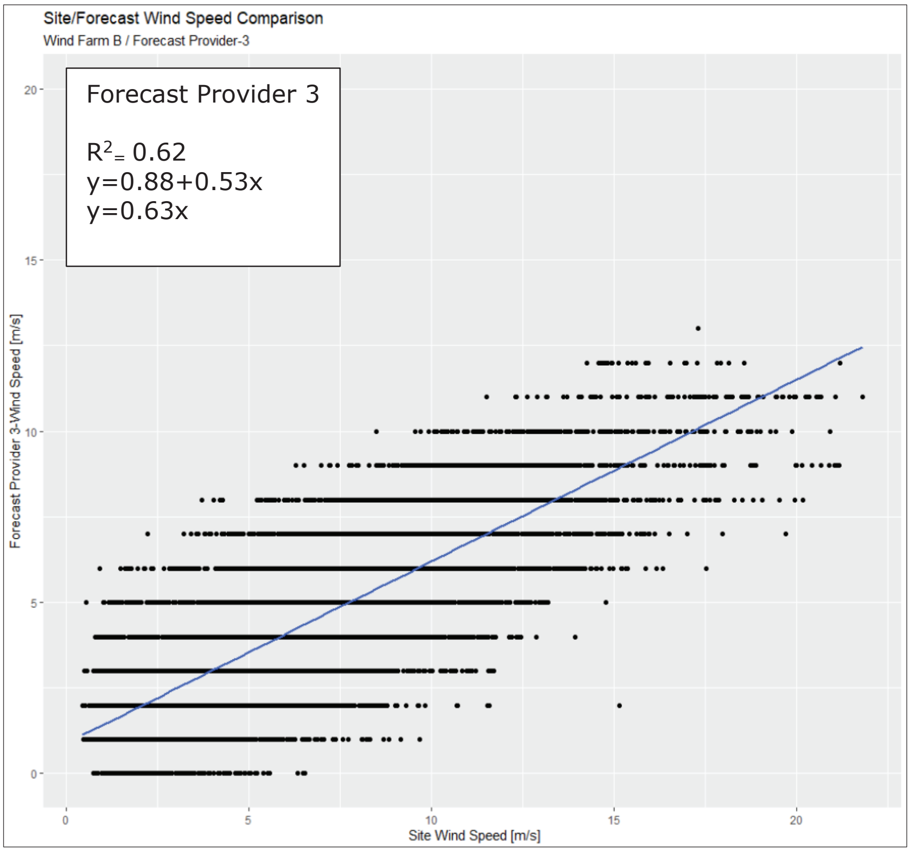

Figure 5.11 Correlation Between Average Wind Speed of Wind Farm B and Wind Speed Forecasting from Forecast Provider 3

Following plots are belonging to comparisons of site wind speed of Wind Farm C and wind speed forecasting from forecast providers 1 and 3. According to Figure 5.12 and Figure 5.13, good correlation were observed between wind speed forecasting from provider 1 and site wind speed at Wind Farm C.

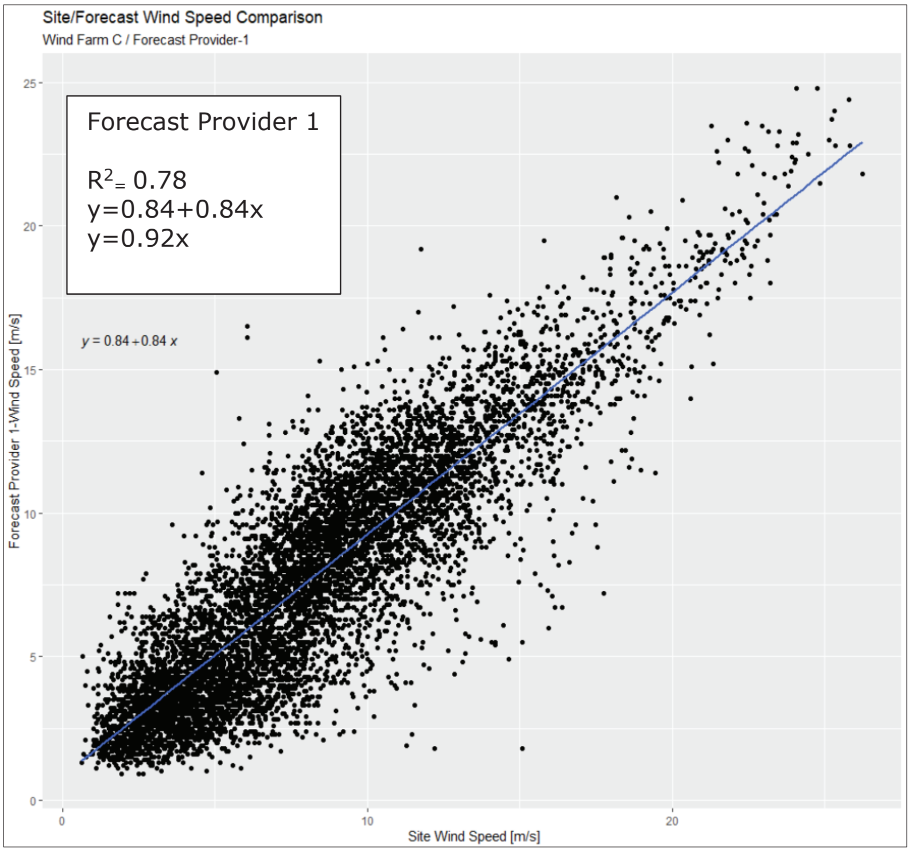

Figure 5.12 Correlation Between Average Wind Speed of Wind Farm C and Wind Speed Forecasting from Forecast Provider 1

As a result of these correlation analyses between site wind speed and related forecasting wind speed, forecasting provider 1 was selected due to the fact that its wind speed forecasting has less deviation from the site wind speed at Wind Farm A, B and C. Based on this, turbulence intensity of the wind farms in the study can be categorized into four parts as from very low, moderate to high and heavy. To investigate average of turbulence intensity of the site, some wind turbines that under exposure of free wind speed at the wind farm can be selected and then turbulence intensity can be categorized based on wind speed forecasting from forecast provider 1, which is representative of the site wind speed for all wind farms.

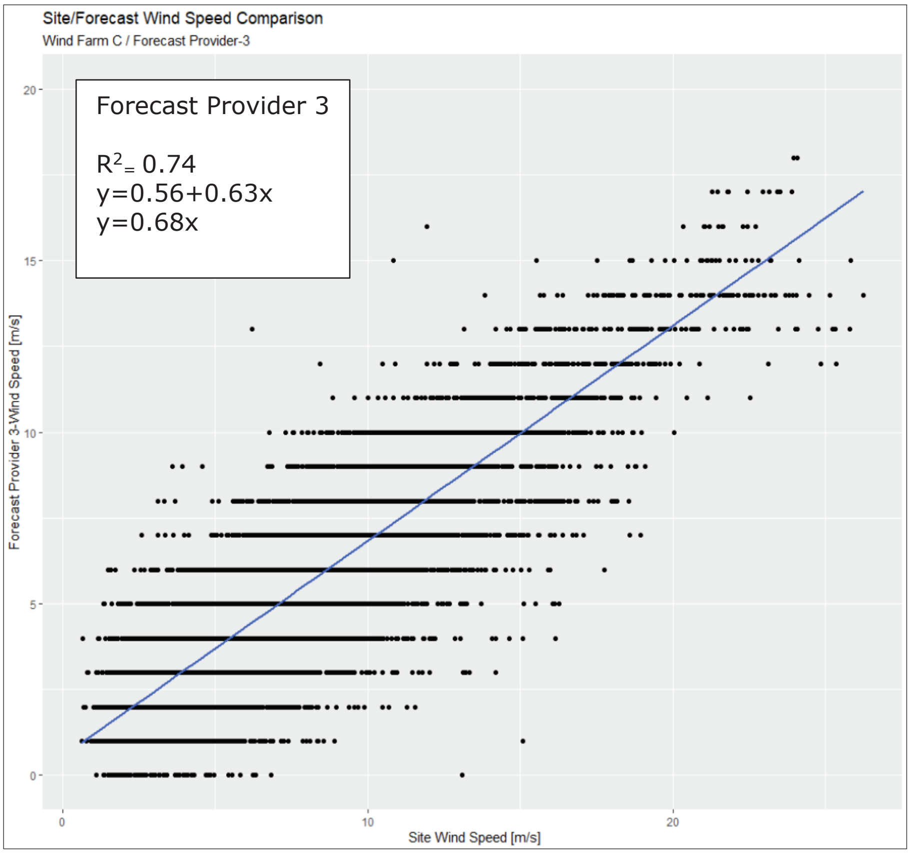

Figure 5.13 Correlation Between Average Wind Speed of Wind Farm C and Wind Speed Forecasting from Forecast Provider 3

Turbine 16 from Wind Farm A, Turbine 8 from Wind Farm B and Turbine 23 from Wind Farm C can be preferred to observe site turbulence without any wake effects caused by other wind turbines at the wind farms. These turbines were selected by taking both wind farm layout and prevailing wind direction at the wind farm into consideration. The layout of the wind farms can be seen in Figure 2.1, Figure 2.2 and Figure 2.3, while wind rose of the wind farms that shows prevailing wind direction can be seen in below Figure 5.14, Figure 5.15 and Figure 5.16.

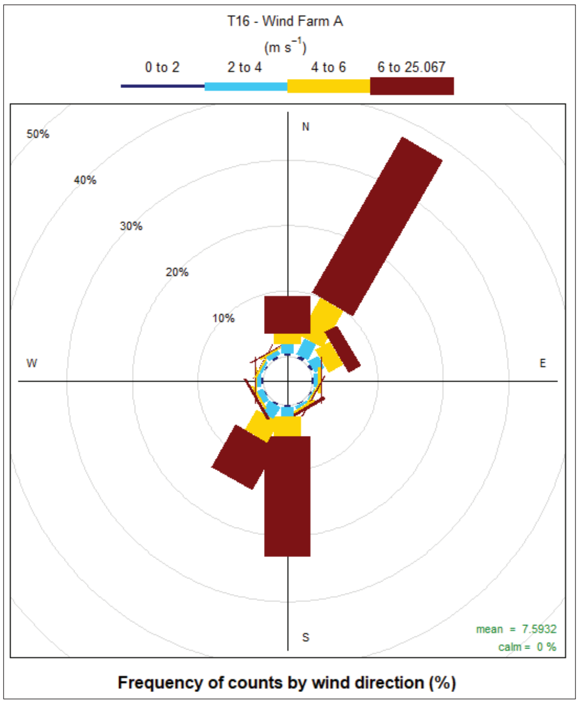

Figure 5.14 Exemplary Wind Rose for Wind Farm A

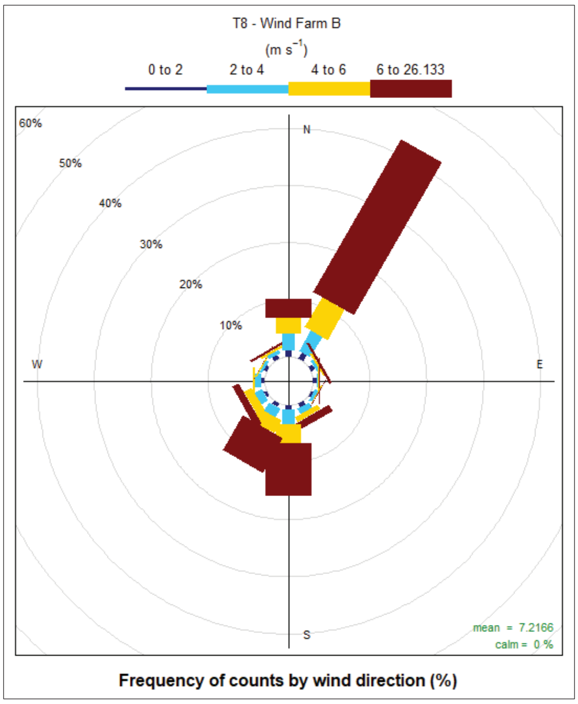

Figure 5.15 Exemplary Wind Rose for Wind Farm B

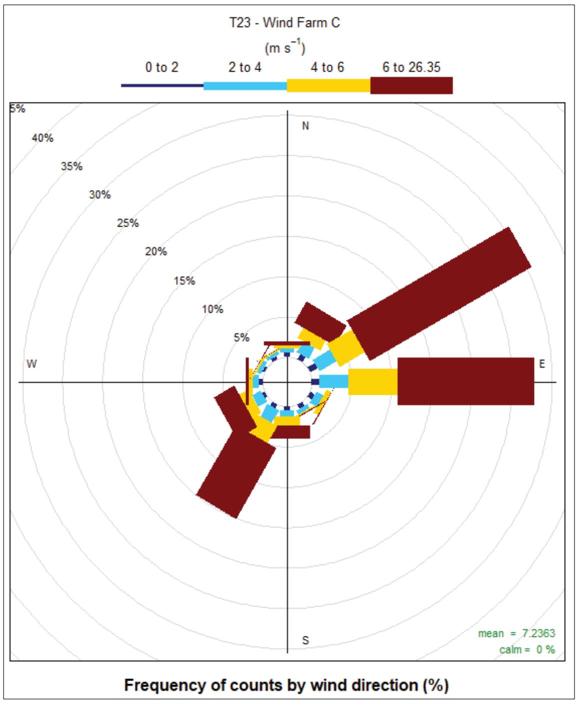

Figure 5.16 Exemplary Wind Rose for Wind Farm C

In order for detecting turbulence intensity range by considering wind speed value, turbulence intensity of Turbine 16 from Wind Farm A, Turbine 8 from Wind Farm B and Turbine 23 from Wind Farm C can be visualized based on wind speed forecasting of the forecast provider 1 for the wind farms. In Figure 5.17, Figure 5.18 and Figure 5.19, x-axis shows wind speed forecasting from forecast provider 1, while y-axis shows turbulence intensity of the turbine. As seen in below figures, low wind speed generally creates high turbulence at the site. Some increase of turbulence intensity at the high wind speed in Figure 5.17 and Figure 5.18 might be caused by sensor failure of the anemometer on the wind turbine nacelle. In general, turbulence intensity is inversely proportional with wind speed. By ignoring this sudden increase of turbulence intensity at the high wind speed, categorization of turbulence intensity can be created based on the exemplary turbulence intensity figures below for the Wind Farm A, B and C.

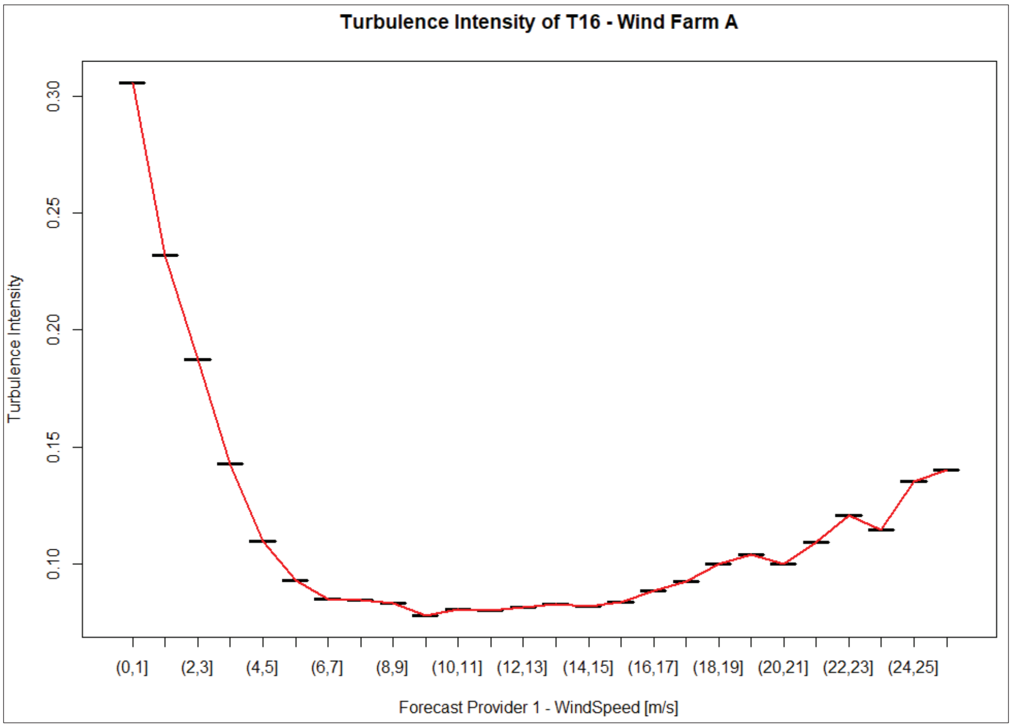

Figure 5.17 Exemplary Turbulence Intensity Graph for Wind Farm A

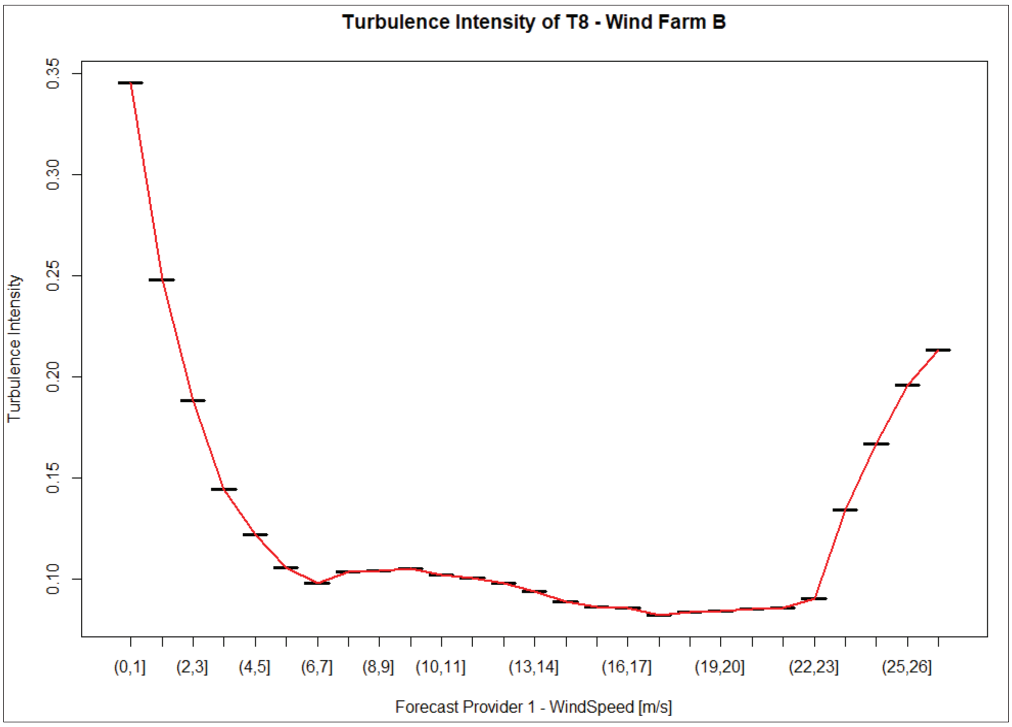

Figure 5.18 Exemplary Turbulence Intensity Graph for Wind Farm B

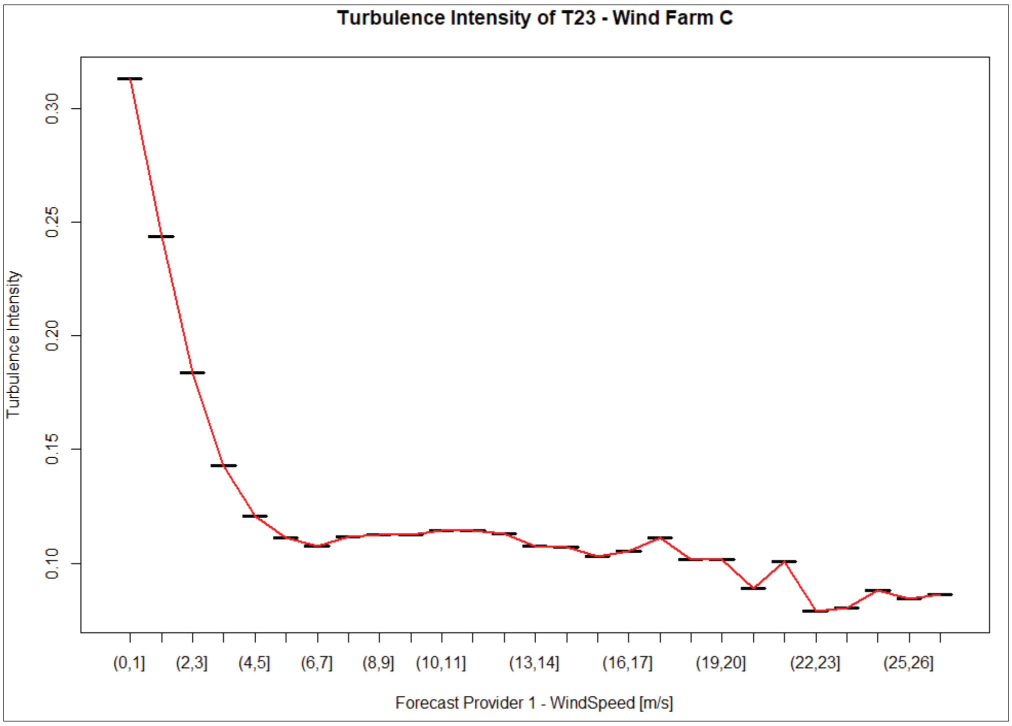

Figure 5.19 Exemplary Turbulence Intensity Graph for Wind Farm C

According to Figure 5.17, Figure 5.18 and Figure 5.19 above, turbulence intensity range can be categorized by wind speed bin. If the wind speed is less than 3 m/s, it points out very high turbulence at the site. If the wind speed is between the range of 3 m/s and 6 m/s, it points out moderate turbulence at the site. The wind speed is between the range of 6 m/s and 12 m/s, it points out high turbulence at the site. The wind speed values that are higher than 12 m/s, it points out low turbulence. Based on this categorization of the turbulence intensity, power production forecasting and SCADA production data of the wind farms can be grouped, and best forecast provider can be selected by considering these turbulence intensities at specific wind speed interval.

Table 5.7 Power Production of Wind Farm A Grouped by Turbulence Intensity

Turbulence Intensity	**Best Forecast Provider**	**Wind Speed Range [m/s]**	**Forecast Provider 1 [MWh]**	**Forecast Provider 2 [MWh]**	**Forecast Provider 3 [MWh]**	**Forecast Provider 4 [MWh]**	**Scada Production [MWh]**
Heavy	4	(0,3]	4027.77	5139.78	4838.65	**4815.99**	4486.05
High	2	(3,6]	24927.24	**31329.11**	30755.80	29303.01	31298.50
Moderate	2	(6,12]	81422.65	**89538.99**	91887.39	85008.21	90478.20
Very low	1	(12,20]	**7152.07**	6345.33	6605.91	6159.18	7749.94

According to Table 5.7, if the Wind Farm A has heavy turbulence intensity due to low wind speed until 3 m/s, forecast provider 4 gives more accurate power forecasting compared to remaining forecast providers, which have more deviation from SCADA power production. If high and moderate turbulence exist at the site, forecast provider 2 can be preferred to acquire more accurate power forecasting, which means less deviation from SCADA power production. If the wind speed is higher 12 m/s and turbulence intensity is very low, forecast provider 1 will provide more accurate power forecasting. By considering the different turbulence intervals based on Table 5.7, new time series of power production forecasting can be created to improve accuracy of forecasting and the time series can be compared with time series of power production forecasting of the forecast providers.

Same methodology is applied for Wind Farm B and C. In Wind Farm B, forecast provider 3 will provide more accurate power forecasting at both heavy and moderate turbulence circumstance. At high and very low turbulence, forecast provider 1 and 2 should be preferred to acquire more accurate power production forecasting of Wind Farm B.

Table 5.8 Power Production of Wind Farm B Grouped by Turbulence Intensity

Turbulence Intensity	**Best Forecast Provider**	**Wind Speed Range [m/s]**	**Forecast Provider 1 [MWh]**	**Forecast Provider 2 [MWh]**	**Forecast Provider 3 [MWh]**	**Forecast Provider 4 [MWh]**	**Scada Production [MWh]**
Heavy	3	(0,3]	1752.30	3451.17	**1250.77**	1896.91	1149.23
High	1	(3,6]	**9424.54**	15185.72	10310.82	12123.19	9472.78
Moderate	3	(6,12]	65227.97	59602.45	**58464.07**	60314.44	58683.47
Very low	2	(12,20]	27025.29	**20130.90**	22427.68	22612.03	21196.06

In Wind Farm C, selected forecast provider for different turbulence circumstance can be seen in Table 5.9 in detail. Forecast provider 4 should be preferred in order to decrease deviation in power forecasting in case of high, moderate and very low turbulence at the site. If heavy turbulence circumstance at the site is expected, forecast provider 3 should be included for acquiring power production forecasting close to SCADA power production.

Table 5.9 Power Production of Wind Farm C Grouped by Turbulence Intensity

Turbulence Intensity	**Best Forecast Provider**	**Wind Speed Range [m/s]**	**Forecast Provider 1 [MWh]**	**Forecast Provider 2 [MWh]**	**Forecast Provider 3 [MWh]**	**Forecast Provider 4 [MWh]**	**Scada Production [MWh]**
Heavy	3	(0,3]	3002.47	5338.51	**3414.62**	3977.85	3327.38
High	4	(3,6]	19161.14	27603.77	22587.57	**24415.44**	23904.19
Moderate	4	(6,12]	160856.30	156933.90	163411.40	**159052.90**	158472.20
Very low	4	(12,20]	100095.10	95064.70	93372.16	**94274.82**	94414.82

Based on different turbulence circumstances at the wind farms, criteria of selection for best power forecasting are determined and then new time series of power production forecasting is created in addition to power production forecasting from forecast providers.

5.3. Wind Farm Power Production Based on Both Temperature and Turbulence Intensity

In previous section, effect of temperature and turbulence intensity on power production forecasting was examined separately. By taking various temperature and turbulence intervals into consideration, new time series of power production forecasting were created to acquire more accurate power forecasting in comparison to power forecasting of forecast providers. However, these two parameters cannot be considered separately in the atmosphere. Therefore, new time series of power forecasting just like in previous sections can be created by taking these parameters together into account. Criterion matrix for best selection of power production forecasting from forecast providers should be analyzed to create new time series of power production forecasting. These matrixes are shared below for Wind Farm A, B and C, respectively.

Table 5.10 Power Production of Wind Farm A Grouped by both Temperature and Turbulence Intensity

Forecast Provider	**Temperature Range [°C]**	**(0,3] m/s, Heavy Turbulence**	**(3,6] m/s, High Turbulence**	**(6,12] m/s, Moderate Turbulence**	**(12,] m/s, Very Low Turbulence**
Best forecast provider by considering temperature only		**Best forecast provider by considering both temperature and turbulence intensity**			

cont. on next page

cont. of table 5.10

3	(-10,-7]	2	2	3	NA*
2	(-7,-4]	4	3	2	1
3	(-4,-1]	1	2	4	1
4	(-1,2]	4	4	4	3
2	(2,5]	1	4	2	1
3	(5,8]	1	3	3	1
3	(8,11]	3	4	2	4
2	(11,14]	3	3	3	NA*
1	(14,17]	1	4	1	NA*
1	(17,20]	4	3	3	1
1	(20,23]	4	2	1	NA*
3	(23,26]	2	2	3	NA*
2	(26,29]	2	4	3	NA*
2	(29,32]	3	4	4	NA*

*data is not available.

Table 5.11 Power Production of Wind Farm B Grouped by both Temperature and Turbulence Intensity

Forecast Provider	**Temperature Range [°C]**	**(0,3] m/s, Heavy Turbulence**	**(3,6] m/s, High Turbulence**	**(6,12] m/s, Moderate Turbulence**	**(12,] m/s, Very Low Turbulence**
Best forecast provider by considering temperature only		**Best forecast provider by considering both temperature and turbulence intensity**			
2	(-5,-2]	4	1	2	3
2	(-2,1]	3	3	2	2
2	(1,4]	3	1	2	4
3	(4,7]	1	3	1	2
3	(7,10]	3	3	2	2
3	(10,13]	3	3	1	2
3	(13,16]	3	3	3	4
4	(16,19]	3	3	2	4
4	(19,22]	3	3	2	4
3	(22,25]	3	3	2	2
3	(25,28]	3	3	2	2
2	(28,31]	3	1	4	4
2	(31,34]	1	2	4	1

*data is not available.

For instance, if turbulence is moderate and temperature is between the range of 4 and 7 at the site, best forecast provider can be selected at this specific range. It can be conducted by comparing power production forecasting and SCADA production of this specific range. Table 5.10 and Table 5.11 above show criterion matrixes for Wind Farm A and B, while Table 5.12 below shows the criterion matrix for Wind Farm C. Turbulence intensity and temperature interval are determined by using wind speed and temperature of forecast provider 1 just like in previous sections.

Table 5.12 Power Production of Wind Farm C Grouped by both Temperature and Turbulence Intensity

Forecast Provider	**Temperature Range [°C]**	**(0,3] m/s, Heavy Turbulence**	**(3,6] m/s, High Turbulence**	**(6,12] m/s, Moderate Turbulence**	**(12,] m/s, Very Low Turbulence**
Best forecast provider by considering temperature only		**Best forecast provider by considering both temperature and turbulence intensity**			
2	(-4,-1]	1	2	2	NA*
2	(-1,2]	3	2	4	1
2	(2,5]	4	2	2	4
4	(5,8]	1	3	4	4
2	(8,11]	3	2	3	1
2	(11,14]	1	4	4	2
4	(14,17]	4	1	4	3
4	(17,20]	4	3	4	3
2	(20,23]	3	3	2	4
4	(23,26]	3	3	3	2
4	(26,29]	3	3	3	3
4	(29,32]	1	2	4	2

*data is not available.

According to Table 5.10, Table 5.11 and Table 5.12, new time series of power production forecasting are created and, then the time series are compared with power production forecasting from forecast providers in order to check if there is an improvement on the forecasting. In following chapter, results, and discussions regarding the improvement on wind power forecasting are shared in detail.

CHAPTER 6

RESULTS AND DISCUSSION

In previous section, power forecasting data from all forecast providers was grouped by both site temperature and turbulence intensity and then new time series of forecasting were created by selecting best power forecasting compared to real power production within specific period. Created new time series were examined by hourly and daily as well as seasonal. Thus, it was observed whether there is considerable improvement on power forecasting with this methodology used in this study.

To check the improvement on power forecasting, time series of power production forecasting data from all forecast providers and new time series of power forecasting data based on atmospheric conditions at the sites were compared with real power production on SCADA by checking their coefficient of determination, in another saying R-squared value that explains how differences in the dependent variable can be explained by a difference in the independent variable. Firstly, relationships among the time series of real power production and forecasting data were investigated as both hourly and daily. Afterwards, the time series data can be grouped by seasonal based on equinox periods. In addition to hourly and daily relationship of the time series with the real power production, seasonal relationship was investigated.

In Table 6.1, relationship based on R-squared value between the time series of real power production on SCADA and the time series of power forecasting data including newly created forecasting data is shared for Wind Farm A. As seen in below table, there is relatively improvement on accuracy of power forecasting for Wind Farm A based on the methodology mentioned along this study. Highlighted part in the table shows best relationship between the time series of power forecasting and real power production.

Table 6.1 Different Time Horizon Based Relationship between Power Forecasting and Production for Wind Farm A

Wind Farm A	**R-square (Hourly)**	**R-square (Daily)**	**R-square (Spring)**	**R-square (Summer)**	**R-square (Autumn)**	**R-square (Winter)**
Forecast Provider-1	0.64	0.74	0.85	**0.89**	0.77	0.63

cont. on next page

cont of table 6.1

Forecast Provider-2	0.70	0.83	0.85	0.87	**0.83**	0.76
Forecast Provider-3	0.69	0.80	0.79	0.87	0.79	0.73
Forecast Provider-4	0.69	0.81	**0.89**	0.86	0.82	0.70
New Forecast-Temperature	0.70	0.83	0.85	0.89	0.80	0.75
New Forecast-Turbulence	0.70	0.83	0.85	0.87	0.82	**0.76**
New Forecast-Combined	**0.71**	**0.84**	0.89	0.88	0.82	0.75

In Table 6.2 and Table 6.3, relationships based on R-squared value between the time series of real power production on SCADA and the time series of forecasting data including newly created forecasting data are shared for Wind Farm B and C, respectively.

Table 6.2 Different Time Horizon Based Relationship between Power Forecasting and Production for Wind Farm B

Wind Farm B	**R-square (Hourly)**	**R-square (Daily)**	**R-square (Spring)**	**R-square (Summer)**	**R-square (Autumn)**	**R-square (Winter)**
Forecast Provider-1	0.74	0.86	0.87	**0.93**	0.88	0.79
Forecast Provider-2	0.72	0.83	0.82	0.90	0.89	0.73
Forecast Provider-3	0.74	0.83	0.87	0.92	0.91	0.74
Forecast Provider-4	**0.75**	0.84	**0.89**	0.85	**0.92**	0.77
New Forecast-Temperature	0.74	0.85	0.88	0.92	0.91	0.75
New Forecast-Turbulence	0.73	0.85	0.85	0.92	0.91	0.77
New Forecast-Combined	0.74	**0.86**	0.85	0.90	0.91	**0.80**

According to Table 6.2 above, improvement on power forecasting, especially on daily basis, under favour of mentioned methodology in this study has been achieved and more accurate power forecasting for Wind Farm B during winter times could be provided by the help of this methodology.

Table 6.3 Different Time Horizon Based Relationship between Power Forecasting and Production for Wind Farm C

Wind Farm C	**R-square (Hourly)**	**R-square (Daily)**	**R-square (Spring)**	**R-square (Summer)**	**R-square (Autumn)**	**R-square (Winter)**
Forecast Provider-1	0.74	0.86	0.84	0.87	0.89	0.82

cont. on next page

cont of table 6.3

Forecast Provider-2	0.79	0.90	0.86	0.88	0.91	**0.90**
Forecast Provider-3	0.78	0.89	0.85	0.83	0.92	0.89
Forecast Provider-4	0.79	0.90	**0.88**	0.87	0.93	0.88
New Forecast-Temperature	**0.79**	**0.91**	0.87	**0.89**	0.93	0.89
New Forecast-Turbulence	0.79	0.90	0.88	0.87	**0.93**	0.88
New Forecast-Combined	0.79	0.91	0.86	0.88	0.93	0.89

According to Table 6.3, any improvement on power forecasting by the help of the mentioned methodology could not be achieved for Wind Farm C so the wind farm is not considered to be exposure to low site temperature and high turbulence intensity unlike Wind Farm A and B. It can be related to non-complexity of the site.

To check improvement mathematically, daily power forecasting based on new combined method for Wind Farm B can be compared with forecast provider 3. Even if the difference depending upon determination coefficient (R2) seems to be small, using new combined method will provide accurate power forecasting up to about 4 GWh/a for Wind Farm B. Numerical values of the improvement on short-term wind power forecasting on daily basis by the help newly developed combined method are shared in Table 6.4. Highlighted values indicate the maximum of improvement value as GWh/a while using newly developed combined method instead of forecast providers. Approximately 10 GWh/a improvement for Wind Farm A by replacing forecast provider 1 with new combined method and approximately 5 GWh/a improvement for Wind Farm B by replacing forecast provider 2 with new combined method were provided on daily basis.

Table 6.4 Improvement in Wind Power Forecasting on Daily Basis

Hourly Basis Profit (versus)	New Forecast Combined Wind Farm A (GWh/a)	New Forecast Combined Wind Farm B (GWh/a)
Forecast Provider-1	**9.8**	1.3
Forecast Provider-2	1.1	**4.7**
Forecast Provider-3	2.0	4.2
Forecast Provider-4	1.6	3.4

Up to this point, it was explained how the accuracy of short-term power forecasting provided by forecast providers was increased by considering atmospheric effects such as temperature and turbulence. In section 5.1. , power forecasting data was grouped based on site temperature and afterwards in section 5.2. , the forecasting data was grouped by turbulence intensity due to the fact that real power production could be had some losses due to these kinds of atmospheric phenomena. However, effects of these two phenomena on the power forecasting are expected to occur at the same time in a single wind farm. For this reason, the power forecasting data was grouped in section 5.3. by considering both site temperature and turbulence intensity together. In order for selecting best forecast provider, grouped power forecasting data was compared with real power production on SCADA within same time interval. Based on grouped forecasting data by both temperature and turbulence intensity, findings regarding some improvement on the forecasts in comparison to other forecast providers were shared in Table 6.5 for all wind farms.

Table 6.5 Comparison of Power Forecasting and Production Based on Cumulative Sum for All Wind Farms (Scaled by SCADA Power Production)

	Wind Farm A [MWh/a]	**Wind Farm B [MWh/a]**	**Wind Farm C [MWh/a]**
SCADA Power Production	1.000	1.000	1.000
Forecast Provider 1	0.877	1.143	1.011
Forecast Provider 2	0.988	1.087	1.017
Forecast Provider 3	1.001	1.022	1.010
Forecast Provider 4	0.935	1.071	1.006
Reference Power Production	1.093	1.166	1.033
New Forecast-Temperature	0.982	1.028	1.011
New Forecast-Turbulence	0.991	0.986	1.004
New Forecast-Combined	0.974	1.019	1.004
Metered Power Production	0.989	0.984	0.986

Reference power production and metered power production were added the same table above. Although, newly developed methodology mentioned in this study has outperformed in terms of accuracy of power forecasting compared to some forecast providers, this kind of comparison based on cumulative sum of annual production like in Table 6.5 would not be sufficient due to the fact that short-term power forecasting on hourly and daily basis are generally carried out in electricity trading market. Therefore, this study has focused on improvement of hourly daily basis wind power forecasting.

Table 6.6 Power Forecasting Review at Different Time Horizon

Period	Best Forecast Provider for Wind Farm A	Best Forecast Provider for Wind Farm B	Best Forecast Provider for Wind Farm C
Hourly	New Forecast-Combined	Forecast Provider-4	New Forecast-Temperature
Daily	New Forecast-Combined	New Forecast-Combined	New Forecast-Temperature
Spring	Forecast Provider-4	Forecast Provider-4	Forecast Provider-4
Summer	Forecast Provider-1	Forecast Provider-1	New Forecast-Temperature
Autumn	Forecast Provider-2	Forecast Provider-2	New Forecast-Turbulence
Winter	New Forecast-Turbulence	New Forecast-Combined	Forecast Provider-2

Hourly power forecasting data was converted as daily and was compared with real power production on SCADA. On daily period, newly developed method considering site temperature and turbulence intensity of Wind Farm A and B has been succeeded to outperform other power forecasting provided by forecast providers. In comparison to turbulence condition of Wind Farm A and B, Wind Farm C may not be suffered from turbulent wind condition during power production operation. Thus, the method considering only site temperature should be used for Wind Farm C in order to increase forecasting accuracy. Based on hourly forecasting accuracy, the newly developed method can also be used for Wind Farm A and C. For Wind Farm B, forecast provider 4 was outperformed compared to the newly developed method, however, the forecasting accuracy that can be related to the determination coefficient in Table 6.2 are very close to combined method that was newly developed. During spring, forecast provider 4 should be preferred for more accurate forecasting. For summer and autumn, forecast provider 1 and forecast provider 2 can be used for Wind Farm A and B, respectively and newly developed method can be considered and included in order for increasing forecasting accuracy for Wind Farm C. At winter times, it is hard to provide accurate power forecasting due the fact that icing conditions at the sites, especially in Wind Farm A and B, are expected to occur and many forecast providers suffer from this phenomenon. Most forecast provider cannot take power production losses due to icing into account. Modelling and forecasting of these kinds of losses are really challenging due to uncertainty of environmental conditions, however, improvement on power forecasting for the wind farms can be performed by including newly developed method mentioned along the study.

CHAPTER 7

CONCLUSION

Wind power forecasting is very challenging due to stochastic nature of atmosphere. In literature, plenty of wind power forecasting models have been developed to improve forecasting accuracy. Most of the forecasting models have been succeeded to provide better results compared to forecasting models on previous studies, however, there is no specific baseline model for benchmarking. The models are generally site-specific, and their applicability is not simple for another wind farm sites. Instead of creating new site-specific forecasting model to the literature, existing forecasting models, which are different NWP models from four forecast providers, have been included to this study on the purpose of improving forecasting accuracy. Schematic presentation of the study can be seen in Figure 7.1.

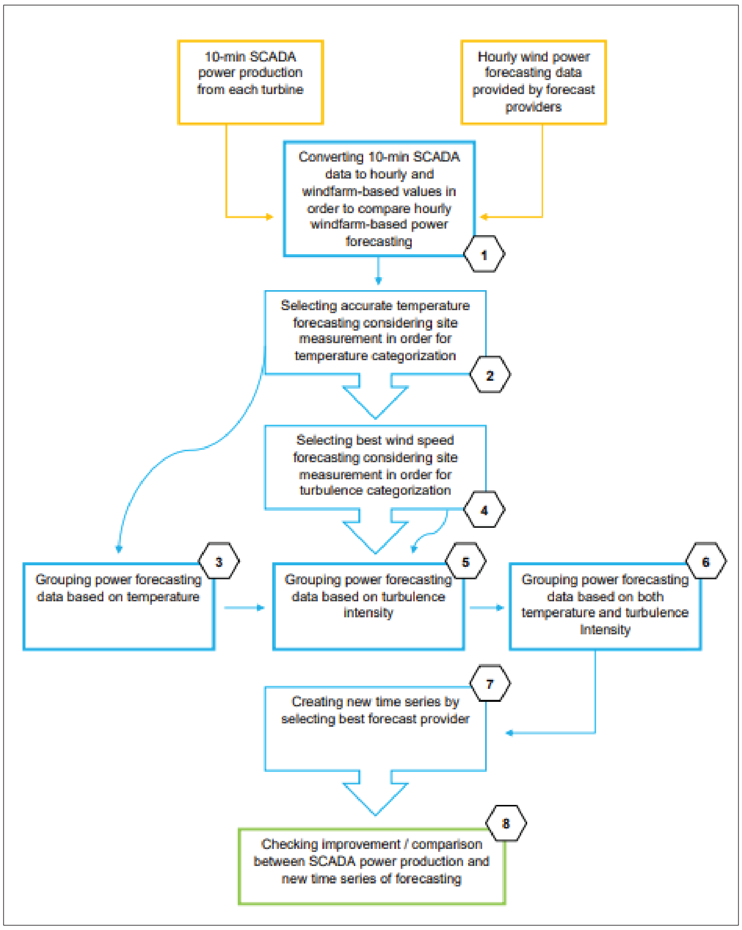

Figure 7.1 Schematic Presentation of the Study

The method mentioned along the study is not a site-specific and can be applicable easefully for another wind farm sites. It should not be forgotten that the method is not a new forecasting model, it can be considered as a cost-effective engineering approach to decrease deviation in power production forecasting. Crucial part of the method is detecting parameters that cause power forecasting errors. Root causes of power production losses that can affect power forecasting accuracy were investigated in the study. There are several parameters regarding power production losses such as turbine availability, grid curtailment, maintenance period, wind direction, icing, turbulence, and extreme wind speed. Icing and turbulence that have significant effect on power production losses are included to this study. Icing and turbulence phenomena can be related to ambient temperature and wind speed of the wind farm site, respectively. As mentioned earlier, this study considers site temperature and turbulence only because these two parameters are responsible for most of the power production losses that leads to power forecasting errors. If the site does not suffer from power production losses originating from temperature or turbulence, method in the study will not improve power forecasting accuracy. In this case, other parameter, wind turbine wake that is also responsible for significant amount of power production losses, should be included to develop the method explained in this study. Due to limited time frame of the study, most important parameters, which are temperature and turbulence, were selected and short-term wind power forecasting has been improved correspondingly. Wake parameter can be examined for further investigation. For future studies, other parameters in addition temperature and turbulence can be included to ensemble learning algorithms, which enable us rapid computation with a lot of variables. In this way, many additional parameters can be studied to observe their effects on wind power forecasting.

REFERENCES

(1) Wu, Y. K.; Hong, J. S. A Literature Review of Wind Forecasting Technology in the World. In *2007 IEEE Lausanne POWERTECH, Proceedings*; 2007. https://doi.org/10.1109/PCT.2007.4538368.

(2) Hanifi, S.; Liu, X.; Lin, Z.; Lotfian, S. A Critical Review of Wind Power Forecasting Methods-Past, Present and Future. *Energies*. MDPI AG August 1, 2020. https://doi.org/10.3390/en13153764.

(3) Soman, S. S.; Zareipour, H.; Malik, O.; Mandal, P. A Review of Wind Power and Wind Speed Forecasting Methods with Different Time Horizons. In *North American Power Symposium 2010, NAPS 2010*; 2010. https://doi.org/10.1109/NAPS.2010.5619586.

(4) Lerner, J.; Grundmeyer, M.; Garvert, M. The Importance of Wind Forecasting. *Renew. Energy Focus* **2009**, *10* (2). https://doi.org/10.1016/s1755-0084(09)70092-4.

(5) Karık, F.; Sözen, A.; İzgeç, M. M. Rüzgâr Gücü Tahminlerinin Önemi: Türkiye Elektrik Piyasasında Bir Uygulama. *J. Polytech.* **2017**. https://doi.org/10.2339/politeknik.369038.

(6) Pelletier, F.; Masson, C.; Tahan, A. Wind Turbine Power Curve Modelling Using Artificial Neural Network. *Renew. Energy* **2016**, *89*. https://doi.org/10.1016/j.renene.2015.11.065.

(7) Foley, A. M.; Leahy, P. G.; Marvuglia, A.; McKeogh, E. J. Current Methods and Advances in Forecasting of Wind Power Generation. *Renewable Energy*. 2012. https://doi.org/10.1016/j.renene.2011.05.033.

(8) Chang, W.-Y. A Literature Review of Wind Forecasting Methods. *J. Power Energy Eng.* **2014**, *02* (04). https://doi.org/10.4236/jpee.2014.24023.

(9) Hyndman, R. J.; Athanasopoulos, G. Forecasting: Principles and Practice. *Princ. Optim. Des.* **2018**.

(10) Cadenas, E.; Rivera, W. Wind Speed Forecasting in Three Different Regions of Mexico, Using a Hybrid ARIMA-ANN Model. *Renew. Energy* **2010**, *35* (12). https://doi.org/10.1016/j.renene.2010.04.022.

(11) Chen, Q.; Folly, K. A. Wind Power Forecasting. *IFAC-PapersOnLine* **2018**, *51* (28). https://doi.org/10.1016/j.ifacol.2018.11.738.

(12) Zhao, Y.; Ye, L.; Li, Z.; Song, X.; Lang, Y.; Su, J. A Novel Bidirectional Mechanism Based on Time Series Model for Wind Power Forecasting. *Appl. Energy* **2016**, *177*. https://doi.org/10.1016/j.apenergy.2016.03.096.

(13) Gallego, C.; Pinson, P.; Madsen, H.; Costa, A.; Cuerva, A. Influence of Local Wind Speed and Direction on Wind Power Dynamics - Application to Offshore Very Short-Term Forecasting. *Appl. Energy* **2011**, *88* (11). https://doi.org/10.1016/j.apenergy.2011.04.051.

(14) Durán, M. J.; Cros, D.; Riquelme, J. Short-Term Wind Power Forecast Based on ARX Models. *J. Energy Eng.* **2007**, *133* (3). https://doi.org/10.1061/(asce)0733-9402(2007)133:3(172).

(15) Lin, Z.; Liu, X. Wind Power Forecasting of an Offshore Wind Turbine Based on High-Frequency SCADA Data and Deep Learning Neural Network. *Energy* **2020**, *201*. https://doi.org/10.1016/j.energy.2020.117693.

(16) Sideratos, G.; Hatziargyriou, N. D. Probabilistic Wind Power Forecasting Using Radial Basis Function Neural Networks. *IEEE Trans. Power Syst.* **2012**, *27* (4). https://doi.org/10.1109/TPWRS.2012.2187803.

(17) Zhang, J.; Yan, J.; Infield, D.; Liu, Y.; Lien, F. sang. Short-Term Forecasting and Uncertainty Analysis of Wind Turbine Power Based on Long Short-Term Memory Network and Gaussian Mixture Model. *Appl. Energy* **2019**, *241*. https://doi.org/10.1016/j.apenergy.2019.03.044.

(18) Bilal, B.; Ndongo, M.; Adjallah, K. H.; Sava, A.; Kebe, C. M. F.; Ndiaye, P. A.; Sambou, V. Wind Turbine Power Output Prediction Model Design Based on Artificial Neural Networks and Climatic Spatiotemporal Data. In *Proceedings of the IEEE International Conference on Industrial Technology*; 2018; Vol. 2018-February. https://doi.org/10.1109/ICIT.2018.8352329.

(19) -, W. C. Application of Back Propagation Neural Network for Wind Power Generation Forecasting. *Int. J. Digit. Content Technol. its Appl.* **2013**, *7* (4). https://doi.org/10.4156/jdcta.vol7.issue4.61.

(20) Singh, S.; Bhatti, T. S.; Kothari, D. P. Wind Power Estimation Using Artificial Neural Network. *J. Energy Eng.* **2007**, *133* (1). https://doi.org/10.1061/(asce)0733-9402(2007)133:1(46).

(21) Chen, N.; Qian, Z.; Nabney, I. T.; Meng, X. Wind Power Forecasts Using Gaussian Processes and Numerical Weather Prediction. *IEEE Trans. Power Syst.* **2014**, *29* (2). https://doi.org/10.1109/TPWRS.2013.2282366.

(22) Jung, J.; Broadwater, R. P. Current Status and Future Advances for Wind Speed and Power Forecasting. *Renewable and Sustainable Energy Reviews*. 2014. https://doi.org/10.1016/j.rser.2013.12.054.

(23) Focken, U.; Lange, M.; Waldl, H.-P. H.-P. Previento - A Wind Power Prediction System with an Innovative Upscaling Algorithm. *Proc. Eur. Wind Energy Conf.* **2001**, No. January.

(24) Lin, Z.; Liu, X.; Collu, M. Wind Power Prediction Based on High-Frequency SCADA Data along with Isolation Forest and Deep Learning Neural Networks. *Int. J. Electr. Power Energy Syst.* **2020**, *118*. https://doi.org/10.1016/j.ijepes.2020.105835.

(25) De Giorgi, M. G.; Ficarella, A.; Tarantino, M. Assessment of the Benefits of Numerical Weather Predictions in Wind Power Forecasting Based on Statistical Methods. *Energy* **2011**, *36* (7). https://doi.org/10.1016/j.energy.2011.05.006.

(26) Pieralli, S.; Ritter, M.; Odening, M. Efficiency of Wind Power Production and Its Determinants. *Energy* **2015**, *90*. https://doi.org/10.1016/j.energy.2015.07.055.

(27) Gylling, N.; Mortensen, N. G.; Petersen, E. L. Influence of Topographical Input Data on the Accuracy of Wind Flow Modelling in Complex Terrain. *Eur. Conf.* **1997**.

(28) U.S. Geological Survey. Shuttle Radar Topography Mission 1 Arc-Second Global. *U.S. Geological Survey*. 2017.

(29) Bingöl, F. Rüzgar Enerji Sistemleri İçin Hava Yoğunluğunun Hesaplanması. *J. Polytech.* **2018**. https://doi.org/10.2339/politeknik.385523.

(30) Sun, P.; Li, J.; Wang, C.; Lei, X. A Generalized Model for Wind Turbine Anomaly Identification Based on SCADA Data. *Appl. Energy* **2016**, *168*. https://doi.org/10.1016/j.apenergy.2016.01.133.

(31) Montgomery, D. C.; Runger, G. C. Applied Statistics and Probability for Engineers. *Eur. J. Eng. Educ.* **1994**, *19* (3). https://doi.org/10.1080/03043799408928333.

(32) Bredesen, R. E.; Vindteknikk, K.; Cattin, R.; Clausen, N. E.; Davis, N.; Jordaens, P. J.; Khadiri-Yazami, Z.; Klintstrom, R.; Krenn, A.; Lehtomaki, V.; Ronsten, G.; Wadham-Gagnon, M.; Wichman, H. IEA Wind Task 19: Wind Energy Projects in Cold Climates. *IEA Wind* **2017**.

(33) IEC. IEC 61400-12-1:2017 Edition 2.0 Wind Energy Generation Systems – Power Performance Measurements of Electricity Producing Wind Turbines. *Int. Stand.*

2017.
(34) ISO. Standard Atmosphere, ISO 2533:1975. *Int. Stand. Organ.* **1975**, *2533*.

APPENDICES

APPENDIX A

R CODE

```
#WIND FARM A
X0_windfarmA_hourly_Autosaved_              <-            read_excel("E:/Scada
Data/windfarmA_windfarmB_windfarmC/windfarmARES/0      -      windfarmA_hourly
(Autosaved).xlsx",  sheet  =  "All_2019",  col_types  =  c("date",  "numeric",
"numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",
"numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",
"numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",  "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric"))

windfarmA19<- X0_windfarmA_hourly_Autosaved_
View(windfarmA19)

names(windfarmA19)[1]<- "date"
names(windfarmA19)[2]<- "scadapower"
names(windfarmA19)[3]<- "reference"
names(windfarmA19)[4]<- "meter"
names(windfarmA19)[5]<- "provider1power"
names(windfarmA19)[6]<- "provider2power"
names(windfarmA19)[7]<- "provider3power"
names(windfarmA19)[8]<- "provider4power"
names(windfarmA19)[16]<- "tempprovider1"
names(windfarmA19)[13]<- "turb"
names(windfarmA19)[21]<- "tempprovider3"
names(windfarmA19)[26]<- "tempprovider2"
summary(windfarmA19$tempprovider1)

length(windfarmA1999$provider3power)
summary(windfarmA1999$tempprovider1)
windfarmA1999   <-   filter(windfarmA19,   windfarmA19$scadapower   >=   0   &
windfarmA19$provider2power >= 0)

ag111<- aggregate(windfarmA1999$provider1power, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))

ag112<- aggregate(windfarmA1999$provider2power, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))

ag113<- aggregate(windfarmA1999$provider3power, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))

ag114<- aggregate(windfarmA1999$provider4power, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))

ag000<- aggregate(windfarmA1999$scadapower, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))

agRRR<- aggregate(windfarmA1999$reference, FUN=sum, na.rm=TRUE,
by=list(cut(windfarmA1999$tempprovider1, breaks=c(seq(from= -10, to = 32, by
=3)), include.lowest=F)))
```

```
ag_finalm <-
cbind.data.frame(ag000$Group.1,ag000$x,ag111$x,ag112$x,ag113$x,ag114$x,agRRR$x
)
View(ag_finalm)

#reference
plot(ag_finalm$`ag000$Group.1`, ag_finalm$`agRRR$x`/ag_finalm$`agRRR$x`, lwd=1,
xlab="Temperature provider1", ylab="Power Forecast (Scaled based on Reference
Power)", ylim=c(0,3.5),main="windfarmA RES (Reference Power)")
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag111$x`/ag_finalm$`agRRR$x`,col="cyan", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag112$x`/ag_finalm$`agRRR$x`,col="green", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag113$x`/ag_finalm$`agRRR$x`,col="blue", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag114$x`/ag_finalm$`agRRR$x`,col="purple", lwd=2)
legend("topright",legend=c("Reference Power",
"provider1","provider2","provider3","provider4"),text.col=c("black","cyan","gr
een","blue","purple"))

#scada
plot(ag_finalm$`ag000$Group.1`, ag_finalm$`ag000$x`/ag_finalm$`ag000$x`, lwd=1,
xlab="Temperature°C - Forecast Provider 1", ylab="Power Forecast (Scaled based
on Scada Power)", ylim=c(0,3.5),main="WIND FARM A (Scada Power)")
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag111$x`/ag_finalm$`ag000$x`,col="cyan", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag112$x`/ag_finalm$`ag000$x`,col="green", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag113$x`/ag_finalm$`ag000$x`,col="blue", lwd=2)
lines(ag_finalm$`ag000$Group.1`,
ag_finalm$`ag114$x`/ag_finalm$`ag000$x`,col="purple", lwd=2)
legend("topleft",legend=c("Scada Power", "Forecast Provider 1","Forecast
Provider 2","Forecast Provider 3","Forecast Provider
4"),text.col=c("black","cyan","green","blue","purple"))

#3 derece bin
mm1<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > -10 &
windfarmA1999$tempprovider1 <= -7)
mm1_yk <- select(mm1, date, provider3power,scadapower,reference)
write.table(mm1_yk,file="temp_1.txt",row.names=F,col.names = T, sep="\t")

mm2<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > -7 &
windfarmA1999$tempprovider1 <= -4)
mm2_yk <- select(mm2, date, provider2power,scadapower,reference)
write.table(mm2_yk,file="temp_2.txt",row.names=F,col.names = T, sep="\t")

mm3<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > -4 &
windfarmA1999$tempprovider1 <= -1)
mm3_yk <- select(mm3, date, provider3power,scadapower,reference)
write.table(mm3_yk,file="temp_3.txt",row.names=F,col.names = T, sep="\t")

mm4<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > -1 &
windfarmA1999$tempprovider1 <= 2)
mm4_yk <- select(mm4, date, provider4power,scadapower,reference)
write.table(mm4_yk,file="temp_4.txt",row.names=F,col.names = T, sep="\t")

mm5<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 2 &
windfarmA1999$tempprovider1 <= 5)
mm5_yk <- select(mm5, date, provider2power,scadapower,reference)
write.table(mm5_yk,file="temp_5.txt",row.names=F,col.names = T, sep="\t")

mm6<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 5 &
windfarmA1999$tempprovider1 <= 11)
mm6_yk <- select(mm6, date, provider3power,scadapower,reference)
write.table(mm6_yk,file="temp_6.txt",row.names=F,col.names = T, sep="\t")
```

```
mm7<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 11 &
windfarmA1999$tempprovider1 <= 14)
mm7_yk <- select(mm7, date, provider2power,scadapower,reference)
write.table(mm7_yk,file="temp_7.txt",row.names=F,col.names = T, sep="\t")

mm8<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 14 &
windfarmA1999$tempprovider1 <= 23)
mm8_yk <- select(mm8, date, provider1power,scadapower,reference)
write.table(mm8_yk,file="temp_8.txt",row.names=F,col.names = T, sep="\t")

mm9<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 23 &
windfarmA1999$tempprovider1 <= 26)
mm9_yk <- select(mm9, date, provider3power,scadapower,reference)
write.table(mm9_yk,file="temp_9.txt",row.names=F,col.names = T, sep="\t")

mm10<-windfarmA1999 %>% filter(windfarmA1999$tempprovider1 > 26 &
windfarmA1999$tempprovider1 <= 32)
mm10_yk <- select(mm10, date, provider2power,scadapower,reference)
write.table(mm10_yk,file="temp_10.txt",row.names=F,col.names = T, sep="\t")

sum(length(mm1$provider3power))+
sum(length(mm2$provider2power))+
sum(length(mm3$provider3power))+
sum(length(mm4$provider4power))+
sum(length(mm5$provider2power))+
sum(length(mm6$provider3power))+
sum(length(mm7$provider2power))+
sum(length(mm8$provider1power))+
sum(length(mm9$provider3power))+
sum(length(mm10$provider2power))

sum(mm1[2])+sum(mm2[2])+sum(mm3[2])+sum(mm4[2])+sum(mm5[2])+sum(mm6[2])+sum(mm
7[2])+sum(mm8[2])+sum(mm9[2])+sum(mm10[2])
sum(mm1[3])+sum(mm2[3])+sum(mm3[3])+sum(mm4[3])+sum(mm5[3])+sum(mm6[3])+sum(mm
7[3])+sum(mm8[3])+sum(mm9[3])+sum(mm10[3])
sum(mm1[5])+sum(mm2[5])+sum(mm3[5])+sum(mm4[5])+sum(mm5[5])+sum(mm6[5])+sum(mm
7[5])+sum(mm8[5])+sum(mm9[5])+sum(mm10[5])
sum(mm1[6])+sum(mm2[6])+sum(mm3[6])+sum(mm4[6])+sum(mm5[6])+sum(mm6[6])+sum(mm
7[6])+sum(mm8[6])+sum(mm9[6])+sum(mm10[6])
sum(mm1[7])+sum(mm2[7])+sum(mm3[7])+sum(mm4[7])+sum(mm5[7])+sum(mm6[7])+sum(mm
7[7])+sum(mm8[7])+sum(mm9[7])+sum(mm10[7])
sum(mm1[8])+sum(mm2[8])+sum(mm3[8])+sum(mm4[8])+sum(mm5[8])+sum(mm6[8])+sum(mm
7[8])+sum(mm8[8])+sum(mm9[8])+sum(mm10[8])
sum(mm1[4])+sum(mm2[4])+sum(mm3[4])+sum(mm4[4])+sum(mm5[4])+sum(mm6[4])+sum(mm
7[4])+sum(mm8[4])+sum(mm9[4])+sum(mm10[4])

######TURBULENCE###############
#sum(windfarmAturb$provider1power);sum(windfarmAturb$provider2power);sum(windf
armAturb$provider3power);sum(windfarmAturb$provider4power);sum(windfarmAturb$s
cadapower)

windfarmAturb1<-windfarmA1999 %>% filter(windfarmA1999$provider1_WS > 0 &
windfarmA1999$provider1_WS <= 3 & windfarmA1999$tempprovider1 > -10 &
windfarmA1999$tempprovider1 <= 32)
windfarmAturb1_yk <- select(windfarmAturb1, date,
provider4power,scadapower,reference)
write.table(windfarmAturb1_yk,file="turb1.txt",row.names=F,col.names = T,
sep="\t")

windfarmAturb2<-windfarmA1999 %>% filter(windfarmA1999$provider1_WS > 3 &
windfarmA1999$provider1_WS <= 6 & windfarmA1999$tempprovider1 > -10 &
windfarmA1999$tempprovider1 <= 32)
windfarmAturb2_yk <- select(windfarmAturb2, date,
provider2power,scadapower,reference)
write.table(windfarmAturb2_yk,file="turb2.txt",row.names=F,col.names = T,
sep="\t")
```

```
windfarmAturb3<-windfarmA1999 %>% filter(windfarmA1999$provider1_WS > 6 &
windfarmA1999$provider1_WS <= 12 & windfarmA1999$tempprovider1 > -10 &
windfarmA1999$tempprovider1 <= 32)
windfarmAturb3_yk <- select(windfarmAturb3, date,
provider2power,scadapower,reference)
write.table(windfarmAturb3_yk,file="turb3.txt",row.names=F,col.names = T,
sep="\t")

windfarmAturb4<-windfarmA1999 %>% filter(windfarmA1999$provider1_WS > 12 &
windfarmA1999$tempprovider1 > -10 & windfarmA1999$tempprovider1 <= 32)
windfarmAturb4_yk <- select(windfarmAturb4, date,
provider1power,scadapower,reference)
write.table(windfarmAturb4_yk,file="turb4.txt",row.names=F,col.names = T,
sep="\t")

#WIND FARM B
library(readxl)
X1_windfarmB_2018_2019 <- read_excel("E:/Scada
Data/windfarmA_windfarmB_windfarmC/windfarmBRES/1_windfarmB2018_2019.xlsx",
sheet = "All_2019", col_types = c("date", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric"))

windfarmB199 <- X1_windfarmB_2018_2019

names(windfarmB199)[1]<- "date"
names(windfarmB199)[2]<- "scadapower"
names(windfarmB199)[3]<- "reference"
names(windfarmB199)[4]<- "meter"
names(windfarmB199)[5]<- "provider1power" #; provider1 <- f1 %>%
filter(f1$provider1power < 50)
names(windfarmB199)[6]<- "provider2power"
names(windfarmB199)[7]<- "provider3power"
names(windfarmB199)[8]<- "provider4power"
names(windfarmB199)[16]<- "tempprovider1"
names(windfarmB199)[13]<- "turb"
names(windfarmB199)[21]<- "tempprovider3"
names(windfarmB199)[26]<- "tempprovider2"

windfarmB1999 <- filter(windfarmB199, windfarmB199$scadapower >= 0 &
windfarmB199$provider2power >= 0, windfarmB199$provider1power < 50)
View(windfarmB1999)

summary(windfarmB1999$tempprovider1)

ag1<- aggregate(windfarmB1999$provider1power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1, breaks=c(seq(from= -5, to = 34, by
=3)), include.lowest=F)))
View(ag1)

ag2<- aggregate(windfarmB1999$provider2power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1, breaks=c(seq(from= -5, to = 34, by
=3)), include.lowest=F)))
View(ag2)

ag3<- aggregate(windfarmB1999$provider3power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1, breaks=c(seq(from= -5, to = 34, by
=3)), include.lowest=F)))
View(ag3)

ag4<- aggregate(windfarmB1999$provider4power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1, breaks=c(seq(from= -5, to = 34, by
=3)), include.lowest=F)))
View(ag4)
```

```
ag0<-      aggregate(windfarmB1999$scadapower,      FUN=sum,      na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1,breaks=c(seq(from= -5, to = 34, by =3)),
include.lowest=F)))
View(ag0)

agR<-      aggregate(windfarmB1999$reference,      FUN=sum,      na.rm=FALSE,
by=list(cut(windfarmB1999$tempprovider1,breaks=c(seq(from= -5, to = 34, by =3)),
include.lowest=F)))
View(agR)

ag_final <- cbind.data.frame(ag0$Group.1,ag0$x,ag1$x,ag2$x,ag3$x,ag4$x,agR$x)
View(ag_final)
write.csv(ag_final, file="ag_final.csv")

#######SCADA POWER######
plot(ag_final$`ag0$Group.1`,ag_final$`ag0$x`/ag_final$`ag0$x`,          lwd=1,
xlab="Temperature°C - Forecast Provider 1", ylab="Power Forecast (Scaled based
on Scada Power)", ylim=c(0,3),main="WIND FARM B (Scada Power)")
#plot(ag0,col="black", type="l")
#axis(side=1, at=seq(-5,34,1))
#lines(ag0$x/ag0$x,col="black", lwd=2)
lines(ag_final$`ag0$Group.1`,ag_final$`ag1$x`/ag_final$`ag0$x`,col="cyan",
lwd=2)
lines(ag_final$`ag0$Group.1`,ag_final$`ag2$x`/ag_final$`ag0$x`,col="green",
lwd=2)
lines(ag_final$`ag0$Group.1`,ag_final$`ag3$x`/ag_final$`ag0$x`,col="blue",
lwd=2)
lines(ag_final$`ag0$Group.1`,ag_final$`ag4$x`/ag_final$`ag0$x`,col="purple",
lwd=2)
legend("topleft",legend=c("Scada   Power",   "Forecast   Provider   1","Forecast
Provider       2","Forecast       Provider       3","Forecast       Provider
4"),text.col=c("black","cyan","green","blue","purple"))

# 3 derece bin
aa<-windfarmB1999    %>%    filter(windfarmB1999$tempprovider1    >    -5    &
windfarmB1999$tempprovider1 <= 4)
aa1 <- select(aa,date,provider2power,scadapower,reference)
write.csv(aa1,file="provider2 -5 4.csv")

bb<-windfarmB1999    %>%    filter(windfarmB1999$tempprovider1    >    4    &
windfarmB1999$tempprovider1 <= 16)
bb1 <- select(bb,date, provider3power,scadapower,reference)
write.csv(bb1,file="provider3 4 16.csv")

cc<-windfarmB1999    %>%    filter(windfarmB1999$tempprovider1    >    16    &
windfarmB1999$tempprovider1 <= 22)
cc1 <- select(cc, date,provider4power, scadapower,reference)
write.csv(cc1,file="provider4 16 22.csv")

dd<-windfarmB1999    %>%    filter(windfarmB1999$tempprovider1    >    22    &
windfarmB1999$tempprovider1 <= 28)
dd1 <- select(dd,date, provider3power,scadapower,reference)
write.csv(dd1,file="provider3 22 28.csv")

ee<-windfarmB1999    %>%    filter(windfarmB1999$tempprovider1    >    28    &
windfarmB1999$tempprovider1 <= 34)
ee1 <- select(ee, date,provider2power,scadapower,reference)
write.csv(ee1,file="provider2 28 34.csv")

sum(aa1$provider2power)+
sum(bb1$provider3power)+
sum(cc1$provider4power)+
sum(dd1$provider3power)+
sum(ee1$provider2power)

sum(length(windfarmB1999$scadapower))
sum(windfarmB1999$reference)
sum(windfarmB1999$provider1power)
```

```
sum(windfarmB1999$provider2power)
sum(windfarmB1999$provider3power)
sum(windfarmB1999$provider4power)
sum(windfarmB1999$provider3er)
summary(windfarmB1999$provider1_WS)

wsti<-       aggregate(windfarmB1999$turb,       FUN=mean,       na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks=c(seq(from= 0, to =20 , by =2)),
include.lowest=F)))
View(wsti)

plot(wsti$Group.1,  wsti$x,  xlab=" provider1  Wind  Speed  ",  ylab="Turbulence
Intensity", main="Turbulence Intensity based on provider1 GL Wind Speed [2 m/s
bin]")
#plot(windfarmB1999$provider1_WS,windfarmB1999$turb, type="p" ,xlab=" provider1
Wind Speed ", ylab="Turbulence Intensity", main="Turbulence Intensity based on
provider1 GL Wind Speed [1 m/s bin]")
lines(wsti$Group.1 ,wsti$x, lwd=3, col="red")
abline(h=0.12, col="red", lwd=1)
abline(v=15, col="blue", lwd=1)

summary(windfarmB1999$provider1_WS)
breaks<-c(0,3,6,12,20)
wsti1<-    aggregate(windfarmB1999$provider1power,    FUN=sum,    na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))
View(wsti1)

wsti2<- aggregate(windfarmB1999$provider2power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))
View(wsti2)

wsti3<- aggregate(windfarmB1999$provider3power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))
View(wsti3)

wsti4<- aggregate(windfarmB1999$provider4power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))
View(wsti4)

wsti5<-     aggregate(windfarmB1999$scadapower,     FUN=sum,     na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))
View(wsti5)

wsti6<-     aggregate(windfarmB1999$reference,     FUN=sum,     na.rm=FALSE,
by=list(cut(windfarmB1999$provider1_WS, breaks, include.lowest=F)))

View(wsti6)

windfarmB_ti<-cbind.data.frame(wsti1$Group.1,
wsti1$x,wsti2$x,wsti3$x,wsti4$x,wsti5$x,wsti6$x)
View(windfarmB_ti)


turba<-windfarmB1999    %>%    filter(windfarmB1999$provider1_WS    >    0    &
windfarmB1999$provider1_WS <= 3)
turba1 <- select(turba,date, provider3power,scadapower,reference)
write.csv(turba1,file="turb_03.csv")

turbb<-windfarmB1999    %>%    filter(windfarmB1999$provider1_WS    >    3    &
windfarmB1999$provider1_WS <= 6)
turbb1 <- select(turbb,date, provider1power,scadapower,reference)
write.csv(turbb1,file="turb_36.csv")

turbc<-windfarmB1999    %>%    filter(windfarmB1999$provider1_WS    >    6    &
windfarmB1999$provider1_WS <= 12)
turbc1 <- select(turbc,date, provider3power,scadapower,reference)
write.csv(turbc1,file="turb_612.csv")
```

```
turbd<-windfarmB1999 %>% filter(windfarmB1999$provider1_WS > 12)
turbd1 <- select(turbd,date, provider2power,scadapower,reference)
write.csv(turbd1,file="turb_13.csv")

sum(turbd$provider4power)
sum(length(turba$provider3power))+
sum(length(turbb$provider1power))+sum(length(turbc$provider3power))+sum(length
(turbd$provider2power))
sum(turba$provider3power)+sum(turbb$provider1power)+sum(turbc$provider3power)+
sum( turbd$provider2power)

#WIND FARM C
X0_windfarmC_hourly <- read_excel("E:/Scada
Data/windfarmA_windfarmB_windfarmC/windfarmCRES/0-windfarmC_hourly.xlsx",
sheet = "All_2019", col_types = c("date", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "numeric", "numeric"))

windfarmC19 <- X0_windfarmC_hourly
View(windfarmC19)

names(windfarmC19)[1]<- "date"
names(windfarmC19)[2]<- "scadapower"
names(windfarmC19)[3]<- "reference"
names(windfarmC19)[4]<- "meter"
names(windfarmC19)[5]<- "provider1power"
names(windfarmC19)[6]<- "provider2power"
names(windfarmC19)[7]<- "provider3power"
names(windfarmC19)[8]<- "provider4power"
names(windfarmC19)[16]<- "tempprovider1"
names(windfarmC19)[13]<- "turb"
names(windfarmC19)[21]<- "tempprovider3"
names(windfarmC19)[26]<- "tempprovider2"

summary(windfarmC1999$tempprovider1)
length(windfarmC1999$provider1power)
windfarmC1999 <- filter(windfarmC19, windfarmC19$scadapower >= 0 ,
windfarmC19$provider2power >= 0 ) #, windfarmB199$provider1power < 50)

ag11<- aggregate(windfarmC1999$provider1power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

ag12<- aggregate(windfarmC1999$provider2power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

ag13<- aggregate(windfarmC1999$provider3power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

ag14<- aggregate(windfarmC1999$provider4power, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

ag00<- aggregate(windfarmC1999$scadapower, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

agRR<- aggregate(windfarmC1999$reference, FUN=sum, na.rm=FALSE,
by=list(cut(windfarmC1999$tempprovider1, breaks=c(seq(from= -4, to = 33, by
=3)), include.lowest=F)))

ag_finalb < -
cbind.data.frame(ag00$Group.1,ag00$x,ag11$x,ag12$x,ag13$x,ag14$x,agRR$x)
View(ag_finalb)
```

```
#reference
plot(ag_finalb$`ag00$Group.1`,  ag_finalb$`agRR$x`/ag_finalb$`agRR$x`,  lwd=1,
xlab="Temperature provider1", ylab="Power Forecast (Scaled based on Reference
Power)", ylim=c(0,3),main="windfarmC RES (Reference Power)")
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag11$x`/ag_finalb$`agRR$x`,col="cyan", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag12$x`/ag_finalb$`agRR$x`,col="green", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag13$x`/ag_finalb$`agRR$x`,col="blue", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag14$x`/ag_finalb$`agRR$x`,col="purple", lwd=2)
legend("topleft",legend=c("Reference                              Power",
"provider1","provider2","provider3","provider4"),text.col=c("black","cyan","gr
een","blue","purple"))

#scada
plot(ag_finalb$`ag00$Group.1`,  ag_finalb$`ag00$x`/ag_finalb$`ag00$x`,  lwd=1,
xlab="Temperature°C - Forecast Provider 1", ylab="Power Forecast (Scaled based
on Scada Power)", ylim=c(0,3),main="WIND FARM C (Scada Power)")
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag11$x`/ag_finalb$`ag00$x`,col="cyan", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag12$x`/ag_finalb$`ag00$x`,col="green", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag13$x`/ag_finalb$`ag00$x`,col="blue", lwd=2)
lines(ag_finalb$`ag00$Group.1`,
ag_finalb$`ag14$x`/ag_finalb$`ag00$x`,col="purple", lwd=2)
legend("topleft",legend=c("Scada   Power",   "Forecast   Provider   1","Forecast
Provider       2","Forecast       Provider       3","Forecast       Provider
4"),text.col=c("black","cyan","green","blue","purple"))

#3 derece bin
bb1<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    -4    &
windfarmC1999$tempprovider1 <= 5)
bb1_yk <- select(bb1, date, provider2power,scadapower,reference)
write.table(bb1_yk,file="temp1.txt",row.names=F,col.names = T, sep="\t")

bb2<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    5    &
windfarmC1999$tempprovider1 <= 8)
bb2_yk <- select(bb2, date, provider4power,scadapower,reference)
write.table(bb2_yk,file="temp2.txt",row.names=F,col.names = T, sep="\t")

bb3<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    8    &
windfarmC1999$tempprovider1 <= 14)
bb3_yk <- select(bb3, date, provider2power,scadapower,reference)
write.table(bb3_yk,file="temp3.txt",row.names=F,col.names = T, sep="\t")

bb4<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    14    &
windfarmC1999$tempprovider1 <= 20)
bb4_yk <- select(bb4, date, provider4power,scadapower,reference)
write.table(bb4_yk,file="temp4.txt",row.names=F,col.names = T, sep="\t")

bb5<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    20    &
windfarmC1999$tempprovider1 <= 23)
bb5_yk <- select(bb5, date, provider2power,scadapower,reference)
write.table(bb5_yk,file="temp5.txt",row.names=F,col.names = T, sep="\t")

bb6<-windfarmC1999    %>%    filter(windfarmC1999$tempprovider1    >    23    &
windfarmC1999$tempprovider1 <= 32)
bb6_yk <- select(bb6, date, provider4power,scadapower,reference)
write.table(bb6_yk,file="temp6.txt",row.names=F,col.names = T, sep="\t")

sum(length(bb1$provider2power))+
sum(length(bb2$provider4power))+
sum(length(bb3$provider2power))+
sum(length(bb4$provider4power))+
```

```
sum(length(bb5$provider2power))+
sum(length(bb6$provider4power))

sum(bb1[2])+sum(bb2[2])+sum(bb3[2])+sum(bb4[2])+sum(bb5[2])+sum(bb6[2])
sum(bb1[3])+sum(bb2[3])+sum(bb3[3])+sum(bb4[3])+sum(bb5[3])+sum(bb6[3])
sum(bb1[5])+sum(bb2[5])+sum(bb3[5])+sum(bb4[5])+sum(bb5[5])+sum(bb6[5])
sum(bb1[6])+sum(bb2[6])+sum(bb3[6])+sum(bb4[6])+sum(bb5[6])+sum(bb6[6])
sum(bb1[7])+sum(bb2[7])+sum(bb3[7])+sum(bb4[7])+sum(bb5[7])+sum(bb6[7])
sum(bb1[8])+sum(bb2[8])+sum(bb3[8])+sum(bb4[8])+sum(bb5[8])+sum(bb6[8])
sum(bb1[4])+sum(bb2[4])+sum(bb3[4])+sum(bb4[4])+sum(bb5[4])+sum(bb6[4])

######TURBULENCE###############
turbaa<-windfarmC1999 %>% filter(windfarmC1999$provider1_WS > 0 &
windfarmC1999$provider1_WS <= 3 & windfarmC1999$tempprovider1 > -4 &
windfarmC1999$tempprovider1 <= 32)
turbaa_yk <- select(turbaa, date, provider3power,scadapower,reference)
write.table(turbaa_yk,file="turb1.txt",row.names=F,col.names = T, sep="\t")

turbbb<-windfarmC1999 %>% filter(windfarmC1999$provider1_WS > 3 &
windfarmC1999$provider1_WS <= 6 & windfarmC1999$tempprovider1 > -4 &
windfarmC1999$tempprovider1 <= 32)
turbbb_yk <- select(turbbb, date, provider4power,scadapower,reference)
write.table(turbbb_yk,file="turb2.txt",row.names=F,col.names = T, sep="\t")

turbcc<-windfarmC1999 %>% filter(windfarmC1999$provider1_WS > 6 &
windfarmC1999$provider1_WS <= 12 & windfarmC1999$tempprovider1 > -4 &
windfarmC1999$tempprovider1 <= 32)
turbcc_yk <- select(turbcc, date, provider4power,scadapower,reference)
write.table(turbcc_yk,file="turb3.txt",row.names=F,col.names = T, sep="\t")

turbdd<-windfarmC1999 %>% filter(windfarmC1999$provider1_WS > 12 &
windfarmC1999$tempprovider1 > -4 & windfarmC1999$tempprovider1 <= 32)
turbdd_yk <- select(turbdd, date, provider4power,scadapower,reference)
write.table(turbdd_yk,file="turb4.txt",row.names=F,col.names = T, sep="\t")

sum(turbaa$scadapower)
sum(length(turbaa$provider3power))+
sum(length(turbbb$provider4power))+sum(length(turbcc$provider4power))+sum(leng
th(turbdd$provider4power))
sum(turbaa$provider3power)+
sum(turbbb$provider4power)+sum(turbcc$provider4power)+sum(turbdd$provider4powe
r)

####################windfarmb########################
View(T5)
View(T8)
library(openair)
windRose(T5, ws = "ws", wd = "wd")
windRose(T8, ws = "ws", wd = "wd", col="jet",key.header = "T8 - Wind Farm
B",key.position = "top")

summary(T8$ws)

TIT8<- aggregate(T8$TI, FUN=mean, na.rm=TRUE,
by=list(cut(T8$ws, breaks=c(seq(from= 0, to =27 , by =1)), include.lowest=F)))
View(TIT8)

plot(TIT8$Group.1, TIT8$x, xlab= "Forecast Provider 1 - WindSpeed [m/s]",
ylab="Turbulence Intensity", main="Turbulence Intensity of T8 - Wind Farm B")
lines(TIT8$x, col="red", lwd=2)

#Time Series
T8windfarmb <- read_excel("C:/Users/Yüksel
Kalay/Desktop/windfarmb.xlsx",sheet="All_2019")
plot(T8windfarmb$T8_WS, type="l",col="red", lwd="1",ylab="Wind Speed [m/s]")
lines(T8windfarmb$dnv_WS, type="l",col="purple")
legend("topright",legend=c("Wind Speed - T8", "Wind Speed -
provider1"),text.col=c("red","purple"))
```

```
#Correlation Plot
ggplot(T8windfarmb,aes(x = T8_WS, y = dnv_WS)) +
geom_point() +
geom_smooth(method = "lm", se=FALSE,formula=y~x-1 ) +
stat_regline_equation(label.y = 25,formula=y~x-1  , aes(label = ..eq.label..))
+
stat_regline_equation(label.y = 27,formula=y~x-1  , aes(label = ..eq.label..))
+
stat_regline_equation(label.y = 29, formula=y~x-1 ,aes(label = ..rr.label..))+
#stat_cor( label.y=17,method="pearson",cor.coef.name="R")
labs(title = "Correlation Plot", subtitle = "",x ="Wind Speed - T8", y="Wind
Speed - provider1")

###########windfarmc###########################
#windRose(T20, ws = "ws", wd = "wd", col="jet",key.header = "T20",key.position
= "top")
summary(T23$ws)
windRose(T23, ws = "ws", wd = "wd", col="jet",key.header = "T23 - Wind Farm
C",key.position = "top")
TIT23<- aggregate(T23$TI, FUN=mean, na.rm=TRUE,
by=list(cut(T23$ws, breaks=c(seq(from= 0, to =27 , by =1)), include.lowest=F)))

View(TIT23)
plot(TIT23$Group.1, TIT23$x, xlab= "Forecast Provider 1 - WindSpeed [m/s]",
ylab="Turbulence Intensity", main="Turbulence Intensity of T23 - Wind Farm C")
lines(TIT23$x, col="red", lwd=2)
#######################windfarma###########################T13
#windRose(T1, ws = "ws", wd = "wd", col="jet",key.header = "T1",key.position =
"top")
summary(T16$ws)
windRose(T16, ws = "ws", wd = "wd", col="jet",key.header = "T16 - Wind Farm
A",key.position = "top")

TIT16<- aggregate(T16$TI, FUN=mean, na.rm=TRUE,
by=list(cut(T16$ws, breaks=c(seq(from= 0, to =26 , by =1)), include.lowest=F)))

plot(TIT16$Group.1, TIT16$x, xlab= "Forecast Provider 1 - WindSpeed [m/s]",
ylab="Turbulence Intensity", main="Turbulence Intensity of T16 - Wind Farm A")
lines(TIT16$x, col="red", lwd=2)
```