UiT The Arctic University of Norway

Faculty of Science and Technology
Department of Physics and Technology

Wind power forecasting as input to day-ahead trading strategies for wind in complex terrain

Julie Therese Svane

EOM-3901 Master's thesis in energy, climate and environment 30 SP June 2022

Abstract

Giving the increasing penetration of intermittent wind power in the liberalized electricity market, wind power forecasting (WPF) is a topic of growing importance [Kariniotakis, 2017]. The number of papers on the field WPF evaluating the statistical performance has increased rapidly, while only a proportion of the former studies focus on the economic benefit of WPF. In this study we have answered how well a set of wind power forecasting (WPF) models works as day-ahead trading strategies for a 54MW wind power park. The performance evaluation is based on both statistic and economic measures. The wind power park is located in Northern Norway in a region with complex terrain and an arctic and coastal climate. The WPF models are applied on weather forecasts provided by two numerical weather prediction (NWP) models i.e., MEPS and AROME Arctic, operated by the Norwegian Meteorological Institute (MET Norway). When applied on MEPS forecasts of wind speed and wind direction, and the statistical performance measures are evaluated over a test period, it is evident that the multilayer perceptron (MLP) model provides the lowest NRMSE of 21.4%. Compared to a current forecasting method of the responsible power trader (ISHK model), the MLP model shows an improvement of 4.0%. Further enhancement of the accuracy of the MLP model is attained by adding air pressure as the third input feature. The resulting NRMSE is 20.9% of installed capacity. This corresponds to a 6.3% improvement compared to the ISHK model, which verify that the MLP model can compete with a current forecasting method of the responsible power trader on statistical measures. When it comes to the economic perspective, given a single-price system of the power market, the naive persistence model surprisingly shows the highest revenue for the power producer. A total revenue of 16.42 MNOK is obtained, where the imbalance revenue accounts for 200 kNOK. However, considering both statistical and economic measures it is evident that the ISHK model is the most effective trading strategy. It provides a total revenue for the power producer at 16.25 MNOK, where the imbalance revenue accounts for 30 kNOK. To summarize, the MLP model and the ISHK model shows the overall best results considering statistic and economic measures, respectively.

Acknowledgements

Contents

1

Introduction

In the aftermath of the pandemic, the humanity is still facing the biggest crises of our time caused by climate change, and the last report of IPCC stress that it happens even faster than first expected [IPCC, 2022a]. To limit global warming, as agreed in the Paris Agreement, deep emission cuts across all sectors and regions are required [IPCC, 2022b]. One climate change mitigation pathway for a regional energy system is to implement a higher penetration of renewable energy sources in the respective electricity system, and at the same time reduce the penetration of fossil fuels. This is seen as a central part of the green transition, which is happening now [Birkelund et al., 2021]. Currently, hydro and bioenergy, with their considerable amounts of flexibility, are recognized as the most important energy sources. Nevertheless, among the renewable energy sources wind are characterized as one of the sources with the greatest future potential, with the most rapid growth and the lowest cost of production [Moriarty and Honnery, 2021] [Wang et al., 2011]. However, wind power production is weather dependent, making the power production variable and intermittent. Several challenges regarding a large-scale integration of wind power are highlighted in the literature, which can become even more challenging in the future. These includes real-time grid operations, power system stability and reliability, transmission capacity upgrades, market design, electricity market clearing, ancillary service requirements and costs, competitive power quality and standards of interconnection. One way to overcome many of these challenges is through accurate wind power forecasting (WPF). Hence, WPF is acknowledged as an important and essential contribution for reliable large-scale wind power integration [Wang et al., 2011].

Taking the example of the Nordic region, and more specifically Norway, we see an electricity system that mostly consist of renewable energy sources already. The Norwegian electricity system consists of over 98 percent installed renewable generation capacity [Langset and Nielsen, 2021]. Nevertheless, the ongoing green transition, coupled with the continues population and economic growth, will increase the energy demand [Wang et al., 2019]. Several energy demand scenarios have been investigated by the Norwegian transmission system operator (TSO), Statnett [2021]. Norway has recognized that wind power provides a significant opportunity for future power generation. The limited onshore potential for hydro and solar power in Norway, coupled with the growing opposition against onshore wind power, makes offshore wind the most realistic source for Norway to increase the power generation in a short-term perspective [Statnett, 2021]. A higher integration of intermittent and varying wind power makes the Norwegian energy system even more weather dependent than today [Birkelund et al., 2021]. This brings us to the importance of wind power forecasting (WPF), making use of the operational weather forecasts [Wang et al., 2011].

1.1   Former research

1.1.1   Renewable power forecasting

Given the increasing penetration of renewable energy, renewable power forecasting is a topic of growing importance [Kariniotakis, 2017]. Alessandrini and Sperati [2017] discuss and compare the results from two important projects dedicated to the field benchmarking renewable energy forecasting: ANEMOS and WIRE. The ANEMOS project was carried out in 2003 and was dedicated to benchmarking of short-term wind power forecasting (WPF) in Europe using numerical weather prediction (NWP) data as input. The objective was to improve the accuracy of the existing technology on the field [Kariniotakis et al., 2004]. A set of state-of-the-art methods were tested for a set of wind power parks located in various types of terrain characteristics e.g., flat vs complex terrain or onshore vs offshore locations. To evaluate the performance of the models as a function of terrain characteristic the ruggedness index (RIX) were used. The performance were evaluated under an common protocol provided by Madsen et al. [2005]. The WIRE project started ten years later in 2013 and had a slightly different focus. The focus was on the requirements for the weather forecast used as input and both wind and solar power prediction systems were investigated. In the case of wind power forecasting, different

methods were tested for one park in complex terrain and one in flat terrain. To mention a few of the models: Support vector machine (SVM), feed-forward multilayer perceptron (MLP), a power curve obtained by linear interpolation and a "nonlinear function approximation between wind speed and direction to wind power output" [Alessandrini and Sperati, 2017].

Renewable prediction systems applied for 6 hours and up to 7 days ahead are mainly based on data from NWP models which has gone through a statistical post-processing. The performance of these prediction systems is therefore very dependent on the errors in the NWP model. Alessandrini and Sperati [2017] states that the NWP data is one of the largest sources of error in short-term prediction systems for renewable energy. The good news is that the NWP models have been improved and show more accurate results today, compared to what we saw back in the 1980s when the research on the filed renewable energy forecasting started. The improvement over years in terms of availability of computational power, meteorological observations and data-assimilation methods have increased accuracy and improved NWP models [Alessandrini and Sperati, 2017].

One of the major findings of both projects within wind power forecasting were how the accuracy depends on the complexity of the terrain. No model performed well for all types of terrain characteristics. A model that shows a good performance in complex terrain will not necessarily perform well in flat terrain. Nevertheless, Alessandrini and Sperati [2017] concluded that machine learning algorithms had the overall best performance within wind power forecasting because they can recreate the non-linear relationship between weather parameters and power output of a park.

1.1.2 Wind power forecasting

Wind power forecasting (WPF) is a multidisciplinary research field, which include many different aspects: meteorology, applied mathematics, artificial intelligence, energy, software engineering, information technology etc. Thus, dealing with all aspects of wind power prediction is a comprehensive task and it is not "plug-and-play". A prediction system that performs well for one site will not necessarily perform well for another site with different topological characteristics as e.g. flat vs complex terrain, or onshore vs offshore wind parks. In other words, wind power prediction is site dependent. It is recommended to devote considerable effort in tuning the models for the specific site to obtain acceptable accuracy [Giebel and Kariniotakis, 2017].

The study of Giebel et al. [2011] and Giebel and Kariniotakis [2017] provides

the state of the art within short term wind power and include both past and present attempts. Short-term wind power prediction consists of several successive stages. Since the first wind power forecasts were made, improvements have been achieved at each stage [Giebel and Kariniotakis, 2017]. The most essential stages of the process are presented by Giebel and Kariniotakis [2017]. Given a prediction horizon exceeding six hours, which is the case of short-term prediction, it always starts with a NWP model. NWP models can provide weather forecasts with a certain temporal, horizontal and vertical resolution. Given the simulation results of an NWP model for a certain time-period, the next stage is to downscale the NWP results to the location of the park. This includes extracting the relevant weather parameters for a certain grid cell and a certain pressure level. Also, the temporal resolution is adjusted to fit the framework of the short-term prediction problem. The next stage is to convert the localized weather parameters into park power production. The final stage according to Giebel and Kariniotakis [2017] is to upscale the predictions to not only yield a single park, but a whole region.

WPF models can be classified in various ways. They can be classified considering the input data, the prediction horizon or the applied methodology Giebel et al. [2011] Hanifi et al. [2020]. The following classification is mainly based on applied methodology. Giebel et al. [2011] distinguishes between two main perspectives within the field short-term WPF: Physical approaches and statistical approaches. A combination is also an approach widely used in operational and commercial models. The physical approaches focus on the physical aspects of wind power. In the end of the process it utilize statistics to reduce the error, which is referred to as model output statistics (MOS). The statistical approaches primarily focuses on statistics and math. One aims to find the relation between inputs and power output. The input can be univariate or multivariate. In the univariate case, there are only one feature. This feature can for instance represent historical wind speed data. In the multivariate case, two or more features are used as input. The features can then represent both historical data and NWP data for the future. Note that the output can also be ether univariate or multivariate. In the case of WPF the output is univariate. A more recent study by Hanifi et al. [2020] divide the statistical models in two main subclassifications: time series based and neural network (NN) based. The ARMA (Auroregressiv Moving Average) is a typical time series based model. When it comes to the (NN) based models Artifical neural networks (ANNs) is common NN based models in WPF.

Giebel et al. [2011] conclude that a typical accuracy for a 36-hour prediction horizon is an RMSE of around 10-15% of the installed wind power capacity. The main error in a short-term prediction model arises from the NWP model. Giebel and Kariniotakis [2017] highlights three strategies to cope with this error source. The first is to use ensemble learners to form the prediction system. The

second is to use several NWP models to provide the input data of the prediction system. Another alternative is to use one NWP model with several different initial conditions to provide different input for the prediction system.

Jung and Broadwater [2014] and Hanifi et al. [2020] both discuss potential options to help improve the forecast accuracy in WPF. The common suggestions includes Kalman filtering, outlier detection, combination of NWP data, selection of input parameters and statistical down-scaling approaches.

Haupt et al. [2017] aims to present the principles of meteorology and NWP relevant for renewable energy forecasting. They state that the meteorological understanding of the wind is a key when working with modelling and forecasting of wind power. The driver of the variability and intermittency of the wind are atmospheric physics and dynamics. Through NWP models these meteorological processes are simulated and can be used to predict weather parameters. All NWP models are sensitive to the model setup. Both initial conditions, model resolution, numerical approximations and choice of physical equations describing the atmospheric conditions, can contribute to noticeable errors between the model output and the real atmospheric conditions [Haupt et al., 2017].

Wang et al. [2019] highlight the important role of accurate wind power curves in a wind power prediction perspective. The paper forwards two different wind power prediction perspectives where an accurate power curve is important. The first perspective has two main steps. The first step is to provide a wind speed forecast. This can be done using various forecasting models with historical wind speed data as input. The second step is to use a power curve to convert the wind speed forecast into wind power. In this perspective, the resulting wind power forecast error can be decomposed into two parts: The error related to the wind speed forecast and the error related to the convention of wind speed to power.

The second perspective utilize an estimated power curve to process outliers in raw wind speed and wind power data. The estimated power curve is based on historical data of wind speed and wind power. In this process the outliers that lays far away from the estimated power curve are eliminated. The processed wind power data together with the historical wind speed data are then used as input to a wind power forecasting model, which gives the final wind power forecast. These two perspectives show the important role of an accurate power curve within wind power prediction. The paper also examines how to pre-processes raw wind and power data and gives a literature review on different methodologies within power curve modelling [Wang et al., 2019].

1.1.3   Prediction methodologies

Nielsen et al. [2006] describes a state of the art wind power prediction system, where a power curve is central. The statistical techniques behind a direction-spesific park power curve, developed through the ANEMOS project, is described.

Kusiak et al. [2009] use five data mining approaches to predict wind park power output over two prediction horizons, one shorter (1 to 12h) and one longer (3 to 84h). The five approaches include the support vector machine for regression (SVM), multilayer perceptron (MLP), radial basis function (RBF) network, regression tree (RT) and random forest (RF). The algorithms were built with two different prediction methodologies: the direct and the integrated prediction methodology. The input of the prediction systems were wind data from two different NWP models. The results of the research showed that the MPL network algorithm outperformed the four others over both prediction horizons.

Tan et al. [2021] use the weather research forecasting (WRF) model for short term wind energy resource prediction. The main purpose of the paper is to assess the performance of the WRF model for predictions of wind speed and wind direction for three different time horizons: 24, 48, 72 hours ahead. However, the paper also provides wind power predictions based on a power curve mapping WRF wind speed to park power. To reduce the systematic errors of the initial power prediction, coming from the WRF model, an artificial neural network (ANN) model is applied using two different methodologies: single-input single-output (SISO) and multi-input single-output (MISO). These methodologies are applied to wind speed and power in individual experiments. The results show that the accuracy of the ANN predictions depends on the season, time horizon, location of turbines, methodology and which input feature that are used. The MISO provides significantly improved wind power predictions at large prediction horizons compared to the initial wind power predictions, especially for the spring, summer and fall.

In the wind energy industry, machine learning is often used to build statistical models that predict wind park power production based on atmospheric input data. The current literature on the field has a major focus on finding the best-performing learning algorithms. However, Optis and Perr-Sauer [2019] states that the choice of variables is more important than the choice of algorithm. They recommend including atmospheric turbulence and stability data as input. They argue that observations and physical models have shown that the presence of these variables may have considerable impact on wind park power production. Among five learning algorithms and nine atmospheric variables, they find that the most significant variable after wind speed is the turbulent kinetic energy. In addition, they find that pressure and temperature is more important than wind direction. The final recommendation based on their findings

is to make turbulence and stability variables standard inputs in wind power prediction models.

1.1.4   Wind power trading

There are many application areas of WPF models. One of many applications is to maximize the profit of electricity traders [Wang et al., 2011]. Over the past few years there has been a growth in installed wind power capacity in Norway and Europe in general. Among many reasons is the green support schemes in form of incentives, e.g., tax credits, feed in tariffs etc. At the same time, the energy cost of wind energy has decreased, so that smaller incentives are required to be competitive in the electricity market. Recently the green support schemes were removed in Norway, which required the wind power producers to participate in the electricity market under the same rules as conventional power plants. In addition to the price risk involved with operation in the Norwegian hydro power market, the wind power producer is financially responsible for their potential imbalances. This suggests that the resulting energy imbalance are corrected at the balancing market operated by the TSO and that the responsible trader take the cost [Morales et al., 2013].

Even if wind power producers must follow the same rules as conventional producers, trading wind power presents considerable differences compared to the situation of conventional power plants. First, the energy source is varying and intermittent due to its stochastic nature. This makes the wind power production uncertain at the time of delivering. Thus, the returns of a wind power producer depend on realization of both the price and the power production. As a sequence of the uncertain power production, the wind power trader must participate in at least two markets. There will be deviations between scheduled and delivered wind power at the day-ahead market due to the long prediction horizon. The power trader can handle the production imbalance in two ways. The first is to correct the schedule at the intraday market, where the power price is known but there is still uncertainty regarding the actual power production volume. The second is to let the imbalance go to the balancing market, which takes place in real-time. This happens automatically if the first option is not used [Morales et al., 2013].

The pricing system implemented for the electricity system is a deciding factor for the revenue and cost opportunities of the power trader. The literature distinguishes between two pricing systems: single-price and dual-price systems. Real-time deviations steaming from the day-ahead schedule in a two-price system are always penalized. In this system the renewable power plants have a disadvantaged position compared to conventional units. However, in a single

price system these deviations are either penalized or rewarded [Morales et al., 2013]. The latter price system is recently implemented in Norway and this study focuses on how this system can provide revenues for the power trader due to its imbalance. Short term wind power forecasting is a central tool for the power traders to make effective trading strategies for the day ahead market [Mazzi and Pinson, 2017].

1.1.5   Fakken wind power park

Fakken wind power park is a 54MW wind power park located in northern Norway. The site of the park is characterized by an arctic, cold and coastal climate, where the surrounding terrain is complex, consisting of fjords, mountains and islands. Previous master thesis and researcher papers at UiT, The Arctic University of Tromsø, have looked at short-term wind power forecasting related to the specific case of Fakken wind power park [Fossem, 2019][Sæther, 2021] [Jacobsen, 2014][Birkelund et al., 2018][Eikeland et al., 2022].

Jacobsen [2014] applied different versions of the Markov Chain model to predict the power output of one wind turbine at Fakken wind power park, with a time horizon of two hours. The model with the best performance had a normalized root mean square error (NRMSE) of 16.84 percent of installed capacity.

Fossem [2019] and Sæther [2021] have both investigated short-term wind power prediction models for several wind power parks located in Northern Norway, including Fakken wind power park. Fossem [2019] investigated different versions of the Markov chain model with a time horizon of two hours. The best performing Markov chain model had a NRMSE of 14.51% of installed capacity. A meteorological data-customized power curve function based on polynomial regression was also proposed and investigated. combining this latter model with the naive persistence model gave the overall best performance with a NRMSE of 13.32%.

Sæther [2021] investigated several time series forecasting algorithms for Fakken wind power park: naive persistence, one classical statistical method, auto regressive integrated moving average - maximum, and two machine learning methods, random forest (RF) and support vector machine (SVM). The algorithms were built using two different methodologies: direct and recursive. Park power predictions were made for the time horizon ranging from 1 hour to 24 hours ahead. The most promising model for Fakken was shown to be the recursive SVR model. For a prediction horizon of 12 hours the NRMSE is 46%.

Birkelund et al. [2018] investigated analog ensemble (AnEn) as a wind power

forecasting model, focusing on the same five wind power parks as Fossem [2019] and Sæther [2021]. The input data is wind data provided by the Norwegian meteorological institute. The prediction horizon ranges from 1 hour to 66 hours ahead. The performance of the AnEn model is compared to the performance of a direction-specific park power curve model (PNWP) which takes the same input as the AnEn model. The AnEn provides a robust park power forecast and outperforms the PNWP model for all wind power parks and for all lead times.

The surrounding terrain has a large impact on the wind resources at the location of the wind power park [Markowski and Richardson, 2011] [Jackson et al., 2013]. Solbakken et al. [2021] evaluate surface winds using WRF for the area of Fakken wind power park. The paper found that the coastal orthography can under certain stability and wind conditions contribute to create complex flow patterns, i.e., dynamically driven winds, which include blocking, gap winds and mountain waves.

1.2   Objectives

The overall objective of this master thesis is to answer how well a set of wind power forecasting (WPF) models works as day-ahead trading strategies for Fakken wind power park. A case study with real data is investigated to better understand how the trading strategies differs, and what value wind power forecasting has in this perspective. This will be achieved by:

- Collecting measured power production output, market prices, and modelled weather data.

- Make datasets that match the operational framework of the day-ahead market and analyse the raw data.

- Apply several models for short-term forecasting: three reference models, a statistical model, two machine learning models and an artificial natural network (ANN).

- Evaluate and compare the performance of the different power estimation models based on statistical and economic considerations.

- Compare the models with an operational model based on statistical and economical measures. Are the models competitive for operational power forecasting?

1.3   Contribution

This master thesis is a contribution to the research field on short-term wind power forecasting at UiT and focus on Fakken wind power park. My contribution is to investigate day-ahead prediction models i.e., the prediction horizon of 12-36 hours ahead, and to evaluate the economic revenue and cost of including these wind forecasting models as the trading strategies of the responsible wind power trader.

This study is structured as follows. Chapter 2 provide theory on wind energy, electricity markets, wind power forecasting and machine learning models. The operational framework, data and method is described in chapter 3. In chapter 4 the results are presented. The results of the raw data analysis and validation of hyper-parameters build up to the main results, namely the performance evaluation of the WPF models. The WPF models are evaluated using both statistic and economic measures. The economic measures explain how the WPF models work as trading strategies. The results are discussed in chapter 5, examining the errors and limitations of the study. Finally, conclusions are drawn and further research are suggested in chapter 6. Note that the descriptions of the pure statistical model is directly copied form my project thesis [Svane, 2021]. This also yields parts of the former research and the theory on wind energy and the electricity market.

2

Theory

2.1 Wind Energy

2.1.1 Conversion of wind energy

Wind is moving air. Air is a fluid with mass and for that reason the moving air has kinetic energy. Considering the wind approaching a fixed area, A, the kinetic power [W] of the wind can be given by equation 2.1.

$$P_k = \frac{1}{2}\dot{m}v^2, \tag{2.1}$$

where v is the wind speed, $\dot{m} = \rho A v$ is the mass flow of the fluid through the area A, and ρ is air density. Substituting the time derivative of the mass with the expression for the mass flow we find that kinetic power can be written as:

$$P_k = \frac{1}{2}\rho A v^3. \tag{2.2}$$

This equation highlights the propositional context between the power and the cubic wind speed. In fact when the wind speed doubles, the power increases by a factor of eight [Wizelius and Earnest, 2011]. Thus, a small increase in the wind speed, results in a larger increase of the power. This makes wind speed the most deciding factor of the power production. Nevertheless, air density and the fixed area also has an impact on the power production. Air density is typically a function of pressure, temperature and molecular composition of air,

which makes these atmospheric variables relevant for the power production [Wizelius and Earnest, 2011] [Wang et al., 2019].

To be able to utilize the power of the wind, we need to capture and covert it to a usable form. For this purpose, we have wind turbines, which can extract a fraction of the power in the wind. A wind turbine can only capture and convert parts of the kinetic power in the wind, P_k, given in equation 2.2. The fraction of the power in the wind that can be captured by the rotor is referred to as the power coefficient, C_p. Considering air that approaches the swept area, A, of a wind turbine. The power extracted by the wind turbine can be expressed as follows:

$$P_t = \frac{1}{2}\dot{m}v^2 C_p = \frac{1}{2}\rho A v^3 C_p = P_k C_p, \qquad (2.3)$$

where the power production is affected by turbine specific variables, in addition to the atmospheric variables described earlier. These turbine specific variables include the sweat area of the turbine rotor and the capacity coefficient of the turbine. The theoretical maximum power that a rotor of a wind turbine can extract is the fraction 16/27, or approximately 59 percent, of the power in the wind. This is known as the Betz-limit and reflect the maximum value for the power coefficient. In practice, this power coefficient will always be lower, due to aerodynamic and mechanical losses that are neglected in the theoretical calculations describing the Betz-limit.

How much power produced by the turbine is based on how much the wind is slowed down behind the rotor. Considering an imaginary stream tube that describes the wind surrounding the wind turbine, we find that the stream tube must expand behind the rotor. This is because the wind approaching the rotor is slowed down, while the mass flow is held constant. For this to hold true, the stream tube needs to expand as the wind decreases. Aerodynamics is the part of fluid mechanics that studies the properties of air [Wizelius and Earnest, 2011]. Further details about aerodynamics and the extraction of energy from the wind is beyond the scope of this study and will not be explained any further.

2.1.2   Power curve

Wang et al. [2019] highlights the important role of wind power curves in a wind power forecasting perspective. The general explanation of a power curve is a function that maps the input to the output. A typical wind power curve illustrates the non-linear relationship between wind speed and turbine power output. Wind turbine manufacturers provide such power curves for different turbine designs.

Figure 2.1 displays a power curve of a Vestas V90 3MW wind turbine. The cut-in wind speed, the rated wind speed and the cut-out wind speed are marked in the figure. The respective values for a V90 3MW wind turbine are 3.5m/s, 15.0m/s and 25.0m/s [Vestas, n.d.a] [Ramadan et al., 2016]. According to the figure, the turbine is not able to generate any power when the wind speed is below the cut-in wind speed. Similarly, the wind turbine will not produce any power when the wind speed is higher than the cut-out wind speed. The turbine is expected to start producing power for wind speeds equal to or exceeding the cut-in wind speed. The power generation of the turbine for wind speeds between the cut-in and the rated wind speed, follow the expression given in equation 2.3. For wind speeds between the rated wind speed and the cut-out wind speed, the turbine in theory can generate constant maximum power output. In general, the power curve of a manufacturer gives an indication of expected power production for a specific wind turbine at a certain wind speed.
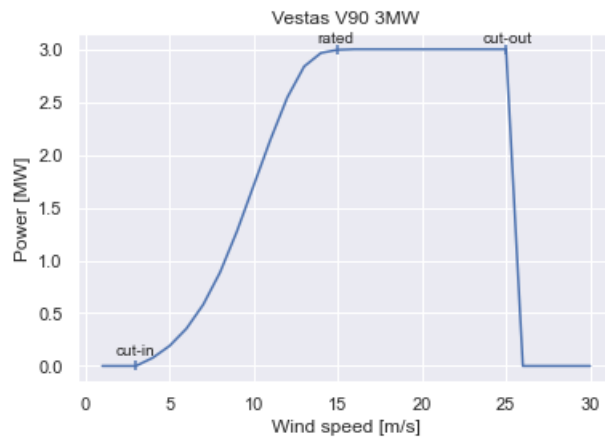
Figure 2.1: Wind power curve for a Vestas V90 3MW turbine with cut-in, rated and cut-out wind speeds marked

The behavior of a wind turbine in practice will not necessarily follow the power curve estimated by the manufacturer. It is strongly dependent on the weather and topography of the turbine's specific location. The power curves provided by the manufacturer are estimated under ideal meteorological and topographical conditions, which rarely represent the specific site of any turbine. In other words, wind turbines with the same design located in different terrain and locations, probably produce different power outputs. [Wang et al., 2019] states that using wind data measured at wind turbines and wind parks are necessary to build site-specific wind power curves that are more accurate compared to

the ones from the manufacturer.

The operational settings for a turbine can be updated or changed after installation. This might in turn modify the power curve of a turbine. New software updates of Vestas wind turbines makes it possible to extend the cut-out and enable gradual shutdown of the turbines. In other words, the abrupt transition at cut-off, are now changed to a gradual reduction in power production. There also exist power updates that can optimize the power output of the turbines for moderate wind speeds [Vestas, n.d.b]. These operational settings was updated for Fakken wind power park in summer 2019.

Beyond the typical wind power curve which maps wind speed to turbine power, it is possible to provide wind power curves based on multiple input variables. For instance, a park power curve that provides power production as a function of both wind speed and direction. Weir [2014] proposed such direction-specific power curves for several wind power parks located in Norway. The park power curve (KVT-PPC) proposed for Fakken wind power park is displayed as two- and three-dimensional heat-maps in figure 2.2. The KVT-PPC is obtained using a pre-defined model in WindPRO called the Park Power Verification model. This model takes in the location of each turbine, the hub height of the turbines, the respective power curves from the manufacturer, the surrounding topography and the wake effect in-between the turbines in the park citepFossem2019short. The KVT-PPC shows small variations in the power curve for each of the 12 wind sectors. Among all wind sectors, the park will produce at maximum power (53 MWh) at lower wind speeds in north (0 degrees) and south (180 degrees) wind sector.
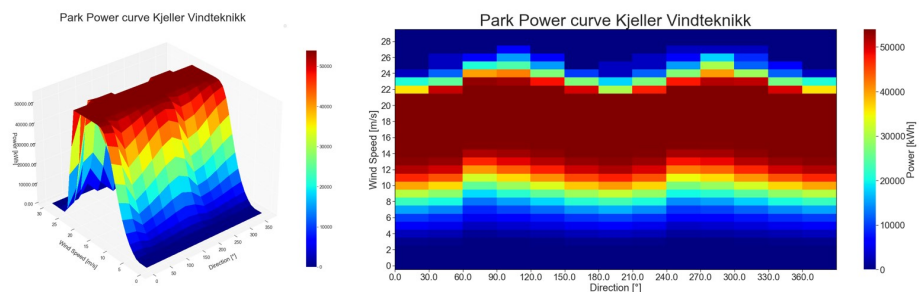

Figure 2.2: Direction-specific power curve for Fakken wind power park illustrated as three-dimensional and two-dimensional heat-maps

2.2   The Norwegian electricity market

In the last section it was established that the wind contains a large amount of power, and that only parts of it can be converted to usable electrical power through turbines. For a consumer to be able to utilize the electrical power, an arrangement for transmission from producer to consumer is needed, the electricity market. It is split in three parts: The power system, the system for electricity trading and the grid tariffs. The purpose of a system for electricity trading is to ensure that the producers get paid for the power they deliver, and that the consumers pay for the amount they use. The electricity trading is the focus of this study. The theory is mainly based on literature by Söder and Amelin [2011], Morales et al. [2013] and Mazzi and Pinson [2017].

Trading of electricity is done in several steps and at different timeframes. The broader steps of the trading are known as the ahead trading, the real-time trading and the post trading. The ahead trading includes all trading that takes place before the real time operation hour. The real-time trading includes all trading that takes place within the real time operation hour, in which the purpose is to maintain a secure operation of the electricity system. After the operation hour has passed the final deviation between scheduled and actual production and consumption for each player are settled in the post trading [Söder and Amelin, 2011].

The steps of the electricity trading proposed above consist of several subsequent markets of distinct time frames. How the subsequent markets are organized differs between countries. Figure 2.3 illustrates the subsequent markets implemented in the Norwegian electricity market [NVE, 2021]. The financial market, the day-ahead market and the intraday market are all markets belonging to the ahead trading. In the Nordic countries the physical ahead trading is done through a joint power pool, called Nord Pool. The trading prices are assured though the Nasdaq financial market [Wizelius and Earnest, 2011]. The real-time trading includes the balancing market where fine adjustments are made by the transmission system operator (TSO) to maintain the balance between consumption and production in the power system. The post trading includes the imbalance settlements and takes please after the operation hour [NVE, 2021][Mazzi and Pinson, 2017].
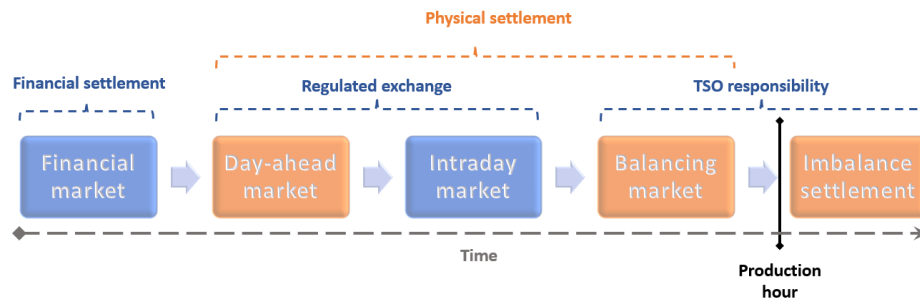
Figure 2.3: The time frames of the Norwegian electricity market [NVE, 2021]

2.2.1 The ahead trading

Mazzi and Pinson [2017] distinguishes between two main trading floors within the ahead trading: short-term markets and medium/long-term markets. A typical short-term market is an electricity pool, e.g., Nord Pool, where physical products are traded on a time horizon ranging from minutes to hours to days. Electricity pools usually contains subsequent trading floors with distinct time frames, i.e., the day-ahead market and the intraday market. A typical medium/long term market, on the other hand, is a so-called future market where forward contracts and options are offered for both physical and financial products. Future markets have a long time horizon, often ranging from weeks to months. These markets may be beneficial for hydro or thermal power plants, that can contract a part of the total capacity to ensure fixed revenues [Mazzi and Pinson, 2017]. However, wind power producers do not have this opportunity due to the intermittency of the wind. The financial market will therefore not be further explained.

The day-ahead market

The day-ahead market enables trading of electric power one day in advance of the production. The following day is then divided in hourly intervals, which means 24 separate trading periods at 1 hour each. The players that want to trade power at the day-ahead market must submit their bids to the electricity pool by a certain deadline. At Nord Pool, which is the Nordic market operator and electricity pool, this deadline is at 12:00 each day [NVE, 2021]. How the bids are designed and how the market-clearing price and volume are calculated vary

between diverse electricity pools and market operators. At Nord pool there are marginal pricing for the bids and a price-cross for the market-clearing price and volume. In addition, the Nord pool markets are separated in different bidding areas, which are established to limit the trading between different parts of the Nordic power system. Currently there are five bidding areas in Norway [NordPool, n.d.]. A day-ahead bid submitted at Nord Pool must specify the price that the player is inclined to contact for a certain amount of power and for a certain hour. In addition, the bid must specify the bidding area in which the power will be injected or extracted. One distinguishes between the bids submitted by consumers and producers. A purchase bid is submitted by a consumer and specifies the maximum price that the consumer is inclined to pay giving the amount and hour. An offer bid is submitted by a producer and, must in a similar manner, specify the minimum price that the producer is inclined to pay.

After the gate of the day-ahead market closes, the market operator matches the submitted purchase and offer bids [Belyakov, 2019]. The purchase bids are sorted in descending order, to form a demand curve, while the offer bids are sorted in ascending order, to form a supply curve. The point where the two curves cross determine the market-clearing price i.e., the spot price and volume. This is illustrated in figure 2.4. At Nord pool this is done individually for all bidding areas. The variable transmission capacity and bottlenecks in different parts of the Nordic power system limits the power flow between bidding areas. The consequence is that the area prices differ [NordPool, n.d.].
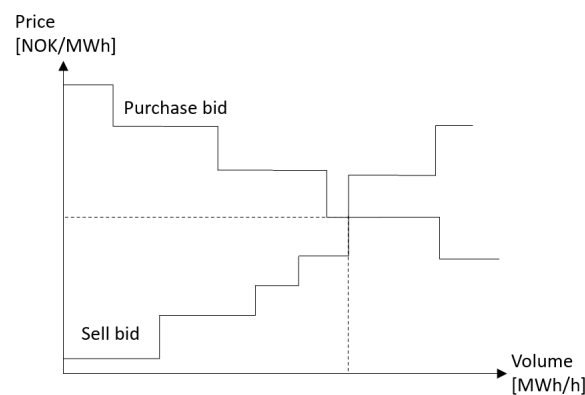

Figure 2.4: A typical supply-demand curve. The system price is set at the cross [Söder and Amelin, 2011]

The intraday market

The intraday market enables trading of electric power during the day of operation. It facilitates fine adjustments of possible imbalances related to the

submitted day-ahead schedule close to the operation hour. The intraday market works as a supplement to the day-ahead market. In general, the intraday markets opens after the day-ahead market closes and adjustments can be done until hours or minutes prior the operation hour [Mazzi and Pinson, 2017]. One advantage with intraday trading is that the spot price is known, another is that the prediction of the power volume is more accurate compared to one day in advance. Nevertheless, intraday trading is not widespread for wind power trading in Norway. This may be related how much time this takes for a trader when it is done manually. Thus, automatic prediction systems are required for the intraday trading of wind power to be more widespread.

2.2.2   The real-time trading

The balancing market

During the real-time operation hour there will be events that disturb the balance settled through the day-ahead and intraday markets. To handle this a final trading stage is established, namely the balancing market. It takes place before the power delivery. The purpose of the balancing market is to ensure instantaneous balance between supply and demand in the power system. For this purpose, it allows trading of reserves and ancillary services in separate session for each trading period. Market participants that have reserves to offer submits their bids in the balancing market before a certain deadline [NVE, 2021] [Monteiro et al., 2009]. The balancing market is operated by the balancing responsible in the country. In Norway Statnett has both the role as TSO and balance responsible [ACER, 2021] [NVE, 2021].

Conventional power plants may act as regulating power in the balancing market, by offering up- or downward regulation of the production. Power plants based on energy sources with a stochastic nature are not suitable as regulating power. This yields wind power. However, the reason for wind producers to participate in the balancing market is to resolve deviations from contracted production [Mazzi and Pinson, 2017]. If no adjustments of the day-ahead bid is done the deviation automatically goes to the balancing market, with no extra work for the power trader. The disadvantage is that the balancing price is not known in advance.

2.2.3   The post trading

The imbalance settlement

The imbalance settlement takes place after a trading period is over and it aims to establish a financial balance in the electricity market [NBS, 2021]. In the process the system operator gets an overview of the past trading period including schedules given at the ahead and real time markets, together with the actual delivery. In most cases there will be a deviation between scheduled and delivered power at the power system level [Söder and Amelin, 2011]. Each participant at the market is responsible for their individual deviation and are penalized with an imbalance cost [Mazzi and Pinson, 2017]. The imbalance cost heavily depends on the imbalance volume and on the pricing-system implemented in the market. Typically, one distinguishes between a single-price system and a dual-price system. In short, and as the names suggests, a single price system penalizes up- and down-regulation of imbalance volumes with the same price, while a dual-price system have different prices for up- and down-regulation [Söder and Amelin, 2011].

Since 2017 the imbalance settlement in Norway has been operated by a common Nordic imbalance settlement (NBS). The imbalance pricing system of production has until recently been dual pricing, while the imbalance pricing system of consumption has been a single pricing system. However, the first of November 2021, single pricing for production was implemented at the NBS [NBS, 2021]. The imbalance cost is now calculated against a balancing price, in the same way as for consumption. This change has been beneficial for power producers since the single price system enables arbitrage opportunities. The imbalance volumes can not only provide costs but also revenues for the power producer. For example, a producer can be rewarded for its deviation, if the producer and the system are of opposite signs. In this case, the deviation contributes to reduce the imbalance in the system. In contrast, the producer will be penalized when the signs are equal. Note that this is not the case of a dual-price system. In this pricing system the power producer is always penalized for its deviation and have no opportunity to receive any revenues for its imbalance [Mazzi and Pinson, 2017].

2.2.4   Trading wind power in power pools

Revenue and Imbalance Cost

In this section the concept of trading wind power at electricity pools is translated into mathematical equations. In the case of wind power trading the

market decisions are made under uncertainty. As initially described in the introduction, there are two uncertain variables involved in wind power trading: the power price and the power production. Revenue and imbalance cost is common concepts in decision-making under uncertainty and they can be used to determine optimal trading strategies for a wind power park. In this section these concepts will be presented using a simplified framework of the trading problem. The simplified framework is defined by the following assumptions, which are necessary to obtain an analytical solution of the revenue maximization problem [Morales et al., 2013] [Mazzi and Pinson, 2017].

Assumption 1: The intraday trading is neglected, which means that the responsible trader of the wind power park trades only in the day-ahead market and the balancing market i.e., the regulating market (RK). Currently this is a reasonable assumption since intraday trading is not widespread for wind power trading in Norway.

Assumption 2: The market prices are known in advance, which makes the production volume the only uncertainty. This is a simplification of the reality where the market prices are not known in advance.

Assumption 3: The responsible trader of the wind power park is risk-natural, which mean that the objective of the trader is to maximize the revenue of the power producer while neglecting possible losses.

Assumption 4: The responsible trader of the wind power park is the price-taker, which means that he or she must accept the prevailing prices in the market and cannot affect them in any way. Thus, the employed trading strategy do not affect the prices in the market.

Assumption 5: Incentives are neglected. This includes price premia added on top of market prices, like for instance feed-in tariffs etc. Thus, the wind power trader undergoes the same rules as conventional hydro power.

Assumption 6: Possible associations with other market participants, like flexible demand or storage possibilities, are neglected for the trading strategies.

Assumption 7: The energy generation is offered at zero marginal cost NOK/unit of production, i.e. the total cost of production does not increase if an additional unit is produced.

Assumption 8: The power production is traded in a single-price system, i.e. up- and down-regulations in the balancing market has the same prices.

Based upon the simplified framework defined by the presented assumptions, formulations of the market revenue and the imbalance cost can be derived. The derivation starts with the total revenue of a wind power producer, which are formulated as the sum of the revenue at each market stage. Since the intraday market stage is neglected, we are left with two market stages: the day-ahead market and the balancing market. The revenue at each stage is given by the product of the exchanged power volume and the respective market prices. Considering a single-price system up- and down-regulations at the balancing market has the same price. Taking all considerations inherent in the assumptions, the total revenue of the wind power producer can be expressed as given in equation 2.4. Let the market price and production volume be referred to as ρ and v, respectively. The notation DA is short for day-ahead market, while B is short for balancing market. The bid is delivered at Nord pool at time t and s denote which time steps ahead that is considered [Morales et al., 2013] [Mazzi and Pinson, 2017].

$$r_s = r_s^{DA} + r_s^B = \rho_s^{DA} v_s^{DA} + \rho_s^B v_s^B \tag{2.4}$$

$$v_s^B = v_s - v_s^{DA} \tag{2.5}$$

According to the proposed assumptions the day-ahead schedule, v_s^{DA}, is the only decision variable, i.e. the variable that the wind power trader can affect, in equation 2.4. This implies that the volumes exchanged at the two market stages, i.e. and v_s^B, has the relation given in equation 2.5, where the imbalance volume, v_s^B, equals the difference between the day-ahead schedule, v_s^{DA} , and the actual production, v_s. Given that the wind power park produce whenever the wind resource is available and the stochastic characteristics of the wind, the imbalance volume, v_s^B, is fixed at the balancing market.

In order to obtain the final formulation for the revenue of a wind power producer in a single-price system, equation 2.5 is inserted into equation 2.4 and rearranged as shown in equation 2.6. The final expression is given in the last line of the equation. The total revenue is now split into two terms: r_s^{PI} and

r_s^{IMB}. The first term, r_s^{PI}, represents the revenue if the trader knows at time t the precise wind production at time t+s. Therefore, this term is referred to as the revenue in the case of precise information at time t. The second term, r_s^{IMB}, on the other hand, represents the imbalance cost, that is the penalties for imbalance creation. For a single price system this term can either be positive or negative. Thus, the trader can either be rewarded or penalized in a single price system for the imbalance.

$$
\begin{aligned}
r_s &= \rho_s^{DA} v_s^{DA} + \rho_s^B (v_s - v_s^{DA}) \\
&= \rho_s^{DA} v_s^{DA} + \rho_s^B v_s - \rho_s^B v_s^{DA} \\
&= \rho_s^{DA} v_s^{DA} + \rho_s^B v_s - \rho_s^B v_s^{DA} + (\rho_s^{DA} v_s - \rho_s^{DA} v_s) \\
&= \rho_s^{DA} v_s - \rho_s^B v_s^{DA} + \rho_s^B v_s + \rho_s^{DA} v_s^{DA} - \rho_s^{DA} v_s \\
&= \rho_s^{DA} v_s - (\rho_s^B - \rho_s^{DA})(v_s^{DA} - v_s) \\
&= r_s^{PI} - r_s^{IMB}
\end{aligned}
\tag{2.6}
$$

Taking the expectation of the final expression in equation 2.6, the so called expected monetary value (EMV) is given, which is a common terminology in decision theory. Equation 2.7 provides the expected revenue of a wind power trader in a single-price market. The EMD constitutes the foundation for optimizing the trading strategies for a wind power trader. The producer can not affect the first term of equation 2.7, so this term is neglected when trying to optimize the bids in a single-price system.

$$
E[r_s] = \rho_s^{DA} E[v_s] - (\rho_s^B - \rho_s^{DA})(v_s^{DA} - E[v_s])
\tag{2.7}
$$

Performance ratio

The total revenue of a wind power trader in a single-price system, as given in equation 2.6, can be used to obtain a performance parameter to evaluate the effectiveness of a market trading strategy. [Mazzi and Pinson, 2017] suggest the performance ration, f_s, calculated for a given time step s. This performance parameter is defined in equation 2.8.

$$
f_s = \frac{r_s}{r_s^{PI}} = 1 - \frac{r_s^{IMB}}{r_s^{PI}}
\tag{2.8}
$$

The performance ratio is normally calculated over a certain time period of N days, like in equation 2.9. The lower limit, zero, of this parameter is obtained when the power trader always is penalized for it is imbalance and opposite for

the upper limit, one, where it is never penalized for any imbalance.

$$f = \sum_{d=1}^{N} \sum_{s=13}^{36} f_{d+s} \qquad (2.9)$$

2.3 Wind power forecasting

A wind power forecast (WPF) is provided at time, t, for a certain time horizon, T, ahead in time. The look-ahead time horizon consists of several time-steps, s. This study use hourly resolution for the time steps and the WPFs are given as point forecasts. For each time step, s, there is given a point forecast, $\hat{p}_{(t+s|t)}$. The point forecast is the estimated power production during the hour t+s using available input at time t [Monteiro et al., 2009].

2.3.1 Prediction horizons

Many forecasting systems, for instance load forecasting systems, are characterized by their prediction horizon with well-defined time limits. However, within wind power forecasting there are not unanimously defined time limits. This master thesis is based on the definition proposed by Monteiro et al. [2009] and Giebel and Kariniotakis [2017]. The articles define three prediction horizons for wind power prediction models: Very short term, short term and medium term.

- The very short-term prediction horizon range over a few hours ahead. According to Giebel and Kariniotakis [2017] this includes all forecasts up to 6 hours ahead. This prediction horizon can be useful for trading at the intraday market and management of ancillary services.

- The short-term prediction horizon ranges up to 48 or 72 hours, and the very short-term limit makes the starting point. This time horizon is useful for trading at the day-ahead market, which usually has a time horizon of 12-36 hours ahead from the gate closure time for bids.

- The medium-term prediction horizon ranges up to several days. Monteiro et al. [2009] set the limit at 7 days. This prediction horizon may be useful for some longer market contracts, e.g., future markets.

The time resolution and horizon of a forecast is related. In a day-ahead wind power forecasting perspective, where the time horizon is 12-36 hours ahead, an hourly resolution is normal. Here the intra time-step variations are neglected to obtain an hourly schedule scheme for the day ahead market.

2.3.2 Reference models

Reference models, or baseline models, have a central role when evaluating the performance of different WPF models. For a model to be useful it should perform better than a defined reference model. A well-known reference model within WPF is the persistence model, which is a naive prediction model. This model assumes that the power production for a future time step, s, is equal to the measured power production at time, t, which denotes the time when the forecast was made. The last measured wind production value is used as the future wind production [Madsen et al., 2005], [Monteiro et al., 2009], [Hanifi et al., 2020]. This statement can be expressed by the following notation proposed by [Monteiro et al., 2009].

$$\hat{P}_{(t+s|t)} = P_t \tag{2.10}$$

The power output for all time slots given by $t + s$, where e.g., $s \in [12, 13, ..., 36]$ for the day ahead market, are predicted to be equal the power output measured at the initial time, t, when the forecast was made. Madsen et al. [2005] states that this naive persistence model performs well up to 4-6 hours ahead and relate this to the scale of changes within the atmosphere.

Nielsen et al. [1998] introduces a new reference model for wind power forecasting. Monteiro et al. [2009] recommends this model for prediction horizons over 3 hours. This includes the day-ahead market. The new reference model is a combination of the persistence and a weighted mean. The weight a_s, is the linear correlation coefficient between p_t and p_{t+k}. The lower the correlation gets, the lower the persistence value is weighted and the higher the mean is weighted. The new reference model can be expressed with the following notation.

$$\hat{P}_{(t+s|t)} = a_s P_t + (1 - a_s)\overline{P} \tag{2.11}$$

The estimated mean of the production, $\overline{P}$, is provided by $\frac{1}{N} \sum\limits_{t=1}^{N} P_t$.

2.3.3   Classification of WPF models

Input-based classification

The classification of WPF models may be based on types of input, as proposed by Giebel et al. [2011]. The paper distinguishes between three types of inputs. They are listed below and illustrated in figure 2.5.

- (1) The SCADA data indicated the measured historical data form the specific site of interest. For instance, measure wind speed and wind direction from a meteorological mast, as well as measured power output from the grid connection point.

- (2) METEO FORECASTS indicates weather data for the considered future time horizon. This data can ether be obtained with a physical or statistical approach. For instance, weather data modelled by an NWP model are referred to as a METEO FORECAST.

- (3) TERRAIN indicates the information that is required to do physical considerations. This normally includes information of the local roughness, orography, atmospheric stability and the layout of the wind power park in the terrain. This information is required to do a downscaling of the input data, which means to refine the data to the hub height [Monteiro et al., 2009].
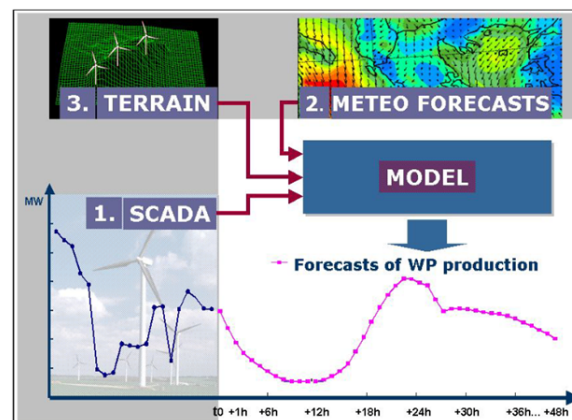

Figure 2.5: Short-term WPF different types of inputs: (1) available on-site measurement data, (2) NWP data (3) data based on terrain and physics [Giebel et al., 2011]

Statistical approach

Wind power forecasting approaches that only takes in measured historical
data (1) are normally considered as very-short-term statistical approaches.
This is illustrated in figure 2.6 [Monteiro et al., 2009]. These models have a
time horizon that ranges up to 6 hours and are based on statistical time-series
approaches, like Auto-Regressive Moving Average (ARIMA). These types of
models can either be univariate or multivariate. In univariate models the only
input is the historical data of the variable one wants to predict, while in mul-
tivariate models additional relevant weather variables are also used as input.
For instance, historical wind speed, wind direction and temperature.
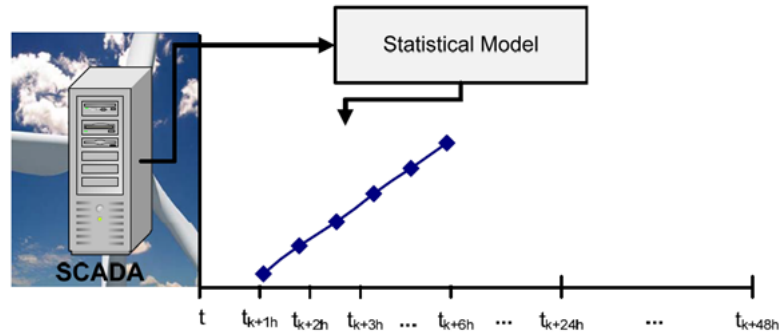
Figure 2.6: very short-term statistical approach [Monteiro et al., 2009]

In general, statistical models can be expressed mathematically as a generic
function, denoted as $f(i)$, added with white noise, e_w. The input, i, of the generic
function can tell us whether the approach should be classified as univariate
or multivariate. The input also indicate which time horizon the model is
appropriate for. Equation 2.12 and 2.13 shows the mathematical expressions of
the univariate and multivariate very short-term statistical models, respectively.
The input of equation 2.12 denotes the historical data of the variable one
wants to predict, namely the power production. In equation 2.13 there are
two input variables. One denoting the power production and one denoting
an exogenous variable e.g., wind speed. The expression can be expanded to
consider additional relevant weather variables.

$$\hat{P}_{(t+k|t)} = f(p_t, p_{t-1}, \cdots, p_{t-n}) + e_w \tag{2.12}$$

$$\hat{P}_{(t+k|t)} = f(p_t, p_{t-1}, \cdots, p_{t-n}, x_t, x_{t-1}, \cdots, x_{t-n}) + e_w \tag{2.13}$$

WPF approaches using NWP data (2) as input are normally considered as short-
term statistical approaches. The prediction horizon typically rages from 6 to 72

hours. In addition to NWP data, measured historical data (1) can be included to form a multivariate input case. The input sources and prediction horizon are illustrated in figure 2.7 [Monteiro et al., 2009]. Equation 2.14 gives the mathematical expression for the short-term statistical models having both NWP and measured historical data as input. What distinguishes this equation from equation 2.12 and 2.13 is that the generic function now incorporates the forecast of the weather variable , $\hat{x}$.

$$\hat{P}_{(t+k|t)} = f(p_t, p_{t-1}, \cdots, p_{t-n}, x_t, x_{t-1}, \cdots, x_{t-n}, \hat{x}_{t+1|t}, \cdots, \hat{x}_{t+s|t}) + e_w$$

$$(2.14)$$
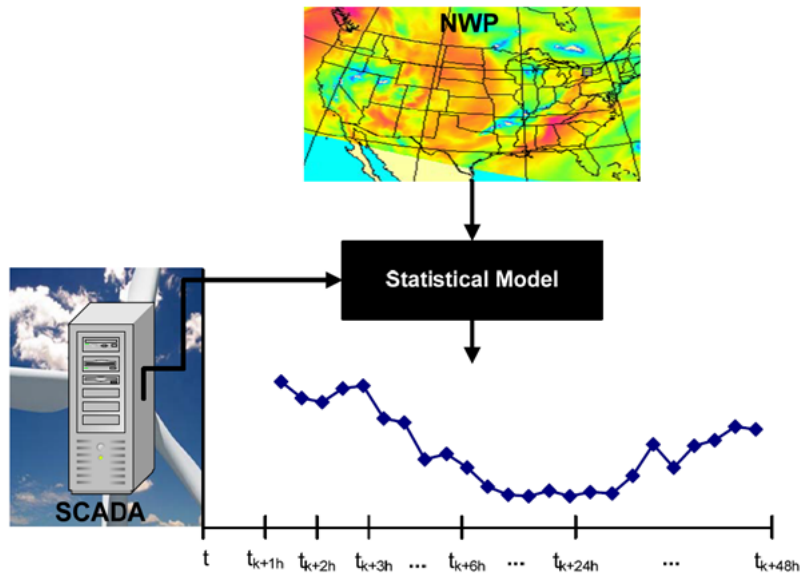
Figure 2.7: Short-term statistical approach [Monteiro et al., 2009]

Physical approach

WPF approaches that takes physical considerations using (3) TERRAIN data as input, in addition to (1) SCADA and/or (2) METEO FORECAST data, are classified as short-term physical approaches. This is illustrated in figure 2.8, proposed by [Monteiro et al., 2009]. Downscaling is typically a part of the physical approaches, where the input data (1) and/or (2) are refined to hub height. For this process information on the local roughness, orthography, atmospheric stability and the layout of the wind power park in the terrain are required.
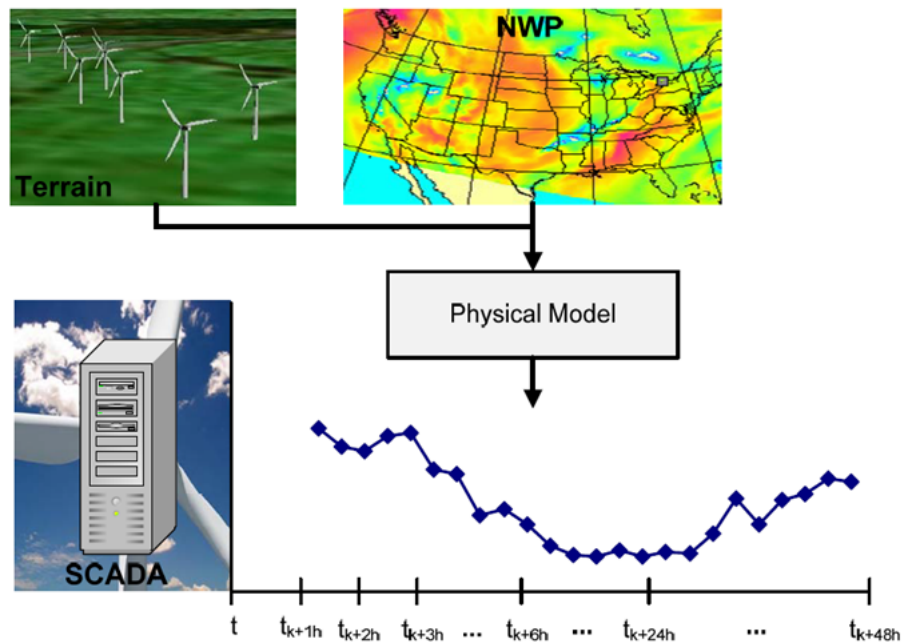
Figure 2.8: Short-term physical approach [Monteiro et al., 2009]

Methodology-based classification

The classification of WPF models may be based on the applied methodology, as proposed by Hanifi et al. [2020]. The paper distinguishes between three methodologies: naive, physical and statistical. The persistence and modified persistence described earlier under reference models are typical examples of naive methodologies. When it comes to the physical methodology, WindPRO offers modules to predict the future power production of a park taking physical considerations [Weir, 2014]. Statistical methodologies can be divided in two subclassifications: time series and machine learning based.

What is important to highlight in this section is the subclassification of the statistical methodologies. Note that despite the different subclassification, all statistical models have the same propose of mapping input to output. Time series based methodologies are normally applied on historical data aiming to estimate parameters of a mathematical model. ARMA (Auroregressiv Moving Average) is a typical example of this. This type of models are particularly applicable for a very-short-term prediction horizon. Machine learning method-ologies are based on learning algorithms that are tuned by comparing model prediction with the on-line measured power production. For learning algo-rithms it is important to split the data in training, validation and testing. The

training is for tuning of model parameters. The validation is for tuning of hyper-parameters. Finally, the testing is for reporting the expected performance of the learning algorithm for unseen data [Alpaydin, 2014]. There exist a wide range of machine learning methodologies. When it comes to WPF, artificial neural networks (ANNs) are common methodologies [Hanifi et al., 2020]. These are also known as deep learning models, which are more advance compared to other statistical techniques within machine learning [Alpaydin, 2014].

2.3.4   Machine learning models

Learning is a so called ill-posed problem, which means that the data by itself is not sufficient to find a unique solution for the problem [Alpaydin, 2014]. Hence, given finite data some additional assumptions are required. In machine learning literature, such assumptions are called inductive biases and are unique for each learning algorithm [Alpaydin, 2014]. For instance, in non-parametric modeling the assumption that similar inputs have similar outputs is an example of an inductive bias. A model can have several inductive biases. The choice of optimizer is another inductive bias. The inductive bias is the unique assumptions inherent in each learning algorithm [Alpaydin, 2014].

Classification of machine learning models

Géron [2019] classify different machine learning systems in broad categories. The classification is based on the three following criteria, which can be combined in any way.

- Is the desired output made available for the system by a supervisor or not? (Supervised versus unsupervised)

- Is the system capable of learning "incrementally on the fly" or not?

- How does the system generalize? Do the system asses the similarity of instances or is it based on model specific assumptions?

For the first criteria Géron [2019] distinguish between two main perspectives. That is supervised and unsupervised learning. In both perspectives there are input data. What distinguish these two perspectives is whether ground truth data is available or not. In supervised learning a supervisor has provided the correct output values. The purpose of supervised learning problems is to map the input to the desired output. Within supervised learning one distinguishes

between regression and classification. In regression problems the output is a number, while in classification it is a class code, e.g., where the only possible outcomes are zero or one [Alpaydin, 2014]. Note that WPF is time series forecasting. Time series forecasting can be framed as a supervised learning problem. Nevertheless it is different to other types of supervised learning problems [Brownlee, 2016]. Observations of a time series has a temporal order that must be preserved. Common ways of referring to time series forecasting are sequence prediction or self-supervised learning. The purpose of sequential prediction is to predict the next value of a time series [Brownlee, 2017].

For the second criteria one distinguishes between batch and online learning. In batch learning the system is trained on all the available data at once. The system is incapable of learning incrementally [Géron, 2019]. This means that the model parameters are updated after a complete search over the whole training set [Alpaydin, 2014]. In online learning, on the other hand, the system is capable of learning incrementally on the fly. An online learning system is gradually trained by feeding it either individual data instances or small groups of data instances. The instances are fed in a sequential order. The small groups of data instances are called mini batches. The advantage is that the system can learn from new data instances that are fed into the system [Géron, 2019]. When training neural networks online learning is usually used [Alpaydin, 2014]. The learning rate is one important parameter of online learning systems and it tells the system how fast it should adapt to new and different data. A system will rapidly be adjusted to new data given a high learning rate, but this also implies that the system tends to let old data slip. A system with a high learning rate is highly based on the latest and newest data the system was exposed to. In the opposite case of a low learning rate, the system adapts to new data in a smaller degree. It learns more slowly compared to the other case. This suggest a system that is less sensitive to noise and/or outliers in new data.

For the third criteria we have generalization, i.e., perform well on unseen instances, which is the overall goal of most machine learning problems. There are two main approaches to generalization: instance-based learning and model-based learning. In instance-based learning new data points are compared to known data points from training. The system can either be implemented to look at identical known data points or to look at similar known datapoints. In the latter case a measure of similarity is used to generalize to new data. In model-based learning, on the other hand, the purpose is to build a predictive model which detects patterns in the training data.

Regression tree

A decision tree store decision rules in a hierarchical data structure that can be associated with the structure of a tree with a root and growing branches. Decision trees are designed according to the divide-and-conquer strategy. In general, this means that a complex problem is broken down into sub-problems which are recursively solved to meet certain conditions. Among machine learning models, decision trees are referred to as white boxes because of the easy way of interpreting them through sets of IF-THEN statements. Alpaydin [2014] states that decision trees in many cases are preferred over more advanced models because it is interpretable. If more complicated models are to be investigated, it is generally recommended to start with a decision tree and use the test performance as a benchmark. When a decision tree is built to predict a single valued number, they are often referred to as regression trees (RT).

Decision trees are non-parametric models. The fundamental assumption is that similar inputs have similar outputs. The nonparametric algorithms put training instances in a "lookup table". In the case of a decision tree the lookup table is structured as a tree, composed of a root node, several internal decision nodes, and terminal leaves. Note that training instances are not stored one by one, they are stored in local regions based on a similarity measure. At each internal decision node, an exhaustive search is done to optimize a node impurity criterion. For instance, in regression the mean square error (MSE) is a common impurity measure, and the optimization criterion is to minimize it. During the training of the model, the input space is divided into local regions based on the similarity criterion and learn local models for each local region. Instead of defining one global model for the whole input-space, several local models are defined based on similar instances identified in the input-space during the training. When a new and unseen sample are fed into a trained tree, the algorithm will pass it to a local region representing similar instances and provide the prediction value stored for this local region. In other words, for each test input the corresponding local model is used to do the prediction [Alpaydin, 2014].

The process of storing the training instances requires memory and computational power, which is why these algorithms are also referred to as instance-based, or memory-based, learning algorithms in the machine learning literature. The complexity of the model is based on the size of the training set. A disadvantage with single decision trees is that they are sensitive to small variations in the dataset. A small change in the data set can give quite different results [Alpaydin, 2014].

CHART training algorythem

Scikit-Learn use the classification and regression tree (CHART) algorithm
to train the tree-based learners. The CHART algorithm starts out with split-
ting the training set in two subsets based on a single feature and a threshold.
The feature - threshold pair that gives the purest subsets are chosen based
on a search. The purest split is found by minimizing a certain cost function.
The impurity measure for the CHART cost function in regression is the mean
square error (MSE). After the first split is made, it splits the subsets using the
same approach, then the sub-subsets and so on, recursively. In this way the
process is repeated at each subsequent level, until it reaches the limitations
which is set for the tree by the hyper-parameters, or until the search does not
give any splits that reduce the MSE. In the case of a RT, each leaf node repre-
sents a real valued number. When a new instance is predicted it traverses the
tree starting at the rot and ends at a leaf node. The numeric value associated
with that leaf node becomes the predicted value for that instance [Géron, 2019].

The CHART algorithm has several hyper-parameters. Some of them are the
maximum depth of the tree (d), the minimum samples required at a leaf node
(e_n), the minimum samples required to do a split (e_s) and the maximum fea-
tures considered at each node (f). These are summarized in table 2.1. The
default values for the hyper-parameters set no restrictions for the tree, which
allows overfitting on the training data. The risk of overfitting is reduced by
restricting the maximum depth of the tree. Restricting the additional hyper-
parameters can also reduce the risk of overfitting. However, there also is a risk
of underfitting, so the restrictions must be chosen carefully through a validation
process [Géron, 2019]. Setting restrictions on a decision tree is often referred to
as pruning in the machine learning literature. Pruning can be implemented pre-
or post- training. In pre-pruning stopping conditions for the hyper-parameters
are applied, also referred to as early stopping. In post pruning backtracking
of a pure tree is used to identify and prune useless nodes [Alpaydin, 2014].
For a regression tree, Alessandrini and Sperati [2017] recommends tuning the
maximum depth of the tree or the minimum number of samples required at a
leaf node.

Random forest

A Random Forest (RF) algorithm is built as an ensemble of decision trees.
Instead of training only one decision tree, multiple decision trees are trained
on random subsets of training data. It is also possible to train each tree on
random subsets of the input features. In both cases, the prediction from each
tree is combined to provide a final overall prediction. The potential benefits

of combining multiple decision trees are a higher accuracy and a reduced instability compared to the case of a single tree [Alpaydin, 2014].

To combine the predictions from each tree the RF algorithm applies bootstrap aggregating, or bagging, as the voting method. Bootstrapping is used to generate a fixed number of slightly different subsets from the original training dataset. The sampling of instances from the original training set is done with replacement, which means that each instance in the dataset can be drawn more than once. In this way one instance can be higher represented then other instances. In summary, several training sets are generated using bagging and they are used to train separate decision trees. During testing the predictions from each decision tree are averaged to give the final prediction.

The generalization ability of the random forest algorithm can be controlled and optimized by tuning several hyper-parameters. There is a balance between over- and under-fitting, in addition to the tradeoff between accuracy and computational power. Table 2.1 summarize the hyper-parameters recommended for validation by Babar et al. [2020]. Note that all hyper-parameters, except the number of trees in the forest, yields the individual decision trees.

Table 2.1: Hyper-parameters for the tree-based models

Hyper-parameter	Description
t	How many trees to grow in the forest
f	How many features to consider in each node
d	What is the maximum depth a tree can grow
e_s	How many elements in a node needed to perform a split
e_n	How many elements needed to make a new leaf-node

Multilayer perceptron

The multilayer perceptron (MLP) is an artificial neural network (ANN) categorized as a feedforward network. It has been established that MLP is capable of estimating any smooth measurable mapping between an input vector and an output vector [Hornik et al., 1989]. Further, the MLP can learn extremely nonlinear functions and can be trained to generalize well on new, unseen data [Gardner and Dorling, 1998]. The multilayer perceptron (MLP) is based upon the principal of stacking multiple perceptrons [Géron, 2019], i.e., the simplest ANN architecture developed by Frank Rosenblatt in 1957 [Rosenblatt, 1958]. The perceptron simply consists of a single layer of parallel threshold logic units (TLUs), i.e., artificial neurons that represents a weighted sum modified by a

step function.

The architecture of a MLP is flexible and depends on the problem investi-
gated. A typical multilayer perceptron (MLP) is composed of one input layer,
one output layer and one or several hidden layers in between these two layers
[Ramchoun et al., 2016], as illustrated in Fig 2.9 [Géron, 2019]. All layers are
fully connected, that is all neurons are connected to every neuron in the previ-
ous and next layer. Each neuron has weighted input connections which form
the weighted sum. The step function, used in the perceptron, is replaced with
another simple nonlinear activation function, like for instance, the rectified
linear unit function (ReLu). It is the activation function that enables the MLP
to estimate highly non-linear functions [Gardner and Dorling, 1998].
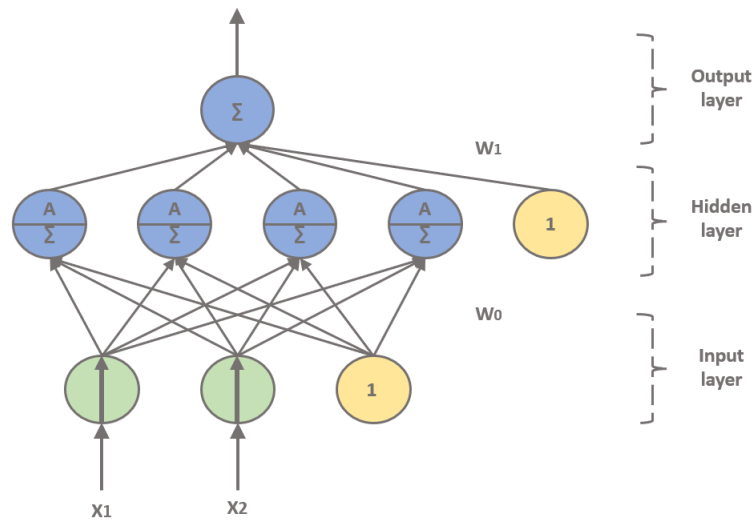
Figure 2.9: Architecture of a MLP with two inputs (green), a bias (yellow), one hid-
den layer of four neurons(blue with activation function) and one output
neuron(blue without activation function)

Through training the MLP learn and select a suitable set of connecting weights
that provides the lowest error. One common training algorithm is the back-
propagation algorithm, developed by Rumelhart et al. [1995]. A gradiant-based
optimization function is used to minimize a certain cost function [Gardner
and Dorling, 1998]. Gradient-decent is one traditional optimizer. The Adam
optimizer is a newer optimizer, proposed by Kingma and Ba [2014].

Going more in detail on the algorithm based on the literature by Géron [2019].
First, the weights of the network are initialized. The training dataset is re-

shaped into input output pairs and divided in mini batches i.e., groups of input output pairs. One mini-batch at a time is fed to the system and the number of epochs decides how many times the system runs through the training dataset. Each mini batch is sent through the network, in a forward pass. The output of each layer in the network is one at a time passed on to the next layer, until the output of the last layer is obtained. All layers are fully connected, and an activation function is applied on the weighted sum in each node. All intermediate results are stored in the network for the backward pass. Before the backward pass takes place, the output error of the last layer is measured based on the chosen loss function. During the backward pass the error gradient across all connection weights are calculated and propagated backwards. The purpose is to detect the error contribution from each layer. Finally, the algorithm performs a gradient-based optimization step to tune the weights and bias in the network based on the backpropagation result. This process of tuning the weights is repeated until the solution of the network, i.e., the weights of the network, converge. The final objective of this backpropagation training algorithm is to minimize a chosen cost function. The above description of the backpropagation algorithm goes under the category on-line learning i.e., the system is capable of learning incrementally on the fly [Gardner and Dorling, 1998].

2.3.5   Statistical performance measures

The error, ϵ of a predicted point is defined as the distance to the true labeled point. In other words, the difference between the measured point, P_t and the predicted point, $\hat{P}_t$, as illustrated in equation 2.15 [Madsen et al., 2005].

$$\epsilon = P_t - \hat{P}_t \tag{2.15}$$

Each prediction error consists of two components: the systematic and the random error of the model. The systematic error is referred to as the bias of the model and can be used to evaluate under or over estimation of the models. The bias can be expressed mathematically as follows [Madsen et al., 2005].

$$BIAS = \frac{1}{N} \sum_{t=1}^{N} (P_t - \hat{P}_t) \tag{2.16}$$

A standard performance measure, commonly used to evaluate performance of prediction models, is the root mean square error (RMSE). The mean square error (MSE) is the average of all the squared errors and the root mean square error (RMSE) is the root of the MSE. The mathematical expression is given in equation 2.17. When squaring the errors like this, the large errors are weighted higher than the small errors, which make the RMSE sensitive to

outliers [Alpaydin, 2014] [Madsen et al., 2005].

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (P_t - \hat{P}_t)^2} \qquad (2.17)$$

Another standard performance measure is the mean absolute error (MAE), which also is a distance measure. It takes the absolute value of the errors, making the sign of the errors irrelevant for the calculation. The mathematical expression is given in equation 2.18.

$$MAE = \frac{1}{N} \sum_{t=1}^{N} | P_t - \hat{P}_t | \qquad (2.18)$$

Both the RMSE and the MAE are scale dependent performance measures. This implies that fair comparisons of different models can only be obtained when the models are used on a data set with a common scale [Hyndman and Koehler, 2006]. The solution to overcome this problem is normalization. The measured time series, P_t , and the predicted time series, $\hat{P}_t$, are re-scaled to lay within the range between 0 and 1, before the performance measures are calculated. Equation 2.19 shows the mathematical expression of the min-max transformation of P_t. The performance measures are now referred to as normalized BIAS (NBIAS), normalized RMSE (NRMSE) and normalized MAE(NMAE). The NRMSE and NMAE gives the percentage of the maximum installed capacity of the park [Shcherbakov et al., 2013].

$$P_t^{nom} = \frac{P_t - P_{min}}{P_{max} - P_{min}} \qquad (2.19)$$

The explained variance score (EVS) can also be used to evaluate the performance of the wind power forecasting models. The EVS, given in equation 2.20, expresses the fraction to which a forecasting model accounts for the variations of the true, or measured, dataset. An EVS value of 1 indicates a model with perfect predictions. The lower the EVS, the worse the predictions are [Hanifi et al., 2020].

$$EVS = 1 - \frac{Var(P_t - \hat{P}_t)}{Var(P_t)} \qquad (2.20)$$

Figure 2.10 illustrates the amplitude and phase errors which are typical in wind speed and power forecasting [Hanifi et al., 2020]. The former appears when there is an over- or underestimation of weather parameters. The latter

appears when there is a time shift between modelled and measured weather data.
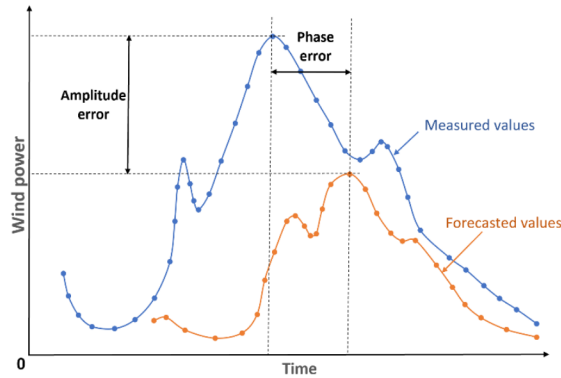
Figure 2.10: Typical amplitude and phase errors in weather forecasts [Hanifi et al., 2020]

The standard correlation coefficient, Pearson's r, is a measure for linear correlations. It is given in the range between -1 and 1, where a value close to 1 has a strong positive correlation, while a value close to -1 has a strong negative correlation and values close to one has no linear correlation. Note that Pearson's r only consider linear correlations. In other words, the coefficient will not necessarily be affected by non-linear relationships. The mathematical expression of the pearson's r is given in equation 2.21.

$$\rho_{a,b} = \frac{cov(a,b)}{\sigma_a \sigma_b},$$ (2.21)

where the numerator is the covariance between a and b. That is the expected value of the product between the deviation between the variables and the respective expected value. The denominator is the product of the standard deviation for each variable [Benesty et al., 2009].

Graphically a strong positive correlation will be recognized as the attribute at the y-axis tends to increase as the attribute at the x-axis tends to increase [Géron, 2019]. When the correlation coefficient is calculated between a series and a time-delayed version of the same series, it is referred to as the autocorrelation coefficient [Hoot et al., 2008].

/3

Method

3.1 The operational framework

A case-study based on real data is investigated with focus on a set of wind power forecasting (WPF) models used as offering, or trading, strategies on the day-ahead market. The operational framework of the trading strategies is set by the following:

- the site of the park and the installed capacity

- the power pool and the associated market frames

- the available data

3.1.1 The site of the wind power park

Fakken is a well established wind power park located at Vannøya, a small island in the Northern part of Norway. The total installed capacity at Fakken wind power park is 54 MW, consisting of 18 Vestas V90 turbines, each with 3 MW rated power and 80m hub height [TromsKraft, n.d.]. This information is summarized in table 3.1. The map in figure 3.1 shows the geographical location of the park at large and small scale. The geographical location is characterised by an arctic, cold and coastal climate. Figure 3.1 shows some of

the surrounding terrain which affects the wind resources at the park. The red dots represents the position of each turbine. Vannøya consists of various terrain, including mountain ranges and flat coastal terrain. The park is located at the south-east part of Vannøya at an elevation of 40-200m, where the terrain is relatively flat compared to the rest of the island. Even though Vannøya is a small island it has several mountain ranges. The highest mountain is 1031 metres above sea level and is located to the west of the wind power park. Vannøya is surrounded by typical coastal features in this region. South-east of the island you find Lyngsfjellan which separates two fjords. This topography creates two natural channels, where the wind potentially can form gap winds under certain conditions.Vannøya is also surrounded by other mountainous islands that may have a large effect on the wind resources at Fakken [Birkelund et al., 2018]. Mountain waves and other dynamically driven wind phenomena has large potential to occur in this type of orography [Markowski and Richardson, 2011][Jackson et al., 2013].

Table 3.1: Fakken wind power park

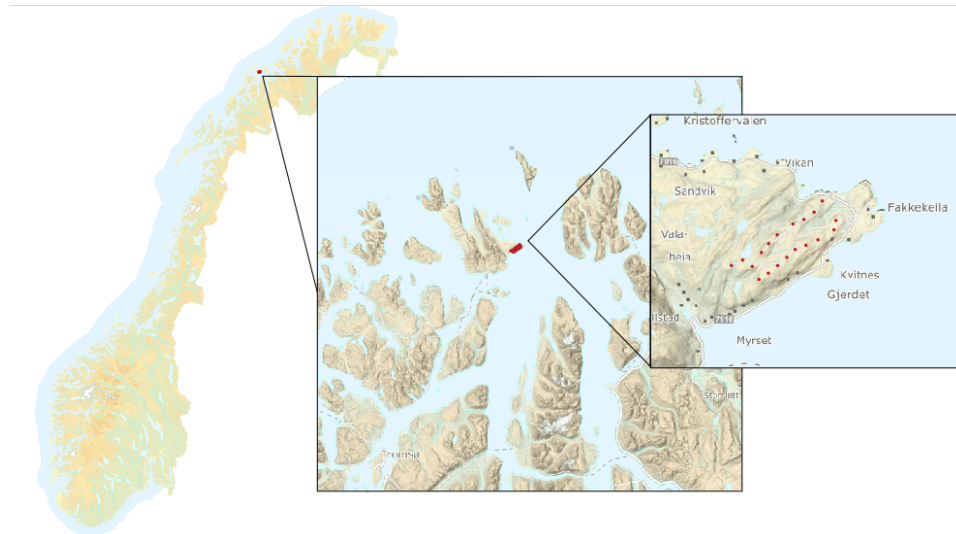| Numb. WTGs	Hub height	Rotor diameter	Installed effect
18	80m	90m	54MW

Figure 3.1: The geographical location of Fakken wind power park

3.1.2 The day-ahead market framework

The closure of the day-ahead market at 12:00 each day limit the data available for predicting the power output of the wind power park for the following day. The purpose is to predict the park power output for each hour the following day. The predictions must be submitted at the power-pool i.e., Nordpool for the Nordic countries, before the gate of the day-ahead market closes. The following assumptions are made for the data preparation for this study:

- The wind power productions forecast is made at t =11:00, which is one hour before the market gate closes. In other words, the prediction horizon range from 13 to 37 hours ahead.

- The weather forecast given at 06:00 is the last available predictions of weather variables. The variables used in this master thesis is meridional wind, zonal wind, temperature and pressure.

- The last available on-site measurement of power and weather variables is provided at t=11:00.

All the assumptions provided in section 2.2.4 applies and help to simplify the trading framework [Mazzi and Pinson, 2017].

3.1.3 Available data

The WPF models are applied on weather forecasts provided by two numerical weather prediction (NWP) models i.e., MEPS and AROME Arctic, operated by the Norwegian Meteorological Institute (MET Norway). Weather data from MEPS and AROME Arctic, and measured park power output data are available for the period ranging from 2017 to 2020. This collection of data form two separate datasets for time series forecasting: MEPS and AA dataset. Measured wind speed and wind direction are available for the entire year of 2017, while only measured wind speed is available for 2020. The market prices of zone NO4 at Nord Pool Spot are provided for both the day-ahead and balancing market in 2020.

3.1.4   Trading strategies

Eight different trading strategies for the day-ahead market are investigated for Fakken wind power park. The trading strategies are listed in table 3.2. Each trading strategy is based on different wind power forecasting (WPF) models. The WPF models are sorted in three categories according to three different methodologies from statistics and machine learning: naive, single-step and multi-step. In addition, a current forecasting method of the responsible power trader is included as a separate category under the name ISHK.

Three of the WPF models are based on naive forecasts methodologies and the remaining four are based on point-forecasting methodologies. Three of the point-forecast methodologies, are implemented as single-step models, which also is referred to as single-step-input, single-step-output (SISO) models. The idea of the SISO methodology is illustrated in figure 3.2. The last point-forecast model is implemented as a multi-step model, or multi-step-input single-step-output (MISO).

Table 3.2: Eight trading strategies based on different WPF methodologies

Strategy 0: operational power forecast	
	Strategy 0: ISKH
Strategy 1: naive forecast	
	Strategy 1A: Average
	Strategy 1B: Persistence
	Strategy 1C: New persistence
Strategy 2: SISO	
	Strategy 2A: NWP-PPC
	Strategy 2B: RT
	Strategy 2C: RF
Strategy 3: MISO	
	Strategy 3: MLP

The trading strategies built on the naive forecast methodology are based on power measurements available before the gate closure of the day-ahead market takes place. These trading strategies are used to describe a reference performance. In day-ahead trading the prediction horizon exceed six hours, which gives numerical weather prediction (NWP) models a central role [Giebel et al., 2011] [Alessandrini and Sperati, 2017]. These day-ahead prediction systems mainly use NWP data as input, leaving its performance dependent on the errors in the NWP model. The trading strategies built on the SISO and MISO methodologies use NWP data provided by the MET Norway i.e., MEPS and
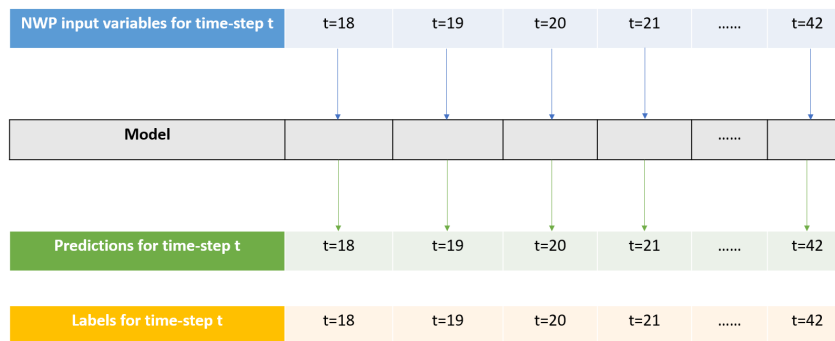
AROME Arctic, as input.

NWP input variables for time-step t	t=18	t=19	t=20	t=21	......	t=42
Model					......	
Predictions for time-step t	t=18	t=19	t=20	t=21	......	t=42
Labels for time-step t	t=18	t=19	t=20	t=21	......	t=42

Figure 3.2: Single step model with NWP data as input and power as output

3.2 Data and preparation steps

3.2.1 Power output data

In this master thesis power prediction are investigated with respect to the total wind park. Historical power output data is required to train the machine learning models. Historical power data for all power parks in Norway are made available for the public at NVEs websites. The power production for Fakken wind power park during the years 2017-2020 is used in this study. The data is measured at the grid connection point and are given at hourly intervals with the unit MWh [Sæther, 2021].

3.2.2 Meteorological weather data

In addition to historical on-site park power measurements, numerical weather prediction (NWP) data is used. The NWP data is provided by the Meteorological Institute of Norway (MET Norway), which daily provides operational weather forecasts four times a day, starting from 00:00, 06:00, 12:00 and 18:00. The NWP model used for the Nordic area goes under the name MetCoOp Ensemble Prediction System, also known as MEPS. MetCoOp stands for the Meteorological Cooperation on Operational Numeric Weather Prediction (NWP) between the Nordic countries: Norway, Sweden and Finland. The weather model (MEPS) runs on an operational domain covering the Nordic area. The domain spans 900 points in the zonal direction and 960 grid points in the

meridional direction. The output of the NWP model is a short time forecast with a 67-hour lead time, with a horizontal resolution of 2.5 km and with 65 vertical pressure levels. In other words, a high-resolution model [Frogner et al., 2019]. There is also a weather forecast model specialized for the European Arctic which is called AROME-Arctic (AA). It is classified as a regional high-resolution forecasting system. The horizontal resolution, number of vertical levels, number of times it is operationally provided, and the lead time are the same for the AA model as for the MEPS model. The operational domain of the AA model covers the European arctic including parts of Northern Norway [Müller et al., 2017]

The historical MEPS and AA forecasts for operational use have been available for the public on the website of MET Norway until recently. The AA forecasts were collected from this webpage in cooperation with Odin Foldvik Eikeland PhD Candidate at UiT The Arctic University of Tromsø. The weather parameters from AA are taken at 10 meters agl. The time series are extracted for the grid cells nearest the latitude and longitude of the measurement mast (70.099310, 20.092690). The MEPS forecasts used in this study were collected by Yngve Birkelund professor at UiT The Arctic University of Tromsø. The weather parameters from MEPS are taken at hub height i.e., 80 meters agl., and the statistical down-scaling approach is a spatial linear interpolation between the nearest grid cells of the measurement mast.

The weather variables of interest in this master thesis are the meridional wind component, the zonal wind component, the temperature two metres above the surface and the air pressure at the surface level. The raw NWP data for the four relevant weather variables are processed to form a set of time series. According to the operational framework set by the day-ahead market the last available weather forecast before the market gate closes is the one starting at 06:00. This forecast is based on meteorological measurements collected until 06:00. Since the data collection process and simulation process of the NWP model takes time, the forecast is provided several hours after the simulation model is started. However, the weather forecast starting at 06:00 each day are used to construct the time series. This is done by storing the 18th-42th data points from each 06:00 weather forecast for each of the weather variables. The time series are aggregated at hourly intervals for the four years: 2017-2020. The result is a dataset consisting of four features represented as columns and $(365 * 24 * 4 =)$ 35040 time-steps represented as rows. Figure 3.3 illustrate the four operational wind speed forecasts given on the 21 of January 2017. Note that the raw wind data from the NWP models is given in zonal and meridional wind components. For this illustration the components are transformed to speed and direction. A 06:00 forecast is illustrated in orange and a black marker at the x-axis at 11:00 illustrates the time when the forecast is provided at the day-ahead market. The black dotted line along the x-axis illustrates the 14th-42nd time steps of the 06:00 forecast, which are extracted to construct the time series.
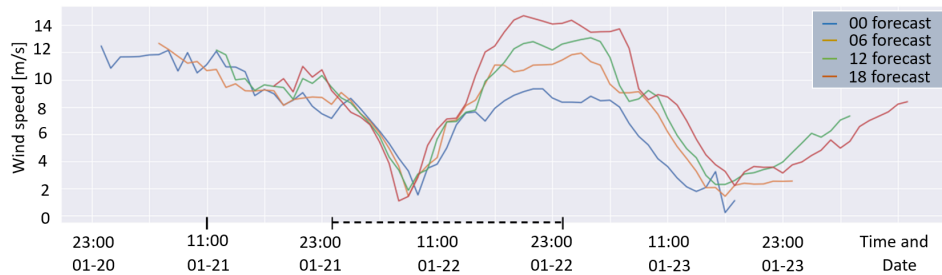
Figure 3.3: The four operational wind speed forecasts given on the 21 of January 2017

3.2.3 On-site weather measurements

On-site measurements of wind speed, wind direction, air temperature and air pressure for the entire year 2017 are provided by Troms Kraft. The location of the measurement mast is given by a latitude of 70.099310 and a longitude of 20.092690. The measured weather is compared with the modelled weather data, to report the performance of the NWP models.

On-site wind speed measurements for each of the 18 turbines are provided by Troms Kraft for the entire year 2020. The measured wind speeds over a case-period are compared with modelled wind speeds, to inspect the error in the input of the WPF models.

3.2.4 Market price dataset

Hourly spot and regulating prices in 2020 for the market region of Fakken wind power park, N04, is provided by Finn Dag Hovem from Ishavskraft. The market price data is used for the economic performance evaluation of the trading strategies. The total revenue and imbalance cost related to each wind power prediction model is calculated. The calculations are based on the one price system implemented in Norway recently. Note that the actual pricing system in 2020 was based on two-prices. The presented calculations will therefore not reflect the true revenue or cost of the power trader in 2020, rather it will reflect what the revenue or cost could have been with the new single-pricing system implemented at the power market today and which is the basis for the future.

The year 2020 has shown to be an exceptional year with the historical lowest electricity prices in Norway. It represent the all-time low. The average power

price of 2020 in Norway was 9.2 EUR/MWh with a reduction of 54% from the last all-time low in 2015. In addition, negative power prices were experienced in the price areas in southern Norway for the first time [Langset and Nielsen, 2021]. It is also worth mentioning that 2021 represent the all-time high power prices, particularly in the southern parts of Norway [Birkelund et al., 2021]. More volatile power prices, as we have seen for 2020-2021, is expected for the coming years as the electricity system becomes more weather dependent.

3.2.5   Preparation steps

The raw data provided by MET and Troms Kraft must be prepared according to the operational framework before using it as input to any of the WPF models. Géron [2019] recommend six data preparation steps for machine learning projects: data exploration, data cleaning, feature selection, feature engineering, data splitting and feature scaling. These steps are used in this study to prepare a park dataset , and each step will be further explained in this chapter. The park dataset consists of four years of data ranging form 2017 to 2020, where the park power output provided by NVE is the target and the predicted weather variables provided by MET are the features.

Data exploration

A pre-analysis on parts of the data is done to gain insight. The part of the data which is analysed is referred to as an exploration set in this thesis. The data of 2017 is chosen as the exploration set as on-site measurements are available for this period. The correlation is explored based on the standard correlation coefficient between measured and modelled time series . The NRMSE is evaluated between the time series of measured and modelled attributes. These measures can tell us how well the NWP model perform and if the attributes can provide useful and qualified information for the machine learning models.

The correlation is in addition visualized using two-dimensional histograms. These plots can give us an idea about where the density of data points is highest [Géron, 2019]. The data relationship between measured park power and different attributes are investigated through two-dimensional histograms. A two-dimensional histogram combines two one-dimensional histograms into one. The histogram of one attribute is following the y-axis on the right-hand side, while the histogram of another attribute is following the x-axis on top. The two-dimensional histogram appears in the grid framed by the x- and y-axis.

The appearance of dark blue color in the two-dimensional histogram indicates a combination of the two attributes that frequently appear, while a lighter blue color indicates that the combination occurs at a much lower frequency. This explanation applies to all the two-dimensional histograms of different attributes for the y- and x-axis.

Feature selection and engineering

One essential goal of this study is to predict the future power production of Fakken wind power park. The power production of any wind power park highly depends on the wind resources in combination with the limitations set by the turbine design installed at the park, as described in equation 2.3. This makes wind a naturally attribute to include. The wind can be described on vector form i.e, zonal and meridional wind, or on component form using wind speed and direction. The NWP wind data is usually given at vector form. In general, it is recommended to input the wind data to a neural network at vector form, while regression trees does not have any recommendations for the input [Giebel et al., 2011]. For this study the wind data transformation giving the best accuracy is chosen.

Besides wind data, there are other weather parameters that can provide useful information. From equation 2.3 we find that density of air, ρ, has an impact on the kinetic power of wind. Deciding factors of air density are air pressure and air temperature, as well as the molecular composition of air. Air pressure and temperature are recommended to use as attributes for wind power forecasting by Optis and Perr-Sauer [2019]. The temperature can serve as an indication to the model about which season it is. This might learn the models to distinguish between winter mounts with potentially high winds and summer mounts with considerably lower winds. The reason for these seasonal variations is linked to the formation of additional low-pressure systems in the winter due to the temperature differences between the poles and equator. Forming low pressure systems is how the nature tries to equalize the temperature differences. This shows how there is a continuous interplay between the wind, temperature and pressure [Donald Ahrens and Henson, 2015]. The terrain surrounding the park and the climate in the area also has turned out to be important factors for the wind regime at the location of the park [Jackson et al., 2013] [Markowski and Richardson, 2011]. For this study weather parameters describing the wind, temperature, and pressure at the location of the wind power park, are evaluated as features.

Data cleaning

The raw MEPS and AA datasets are inspected for missing values and certain obvious anomalies. A cleanup is done to fix, remove or fill in for the data points [Géron, 2019]. Both the on-site measurements provided by Troms Kraft and the NWP data provided by MET Norway has time periods where data is missing. Most statistical and machine learning algorithms does not work well with missing values. In this study all missing values that are identified in each time series are filled with the last available valid value in the respective series.

The raw power data contain a lot of uncertainties, especially in from anomalies. The reasons for anomalies are potentially many and complex and will not be examined further. In this study a physical explanatory model for some obvious anomalies, referred to as bottom-curve stacked by Wang et al. [2019], are applied. They represent the samples were the power production is zero or negative while the wind speed is high, as illustrated in figure 3.4. If the wind speed is in between cut-in and cut-out wind speed while the park power production is zero, the data for this timestep is identified as an outlier. All identified anomalies are removed from the training part of the dataset, but note that this is not applied for the test dataset. The advantage of removing anomalies is that the wind power prediction models only learn from normal and valid data, and the model may predict higher power values than before.
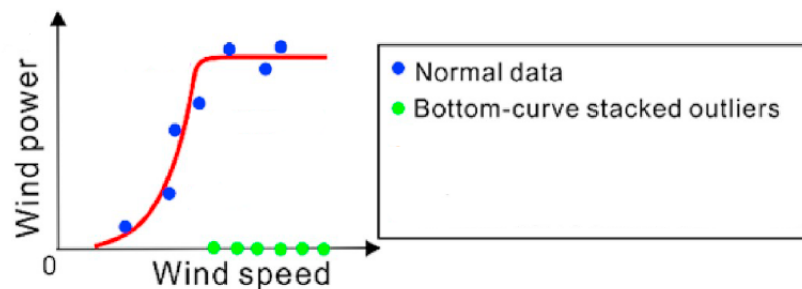
Figure 3.4: Classification: normal data vs bottom-curve stacked anomalies [Wang et al., 2019]

Split the data

In most machine learning projects the aim is not to replicate the available historical data but to provide accurate predictions for new cases [Alpaydin, 2014] [Géron, 2019]. One want the model to be able to generate the right output for a new instance, despite that the correct output has not been learned in the training. This concept is called generalization. In order to obtain the best

generalization for a model we need to have access to data outside the training. A common way of simulating this is to divide the available data into training, validation and test sets. It is highly important to keep the validation and test sets unseen for the model throughout the training. When the validation is done, this is also seen as part of the training. This is to avoid overfitting where the model does accurate predictions on the training set but fails on unseen data.

The number of samples that are used for learning the model is dramatically reduced when using the traditional splitting where the available data is divided into three parts. One way of solving this issue is by using cross-validation [Géron, 2019]. One common way of doing cross-validation is to divide the training set in a number of (f) smaller datasets. These are often referred to as folds. In the validation process the folds are placed in a loop, such that each fold is used for validation while the remaining data are used for training. Then the model is trained for f-1 of the folds and validated for the fold that was held outside the training [Alpaydin, 2014]. In time series prediction where a single-step-input to single-step-output methodology is applied, such cross-validation will work fine. However, if moving to multistep-input or output, which means that several time steps are considered, one should be careful using cross-validation. This is because the sliding window at some time will include time steps that originally are not aligned in time.

Table 3.3 shows the splits applied for the two park datasets in this study. Taking the MEPS dataset as example. It spans four years of data ranging from 2017 to 2020. The last year, 2020, which counts for 25% of the total dataset, is set aside for testing leaving the remaining 75% for training and validation. The hyper-parameters are tuned using a grid search on the validation set.

Table 3.3: The split of the MEPS and AA dataset

Training	Validation	Test
80% of 2017-2019 data	20% of 2017-2019 data	whole 2020

Feature scaling

When the features have very different scales, machine learning algorithms normally does not perform well. In such cases it is important to scale the features to lay within the same range. Normalization, also known as min-max scaling, is one way of scaling the features. The values are re-scaled to fit the

range between 0 and 1. Equation 2.19 shows the mathematical expression of the transformation. One important note is that the scaling takes place after the data splitting, because the scaling parameters i.e., maximum and minimum, can only be based on the training data, not the full dataset. This is in order to hold the test data unseen and new for the models until testing. Another note is that the scaling parameters must be saved for later, so that inverse transformation can be done later after predictions are done.

3.3   Model setup

3.3.1   Naive forecast methods

Three naive forecast models are implemented based on park power measurements available before the gate closure of the day ahead market. The models include the average power production over the training set, the well-known persistence model and the modified version of the persistence model, first proposed by Nielsen et al. [1998]. That is the AVG model, the PE model and the NPE, respectively. These models define the reference for model performance, and other models are evaluated with respect to improvement compared to this reference.

The reference models are implemented using standard built-in functions in python. When implementing the naive persistence model, the measured park power output at 11:00 provided by NVE is used as the predicted value for each hour the following day. This method does not require any training data, it only requires the last available park power measurement.

The modified persistence (NPE) model given in equation 2.11 require a training set in order to estimate the mean power output, $\bar{p}$, and the correlation coefficient, a_s. These are estimated between p_t and p_{t+s} for each time step , s $\in [14, 38]$. The correlation coefficient for each time step can be visualized by the auto correlation plot given in figure 3.5. The bins corresponding to the 24 time-steps of interest are marked.
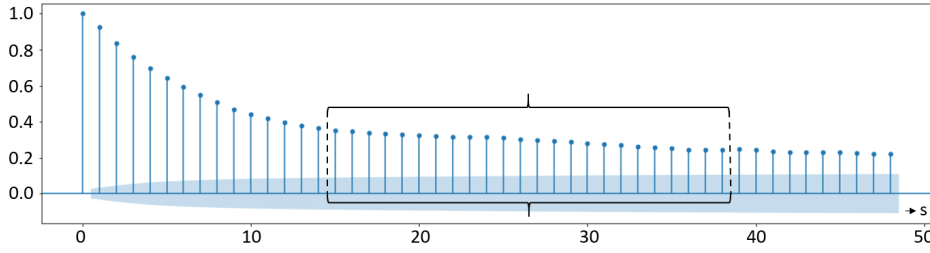
Figure 3.5: Auto correlation based on the park power time series for the traning period ranging from 2017 to 2019

3.3.2   NWP-PPC

The direction-specific park power curve provided by Kjeller Vindteknikk [Weir, 2014], given earlier in figure 2.2, is used as inspiration for a NWP park power curve estimation method (NWP-PPC). In this study each power curve is obtained from statistical curve fitting on the training data, where the purpose is to find the curve that best fits the dataset. The statistical fitting methods investigated in this study are the average and the median of all power outputs inside 12 wind sectors and 30 wind levels, with a resolution of 30 degrees and 1 m/s, respectively. This curve fitting process is based on the training dataset. The input features of the model are the NWP wind speed and NWP wind direction, while the target are the park power output. A tuned direction-specific PPC can be used for park power predictions using NWP data as input. The power output of the park can then be obtained as following [Birkelund et al., 2018].

$$\hat{P} = PPC[WS_{nwp}, WD_{nwp}] \tag{3.1}$$

Where $\hat{P}$ is the predicted park power output, WS_{nwp} is the NWP wind speed, WD_{nwp} is the NWP wind direction and PPC is the direction-specific power curve. The PPC map the inputs, which are the NWP wind speed and the NWP wind direction, to the output ($\hat{P}$) using linear interpolation.

3.3.3   Regression Tree

The regression tree algorithm is implemented as a single-step-model where the predicted weather, for a certain time step ahead, is fed into the model to get the wind power prediction for the same time-step. This methodology is illustrated in figure 3.2. The regression tree is built using the CHART algorithm through the scikit learn library in python. The CHART algorithm has several hyper-parameters as described in section 2.3.4. Pre-pruning is applied to avoid growing a complex and overfitted regression tree. For the results of this study

pre-pruning, i.e. early stopping, is applied for the max-depth (d). The additional hyper-parameters are kept at their default values. For the results presented, the min-samples-leaf is not validated since it did not provide any performance improvements. A grid search is applied for the pre-defined validation split, described in section 3.2.5, to tune the d parameter The search space is set to integers between 1 and 20. The search algorithm use the MSE regression loss to evaluate each parameter.

3.3.4   Random Forest

The Random Forest (RF) model is also implemented as a single-step-model where the predicted weather, for a certain time-step ahead, is fed into the model to get the wind power prediction for the same time-step. The RF is an ensemble of RTs, where each RT is based on the CHART algorithm from scikit learn. The max-depth (d), the min-samples-leaf (e_n) and numb-of-trees (t) are tuned to avoid over-fitting of the RF model. The latter parameter is special for ensemble learners to decide the number of estimators needed. A randomized grid search is applied at the pre-defined validation set to tune the hyper-parameters. The search space for the d parameter is set to the integers ranging from 5 to 15. For the e_n parameter it is set to [1,2,4,6,8,10,50]. The search space of the number of trees in the forest is set to [50, 100, 200, 300, 400, 500].

3.3.5   Multilayer perceptron

The multilayer perceptron (MLP) model is implemented as multi-step input single step output (MISO) model. The predicted weather for three time-steps is fed into the model to get the wind power prediction for the third, and last, time-step of the sequence. This methodology can be referred to as sequence learning and differs from the single-step learning methodology, which were used for the NWP-PPC, RT and RF model. The sequential application programming interface (API) of Tensorflow-Keras is used for this purpose.

The initial dataset described in section 3.2.5 is created with one output column and several input columns, such that a one step-model can learn how to predict the output from the input. However, to use the same dataset for sequence learning it need to be transformed into input output pairs, where the input is a window and the output is a single value. In this study, each window contains a sequence of three-time-steps for each feature. When two features e.g., NWP wind speed and NWP wind direction, are used, the windows contain three-time steps for two features. The output is the power for the third and last time-step represented in the window. For the MLP to be able to learn from the

created input output pairs, the input is given as a one-dimensional vector. Thus, the window is reshaped to form a vector. The consequence of transforming the initial dataset into a sequence learning dataset is that some samples are removed. With a three time-step window, the two first observation of the initial dataset are dropped and there might also be samples at the end of the dataset that are dropped. This make the training data for the MLP slightly different from the training data used for previous point-forecasting models (NWP-PPC, RT and RF) [Géron, 2019] [Brownlee, 2017].

The hyper-parameters chosen for fine-tuning are the number of hidden layers, the number of neurons in each hidden layer and the learning rate. The tuning is done using a randomized grid search algorithm for the predefined validation set. Table 3.4 summarizes the search space of the hyper-parameters for each model. For the optimization the Adam optimizer is chosen. The hidden activation is ReLu and the loss function is the MSE. These design choices are inspired by Géron [2019].

Table 3.4: Grid search space of hyper-parameters for the RT, RF and MLP model

Model	hyper-parameter	Validation grid space
RT	max-depth	integers $\in [5, 15]$
RF	numb-of-trees	50, 100, 200, 300, 400, 500
	max-depth	integers $\in [5, 15]$
	min-samples-leaf	1,2,4,6,8,10,50
MLP	# input neurons	None
	n-neurons	range from 1 to 100
	n-hidden	[0, 1, 2, 3]
	learning-rate	[10e-5,10e-3,10e-1,10]
	# output neurons	1
	Hidden activation	ReLu
	Loss function	MSE

3.3.6 The ISHK model

The ISHK model is a current forecasting method used as trading strategy by the responsible power trader for Fakken wind power park. It represent the bids submitted at Nord Pool in 2020. The model is based on operational NWP data. It is not given how the forecast is made, or which input features that are used for training and during operation. Nevertheless, the normalization of the performance measures makes it possible to compare the operational results with the academic results.

3.4   Performance evaluation

The performance of the trading strategies are evaluated using several performance measures. Some are based on distance measures commonly used in machine learning, and some are based on economic measures often used in a trading perspective. More precisely, the wind power prediction models behind each market trading strategy is evaluated using the statistical performance measures as recommended by Madsen et al. [2005]. This includes the normalized bias (NBIAS), the normalized mean absolute error (NMAE) and the normalized root mean square error (NRMSE). These measures indicates how much error typically is provided in the forecasting method.

Each market trading strategy is also evaluated according to revenue for the wind power trader in a single price system. As described in the theory and repeated in equation 3.2, the total revenue is the difference between the precise information term, r_s^{PI}, and the imbalance term, r_s^{IMB}. The economic performance parameter is based on the ratio between these two terms and will be used to evaluate the effectiveness of the different trading strategies [Mazzi and Pinson, 2017].

To make the two terms: the precise information term, r_s^{PI}, and the imbalance term, r_s^{IMB}, of equation 3.2 easier to interpret the variable used in the theory part is replaced with more intuitive names. The market prices and the power volumes are denoted as P and V, respectively. The subscripts denoting the day-ahead market (DA), the balancing market(B) and the time step (s) are kept as in the theory part. The wind power production measured for the hourly interval is now denoted by V_s. The new representation is given in equation 3.3 and 3.4.

$$r_s = r_s^{PI} - r_s^{IMB} \tag{3.2}$$

$$r_s^{PI} = P_s^{DA} * V_s \tag{3.3}$$

$$r_s^{IMB} = (P_s^B - P_s^{DA})(V_s^{DA} - V_s) \tag{3.4}$$

The values of r^{PI} and r^{IMB}, for one year can be calculated as following, where $d \in [1, 365]$ are days and $s \in [13, 36]$ are time steps. These expressions will be used to give the total revenue for each trading strategy over the entire 2020.

$$r^{PI} = \sum_{d=1}^{365} \sum_{s=13}^{36} r_s^{PI} \tag{3.5}$$

$$r^{IMB} = \sum_{d=1}^{365} \sum_{s=13}^{36} r_s^{IMB} \tag{3.6}$$

$$r = r^{PI} - r^{IMB} \tag{3.7}$$

The performance ratio is calculated for the entire 2020, as in the following equation,

$$f = \frac{r}{r^{PI}} = 1 - \frac{r^{IMB}}{r^{PI}}. \tag{3.8}$$

/4

Results

4.1   Raw data analysis

4.1.1   The MEPS dataset

Time series

Figure 4.1 shows the five time series of the MEPS dataset: park power output, wind speed ($\hat{ws}$), wind direction ($\hat{wd}$), air pressure ($\hat{P}$) and air temperature ($\hat{T}$). Only the training period ranging from 2017 to 2019 is illustrated. The test period (2020) is held outside the analysis. Keeping it unseen until the performance evaluation.

The temperature shows a strong seasonal variation, with low temperatures in the winter and high temperatures in the summer. Aligned with the peeks in temperature, the wind speed and park power output drop. Low wind speeds and power production are associated with the warmest summer months. The wind speed shows small seasonal variations. There are higher wind speeds in the winter compared to the summer.

Since the wind direction is a cyclic parameter, some problems arise when displaying the time series. For the hours where the wind direction changes from 0 to 360 degrees, or visa-versa, large variations will appear in the time series plot, while in reality the change is insignificant. No seasonal variations are observed for the wind direction. The air pressure varies less in the summer

months compared to rest of the year. The frequency of the variations might be higher in the summer months, but the range of the variations are smaller than the rest of the year. Coupling the observation of the figure, the months with the highest park power production are the months with the lowest temperatures, the most stable pressure systems and the highest winds.
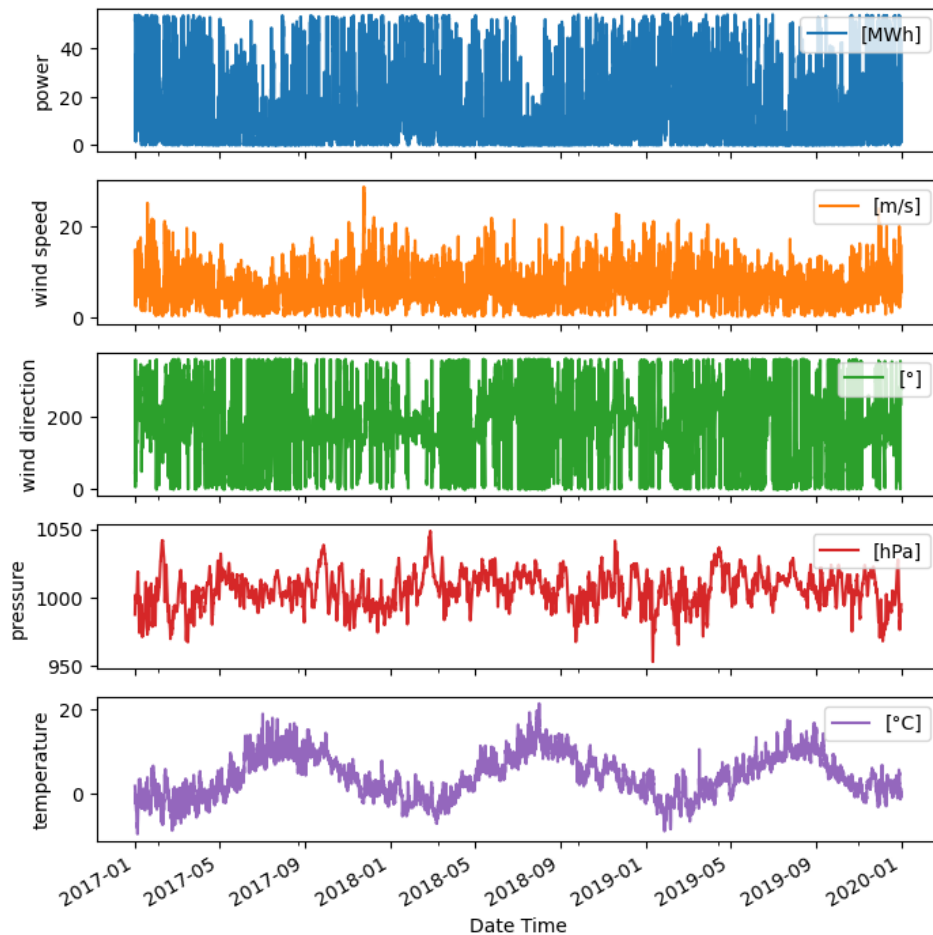
Figure 4.1: Time series over the training period(2017-2019) for the MEPS weather parameters

In a day-ahead forecasting perspective, the daily variations of the input parameters are important and essential to investigate. Figure 4.2 illustrates daily variations of the modelled weather parameters for three consecutive days in October 2017 and compare them to on-site measurements. The variations and intermittency of the wind is visualized.

For the three days, amplitude and phase errors are observed. The latter especially appears in cases of wind speed ramps i.e., large and fast variations in wind speed. Around 12:00-14:00 the October 16th a ramp-like decrease is observed two hours earlier for the modelled wind speed compared to the measured wind speed. This is a typical phase error. Earlier the same day, the ramp amplitude of the modelled wind speed was considerably lower than for the measured. At midnight the October 17th the wind direction appears to have a large and fast variation. Physically this is an insignificant variation because of to the circularity of the parameter.

The air pressure typically has variation with smooth transitions from low to high pressure levels. In this example, the modelled air pressure follows the same pattern as the measured parameter, with a relatively constant amplitude error. The air temperature typically shows a peak in temperature during a day. In this case the modelled and measured temperature has small deviations for the first and last day. For the day in the middle, October 16th, there are both amplitude and phase errors observed.

As described in section 3.2.2 daily time series are sequentially combined to form the NWP datasets. At midnight there is a distinction between the forecast value with a 36-hour prediction horizon and the more updated forecast value with an 18-hour prediction horizon. This can be seen in the figure below.
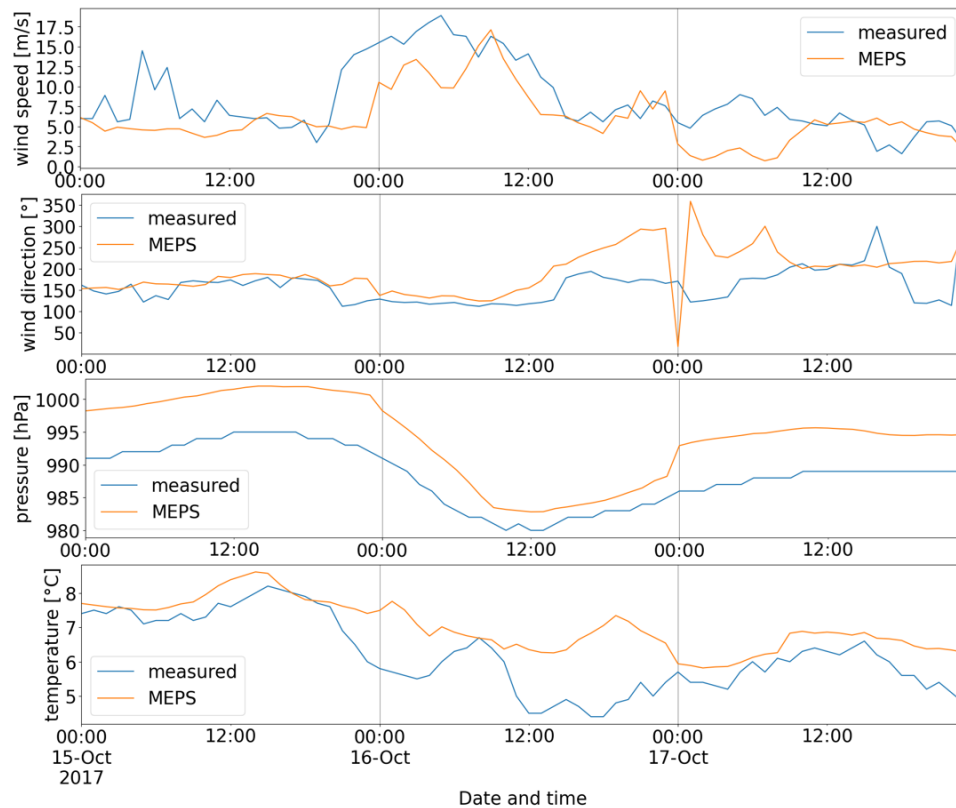
Figure 4.2: Daily variation of the MEPS weather parameters

Wind rose and terrain

Figure 4.3a shows the wind rose i.e., the circular histogram, or rose diagram, for Fakken wind power park. It is based on the modelled wind speed and direction from MEPS during the training period i.e., 2017 to 2019. The wind rose illustrates the wind modelled resources and gives a realistic illustration of the cyclic wind direction, in contrast to the time series.

Two dominant wind directions, 135 degrees and 200 degrees, are illustrated. These directions are referred to as South-East (S-E) and South-South-West (S-S-W), respectively. In the S-S-W sector the low to moderate wind speeds, given in blue tones, most frequently occurs. This is also the situation for the S-E sector. In addition, the maximum wind speeds, marked with yellow bins, are represented in the S-E sector. The sectors where the wind has the lowest occurrence are the east (E) and west (W) sectors. Compared to these sectors, the wind is well represented in the north (N) sector. However, compared to the S-S-W sector, the frequency in the N sector is 50% lower.
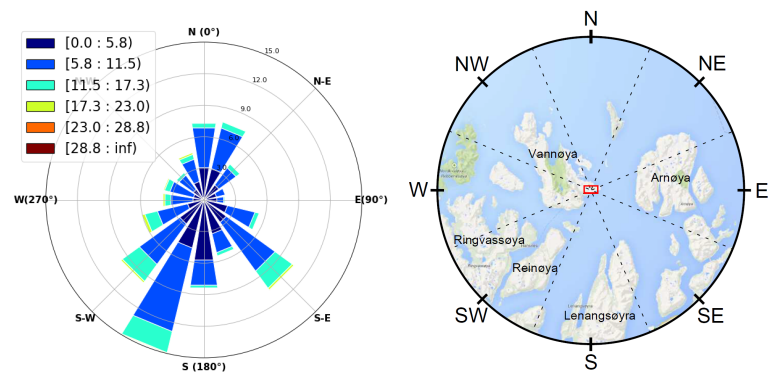
(a) Wind rose based on MEPS wind speed and direction during 2017-2019

(b) The surrounding topography and terrain of Fakken wind power park (red) [Jacobsen, 2014]

Figure 4.3b, provided by Jacobsen [2014], illustrates the terrain surrounding Fakken wind power park. Fakken wind power park is centered in the middle of the figure and is marked with red. This figure can help to set the wind rose in context with the surrounding terrain and roughness. Vannøya consists of various terrain, including mountain ranges and flat coastal terrain. The wind power park is located at the flat south east part of the island with one of the mountain ranges in the west sector. Vannøya is surrounded by several mountainous islands: Arnøya, Lenangsøya, Ringvassøya and Reinøya.

Coupling figure 4.3a and 4.3b, it is evident that the two fjords on each side of Lenangsøya plays a dominant role for the modelled wind resources at the park. The open sea in the north sector can also explain the fairly good wind resources in that direction. The mountain range in the west sector and the mountainous island Arnøaya in the east sector, can explain the low occurrence of wind in these sectors. This type of terrain facilitates dynamically driven winds, like blocking, gap winds and mountain waves, given the right respective atmospheric conditions [Solbakken et al., 2021].

Statistics and correlation

In this section an exploration period of the MEPS weather data is analyzed and compared to on-site measurements. The exploration period is set to 2017, since measurements are available for this period. The exploration data is raw data without any pre-processing steps applied.

Table 4.1 provides descriptive statistics for each time series in the MEPS exploration dataset. For the modelled wind speed and air temperature there are remarkable deviations from the respective measured parameters. The maximum measured wind speed over the exploration period is five m/s higher than for the modelled wind speed, which is significant. Hence, the mean for the measured wind speed is naturally slightly higher than for the modelled parameter. From the standard deviation it is observed that the measured wind speed data is more widely spread in relation to its mean, in comparison to the modelled data. The temperature measurements for January to late April are unreliable and are therefore kept outside the statistical measures. Nevertheless, there are deviations for the temperature. The modelled mean are higher than the measured, while the spread in relation to the mean value is less for the modelled data compared to the measured data. For air pressure the modelled data has a higher mean but lower standard deviation, compared to the measured data. The cyclic wind direction requires circular calculations of the mean and standard deviation. These are not provided in this thesis because it is not a crucial stage for further research.

Table 4.1: Descriptive statistics of the time series in the MEPS dataset for 2017, compared to measurements for the same year

	count	mean	std	min	max
ws	8760.0	7.7	4.6	0.0	34.1
$\hat{ws}$	8760.0	7.1	3.9	0.1	28.8
wd	8760.0	-	-	0.0	359.0
$\hat{wd}$	8760.0	-	-	0.0	360.0
P	8760.0	997.7	13.0	963.0	1038.0
$\hat{P}$	8760.0	1002.5	12.8	967.1	1041.7
T	5952.0	5.9	4.7	-5.9	19.2
$\hat{T}$	5952.0	6.4	4.3	-3.2	19.0

Table 4.2 gives the linear correlation coefficient and the NRMSE between measured and modelled MEPS weather parameters, excluding the cyclic wind direction. The strongest positive correlation and the lowest error measure are found for the pressure parameter. It shows a correlation coefficient of 0.986 and a NRMSE of 0.002. In other words, a significant strong linear correlation and almost no errors. The measures for the wind speed parameter also show a quite strong positive correlation with a correlation coefficient of 0.708, and a relatively low NRMSE of 0.105.

Moreover, the table includes the statistical measures for the zonal (U) and

meridional (V) wind. The modelled and measured zonal wind, typically east-
erly or westerly winds, has a stronger positive correlation than the meridional
wind, typically southerly or northerly winds. In other words, the largest devia-
tion is associated with the sector where the wind most frequently occur.

Table 4.2: The linear correlation coefficient and the NRMSE for pairs of measured and
modelled weather parameters for the MEPS dataset

parameter	NRMSE	correlation coefficient
ws	0.105	0.708
U	0.163	0.844
V	0.164	0.735
P	0.002	0.986
T	0.209	0.690

The correlation results are visualized using two-dimensional histograms. Fig-
ure 4.4a, 4.4b, 4.4c and 4.4d shows the two-dimensional histograms for wind
speed, wind direction, air pressure and temperature, respectively. The visual-
ization of the wind speed, air pressure and temperature coincides with the
linear correlation coefficient given in table 4.2. The very strong positive corre-
lation of air pressure appears as a narrow diagonal line in the two-dimensional
histogram. The modelled air pressure tends to increase as the measured air
pressure increases. The appearance of more scattered or spread bins along the
diagonal represents a slightly weaker correlation, which is the case for wind
speed. The bins are more widely spread around the diagonal, and they are
centered around low wind speeds. Low to moderate wind speeds are most
frequently occurring at the site.

The fact that the wind direction is a cyclic parameter can help to under-
stand and interpret the two-dimensional histogram of the wind direction in
figure 4.4b. The most frequently occurring bins lays on the diagonal, but there
are some strong bins appearing in the upper left and lower right corner as
well. For instance, the values in the upper left corner are measured to be zero
degrees while the MEPS model predict the direction to be 360 degrees. Since
the wind direction is a cyclic parameter, these values represent the same di-
rection. This means that the bins appearing in the upper left and lower right
corner physically lies close to the diagonal. Thus, physically there is a stronger
positive correlation between measured and modelled wind direction.

(a) wind speed

(b) wind direction

(c) air pressure

(d) air temperature
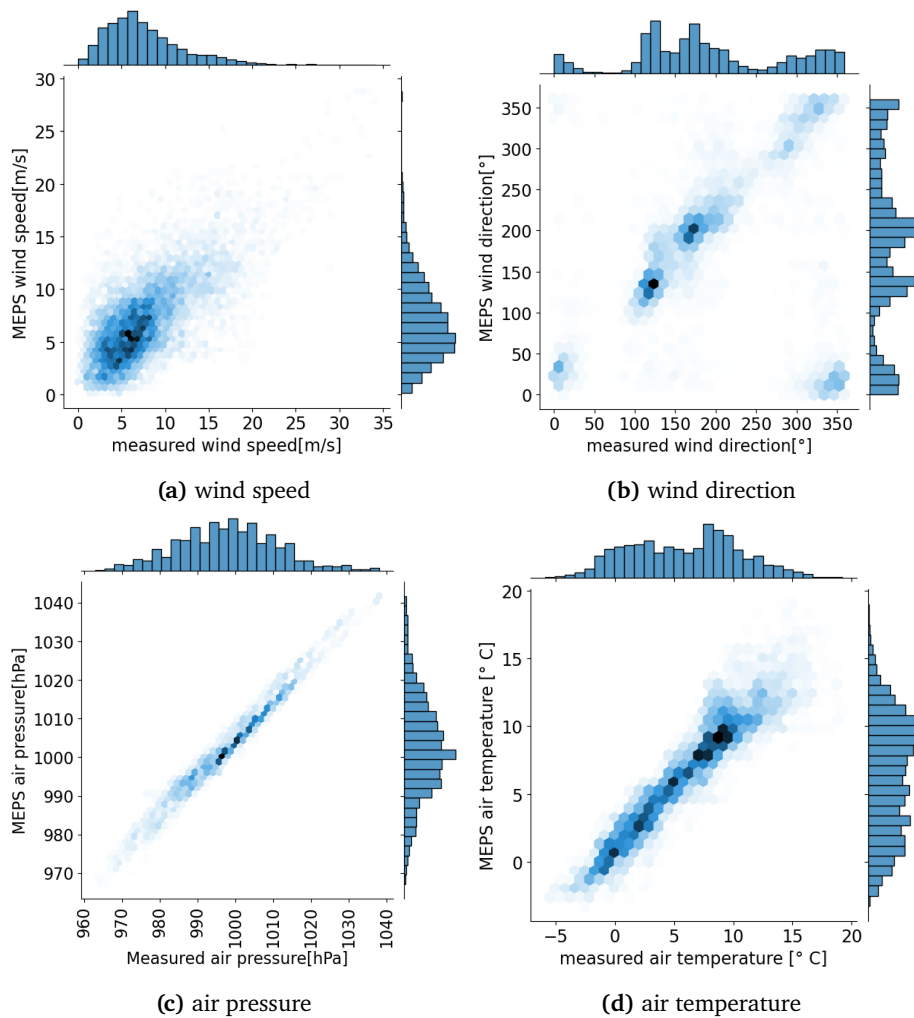
Figure 4.4: Two-dimensional histograms comparing measured and modelled weather parameters from MEPS

In this study the overall objective of the statistical and machine learning models is to map NWP input to park power production output. Therefore, this relationship is initially explored and interpreted trough two-dimensional histograms. Figure 4.5a and 4.5b illustrates how measured park power output is related to modelled wind speed and wind direction from MEPS, respectively. The non-linear relationships can be identified in these plots.

Figure 4.5a shows that measured park power and modelled wind speed have a relationship that resembles the form of a typical wind turbine power curve. Still, where a manufacturer power curve for a turbine follows a distinct line, the bins has a wide spread around the expected power curve. The most fre-

quently occurring bins are found for low wind speed and low power values. The absolute maximum frequency is found for low wind speed combined with zero production.

Figure 4.5b shows the less intuitive relationship between park power production and modelled wind direction from MEPS. Nevertheless, it highlights the dominant wind directions at the site: S-E and S-S-W, which agrees with the wind rose in figure 4.3a. The figure also shows that for low to moderate park power production the wind is most frequently blowing in the S-S-W sector. For maximum park power production, the wind is most frequently blowing in the S-E sector. This can also be found in the wind rose given in figure 4.3a, where the low to moderate wind speeds (blue tones) are most frequently occurring in the S-S-W sector, and the highest wind speed (yellow) in the S-E sector.

(a) wind speed   (b) wind direction
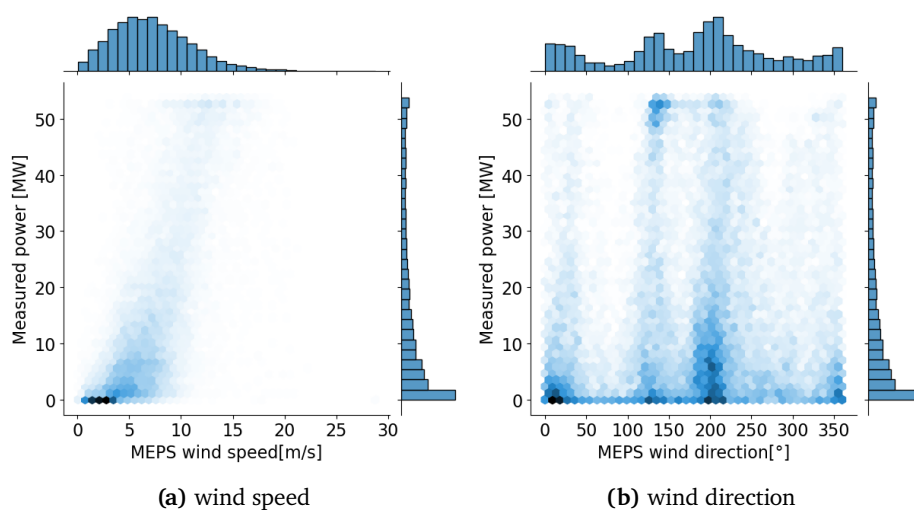
Figure 4.5: The relation between measured park power and modelled weather parameters from MEPS

4.1.2   The AROME Arctic dataset

Statistics and correlation

The analysis of the AA weather data provided in this section is based on the exploration period of 2017. The raw AA weather data are compared to on-site measurements for the same period. Table 4.3 provides descriptive statistics for each time series in the AA exploration dataset. The modelled air pressure and air temperature shows small deviations compared to measurements, while

there are notable deviations for the wind speed. The maximum modelled wind speed during the exploration period is eight m/s lower than for measured, which is a significant deviation. The average of modelled wind speed is more than one m/s lower than for measured. In addition, the modelled wind speed is less spread in relation to the average value. Note that the problematic temperature measurements are also her kept outside and the circular calculations for the cyclic wind speed are not provided.

Compared to the statistics of the MEPS dataset, the AA dataset has a larger deviation to the on-site measurements for most weather parameters. It should be noted that the time series of the AA dataset are taken at 10 m above ground level, while the MEPS dataset is taken at hub height. The measurements are also taken at hub height. This might explain the larger deviations observed for the AA dataset.

Table 4.3: Descriptive statistics of the time series in the AA dataset for 2017, compared to measurements for the same year

	count	mean	std	min	max
ws	8826.0	7.7	4.6	0.0	34.1
$\hat{ws}$	8826.0	6.4	3.4	0.1	26.4
wd	8826.0	-	-	0.0	359.0
$\hat{wd}$	8826.0	-	-	0.1	360.0
P	8826.0	997.8	13.0	963.0	1038.0
$\hat{P}$	8826.0	1006.4	13.1	973.7	1047.4
T	5952.0	5.9	4.7	-5.9	19.2
$\hat{T}$	5952.0	5.8	4.0	-3.4	15.9

Table 4.4 displays the linear correlation coefficient and the NRMSE between measured and modelled weather parameters. The cyclic wind direction is excluded. Again, air pressure shows significant strong positive correlation whit a correlation coefficient of 0.971. It also shows a low distance measure with NRMSE of 0.3% of installed capacity. Comparing the measures of the AA dataset with the MEPS dataset, the AA wind speed and AA temperature gives a somewhat stronger positive correlation and a slightly lower error than for the respective MEPS parameters. However, the opposite is observed for air pressure.

Table 4.4: The linear correlation coefficient and the NRMSE for pairs of measured and modelled weather parameters for the AA dataset

parameter	NRMSE	correlation coefficient
ws	0.102	0.713
P	0.003	0.971
T	0.201	0.719

(a) wind speed

(b) wind direction

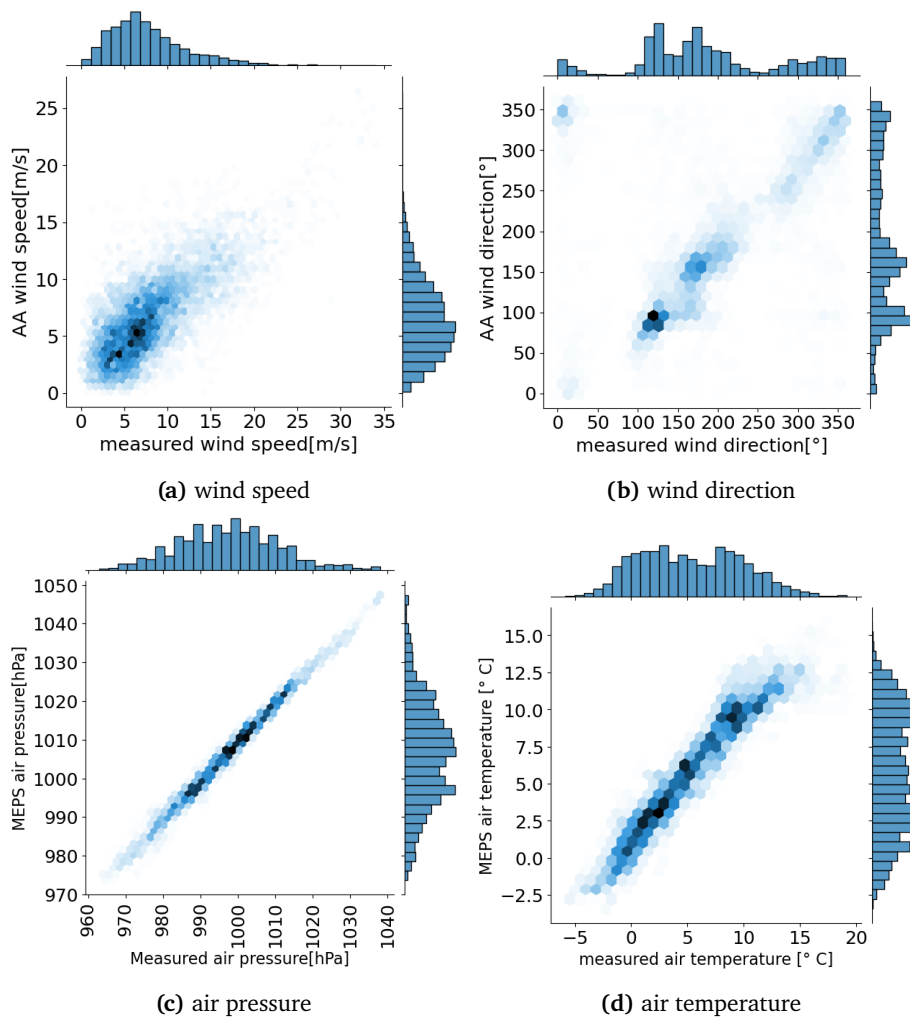**(c)** air pressure

(d) air temperature

Figure 4.6: Two-dimensional histograms comparing measured and modelled weather parameters from AA

Figure 4.6a, 4.6b, 4.6c and 4.6d visualizes the correlation for wind speed, wind direction, air pressure and air temperature, respectively. The correlation plots

for the AA weather parameters appears similar to the ones displayed for the MEPS weather parameters. In figure 4.6a the bins with the highest frequency in dark blue tones are stacked close to the diagonal and less frequent bins are spread around the diagonal for low to moderate wind speeds. The high wind speeds are less frequently occurring, and the bins are more widely spread. Note that the two distinct peaks or bins observed for the MEPS wind direction in figure 4.4b appears more blurred in figure 4.6b for the AA wind direction. It seems like the AA wind direction has a weaker correlation, compared to the MEPS wind direction. The most frequently occurring bins of air pressure form a narrow diagonal line, while for the temperature they are more widespread around the diagonal line and especially for high temperatures.

4.1.3   The market price dataset

Figure 4.7 shows the raw market price data for the entire 2020 as a function of time. Already, from the time series we observe volatile market prices, which also is expected for the future. The appearance of two peaks is notable. These representing single hours with high prices. Figure 4.8 zoom in at these extremes.

Figure 4.7: The spot and regulating prices for N04 for each hour in 2020

Figure 4.8: Extremes in spot and regulating prices for NO4 in 2020

4.2 Model parameters from training data

4.2.1 Correlation results for modified persistence model

Table 4.5 displays the model parameters of the modified persistence (NPE) model. The model parameters are based on power measurements for the training period 2017 to 2019. The average power production over the three years were 17.8 MWh. The correlation coefficient for the 14th lagged version of the power time series is found to be 0.352 and for each following lag the linear correlation decreases slowly to approximately 0.25.

Table 4.5: Model parameters for the modified persistence model

model parameter	value(s)
$\bar{p}$	17.815
a_s for s $\in [14, 38]$	[0.352, 0.346, 0.340, 0.334, 0.330, 0.325, 0.320, 0.316, 0.316, 0.314, 0.309, 0.302, 0.298 , 0.294, 0.287, 0.280, 0.276, 0.269, 0.262, 0.258, 0.254, 0.246, 0.243, 0.245]

4.2.2   Direction-specific power curve

The heat-map in figure 4.9 illustrates the direction-specific power curve (NWP-PPC). It is based on the median since it provided the best results, compared to the average. The NWP-PPC spans 12 wind sectors and 30 wind levels, with a resolution of 30 degrees and 1 m/s, respectively. In general, we observe that the power production increase as the wind speed increase beyond the cut-in wind speed for all wind sectors. However, there are remarkable differences between the power curves in the 12 wind sectors.

In the wind directions between S-E (135 degrees) and S-W (225 degrees) the park is expected to produce close to maximum (54 MWh) for moderate wind speeds. However, for the wind directions between approximately S-W and N-W (315 degrees) the park produces in general under 35MWh for moderate wind speeds. Significant differences are also observed between the wind sector spanning 90 to 120 degrees (E to ESE) and 120 to 150 degrees (ESE to SSE). In the former wind sector, the power production will never reach the maximum power production, and for wind speeds over 14 m/s there are too few data points. The latter wind sector, on the other hand, include one of the dominant wind directions (S-E), thus it carries many data points and only one grid cell miss data. It shows a power curve that gradually increase until it reaches approximately 12 m/s and then has a relatively constant power production before it decreases again at around 18 m/s.
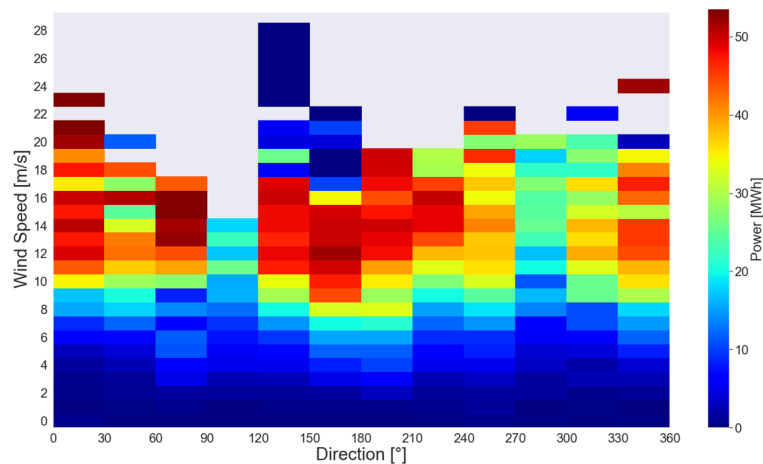

Figure 4.9: Direction-specific power curve based on rolling median of the MEPS dataset. White indicate missing values.

4.2.3 Tuned hyper-parameters for ML models

Tuning the machine leaning models on each dataset as described in chapter 3.3, result in the hyper-parameter values provided in table 4.6. The hyper-parameter values are provided for each machine learning model specifying the dataset and the combination of weather parameters. Using wind speed and wind direction as input features for the MEPS and AA dataset results in a RT model with a max-depth of 7 and 6, respectively. The RF model is only investigated for the MEPS dataset with wind speed and wind direction as input features. The resulting values for the hyper-parameters are a max-depth of 7, a min-samples-leaf of 50 and numb-of-trees of 100. The tuned MLP models with wind speed and wind direction as input features, thus two input neurons, have two to four hidden layers with 40 to 60 neurons in each. The tuned MLP models with three input neurons have three to four hidden layers with approximately 70 to 90 neurons in each.

Table 4.6: Tunes hyper-parameters for the machine learning models

Model	hyper-parameter	MEPS ws wd	AA ws wd	MEPS ws wd T	MEPS ws wd P
RT	max-depth	7	6		
RF	numb-of-trees	100			
	max-depth	7			
	min-samples-leaf	50			
MLP	# input neurons	2	2	3	3
	n-neurons	42	57	91	73
	n-hidden	2	4	4	3
	learning-rate	10e-3	10e-3	10e-2	10e-3
	# output neurons	1	1	1	1
	Hidden activation	ReLu	ReLu	ReLu	ReLu
	Loss function	MSE	MSE	MSE	MSE

4.3   Application of wind power forecasting models

Eight wind power forecasting (WPF) models are used as basis for eight day-ahead trading strategies. The WPF models includes three naive forecasting models, four point-forecasting models and one operational power forecast. These are applied for Fakken wind power park over the test period of 2020.

4.3.1   Case-study: High wind speed event

A high wind speed event is investigated in this section. It exemplifies the behavior of the WPF models, and how the behavior coincides with the input. The example period range over three days from December 25th to December 27th, 2020. The WPF models are applied individually on the MEPS and AA dataset. Both datasets include wind speed and wind direction as input features.

Figure 4.10a displays modelled and measured wind speeds for these three days. Note that measurements of wind direction are not available for this period. The first day shows varying wind conditions. The next day shows steady strong winds. The strong winds prevail until the appearance of a wind ramp i.e., large and fast wind speed variations, at the end of the example period. During the measured high winds both NWP models predict wind speeds that lays five to ten m/s lower than the measurements. During the wind ramp of very high winds the AA model predict wind speeds to be three to five m/s higher than the MEPS model.

Figure 4.10b shows the park power predictions of the naive forecasting models: PE, AVG and NPE. These are based on on-site park power measurements available before the gate closure of the day-ahead market. The PE model assume the production measured at 11:00 one day in advance will persist for the following day. This appears as a constant production value for each day. The average over the training period (AVG) is displayed as a horizontal line. Thus, the same power production is assumed for each hour in 2020. The NPE model, shows better performance compared to the individual models. In practice, the NPE model uses a weighting between the AVG model and the PE model. The prediction of the first hour the next day is 35% based on the PE model and 65% on AVG model, as found in table 4.5. The plotted NPE forecast lay closer to the average since the lagged version of the power time series has less and less impact on the prediction. For the following 24 hours the 35% decrease to circa 25%, explaining the slope in the plot.

Figure 4.10c shows the park power predictions of the point-forecasting models: NWP-PPC, RT, RF and MLP. These are based on the MEPS dataset including wind speed and wind direction as input features. Comparing the power predictions with the corresponding input wind speed, it is evident that they coincide. The cubic relationship between wind speed and power is also clearly illustrated. Small temporal variations in the wind speed forecast, result in large variations in the power forecast.

During the wind ramp at the end of the case period the observed wind speed exceeds the cut-out wind speed of the turbines. According to the power curve of the turbines (figure 2.1), it is expected zero production of the park. The MEPS model fails to predict this ramp event. Instead, it predicts a wind speed around 20 m/s. As a result, the WPF models predicts that the park will produce about 30 MWh, while in reality it did not produced. For the WPF models to behave correctly in such cases, it requires an NWP model that can resolve the high wind speed event with a certain accuracy.

It is observed that the tree-based models are sensitive to variations in the input data. This is illustrated in figure 4.10c between December 26th 12:00 and December 27th 06:00. It is especially seen for the RT model, and to some extent for the RF model. The RT model shows three local minima in a row within approximately eighteen hours, while the wind speed barely changes. The power predictions drop 10MWh from one hour to the next. The next hour it increases 10MWh back again. The remarkable pairs of drops and increases can be coupled to very small variations in the wind speed. The explanation is that the RT model associate the changed wind speed with a 10MWh lower power production.

Figure 4.10d shows the park power predictions of the RT and MLP model, based on the AA dataset. This dataset also includes wind speed and wind direction as input features. As for the MEPS dataset, the RT model is sensitive. It acts as a contrast to the smooth MLP model. What is truly notable, is the behavior of the models during the wind ramp. Previously we saw that the AA model predicts a higher wind speed during this wind ramp, compared to the MEPS model. The small increase in the wind speed forecast contributes to reduce the WPF by 50%. This makes the WPF models more accurate on the AA dataset, compared to the MEPS dataset during the wind ramp.
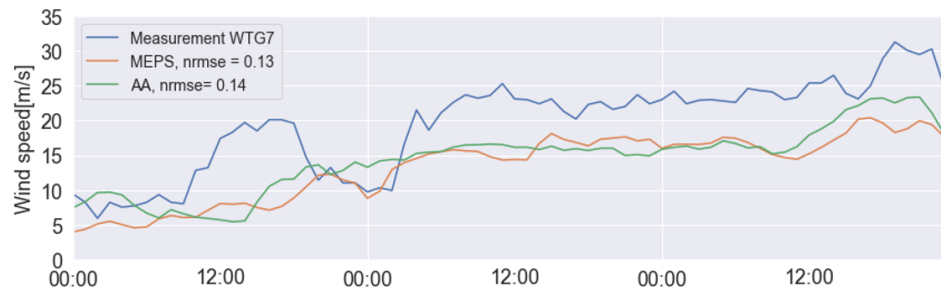
(a) Modelled vs measured wind speed
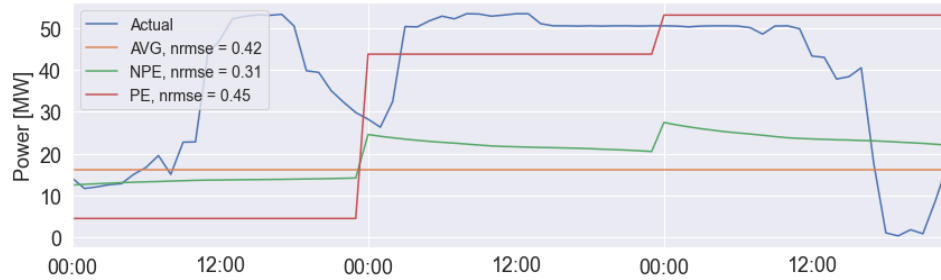
(b) naive power predictions
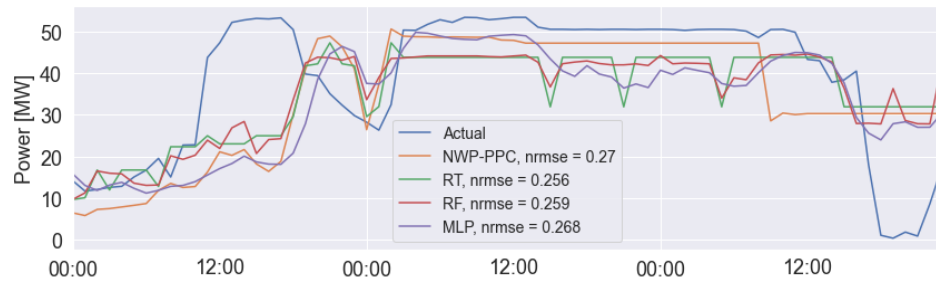
(c) point-forecasts of park power production based on the MEPS dataset including wind speed and wind direction as features

(d) point-forecasts of park power production based on the AA dataset including wind speed and wind direction as features
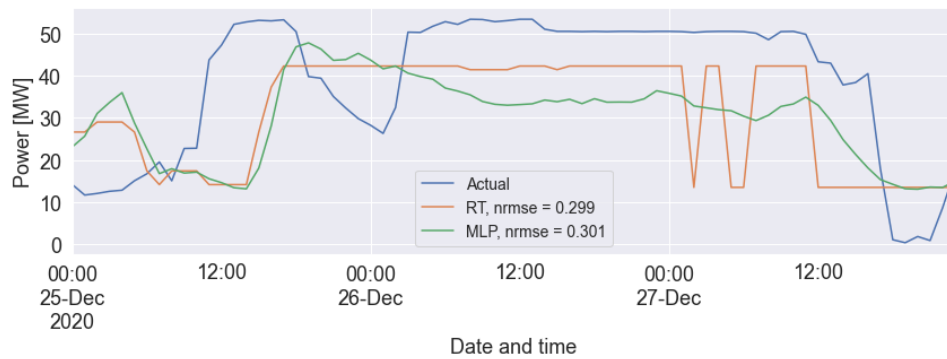
Figure 4.10: Case-study high wind event December 25th to 27th 2020

4.3.2   Statistical performance evaluation

The performance of the eight WPF models are evaluated using the statistical performance measures given in chapter 2.3.5. Table 4.7 summarize the statistical performance measures for the test period of the MEPS dataset. Wind speed and wind direction is used as input features.

Among the naive forecast methodologies, the well-known persistence (PE) model shows the best performance. The NRMSE is 37% and NMAE is 27% of installed capacity. This is set as a reference for the point-forecasting models. In contrast, the AVG model, which is the average power production over the training period, shows the worst performance. Despite the recommendation of using the modified persistence (NPE) model for day ahead WPF forecasting, it shows a worse performance than the standard PE model [Monteiro et al., 2009]. In practice, the NPE model uses a weighting between the AVG model and the PE model. When coupling this with the high NRMSE of the AVG model it is evident that the NPE provides a higher NRMSE than the PE model.

None of the naive forecasting models can recreate the variation of the true power prediction as reflected through the EVS. For instance, the AVG model is a constant value for all time steps and has no variations at all. Therefore, it has a EVS of zero. The AVG model also tends to largely overestimate the power production of the park. The PE model on the other side has a very low normalized bias (NBIAS).

Analyzing the WPF models built as single-step models, the RT model shows the best performance with a NRMSE and NMAE of 21,7% and 16,4%, respectively. Setting the PE model as reference, the NRMSE of the RT model represents 41% improvements. The direction-specific park power curve (NWP-PPC) model shows the worst performance among the point-forecasting models. Nevertheless, the improvement parameter setting the PE model as reference is 35% for the NWP-PPC model. As reflected through the EVS the point-forecasts better represents the variation of the true power prediction compared to the PE model. The NBIAS indicates that all models except the ISHK model, tends to overestimate the power production slightly.

Analyzing the distance measures of the MLP model built as a multi-step model, and comparing the results across all methodologies, reveals that the MLP model has the best performance. The NRMSE is 21,4% and the NMAE is 16,1% of the installed wind power capacity. Taking the example of the NRMSE and setting the PE model as reference, this corresponds to an improvement of 42%. The MLP model also shows the best potential for recreating the variation of the true power prediction, with an EVS of 0.55.

Setting the operational power forecast i.e., the ISHK model, as reference the MLP is the only model that shows a better performance on both the NRMSE and the NMAE. However, looking only at the NRMSE, which penalize large errors, the RT and RF model also performs better than the ISHK model.

Table 4.7: Statistical performance measures on the MEPS dataset using wind speed and wind direction as input features

Model	NRMSE	NBIAS	NMAE	EVS
ISHK	0.223	0.003	0.162	0.506
AVG	0.747	-0.677	0.677	0.000
NPE	0.406	-0.272	0.355	0.102
PE	0.370	-0.010	0.272	-0.360
NWP-PPC	0.240	-0.033	0.184	0.438
RT	0.217	-0.033	0.164	0.542
RF	0.220	-0.054	0.166	0.548
MLP	0.214	-0.024	0.161	0.550

4.3.3   Economic performance evaluation

The effectiveness of the eight day-ahead trading strategies are evaluated using the economic performance measures given in chapter 2.2.4. Table 4.8 displays the economic performance measures over the test period of the MEPS dataset. Wind speed and wind direction is used as input features. The NRMSE is included for a comparison between statistical and economic measures.

The revenue in the case of precise information at the time the day-ahead bid is submitted to Nord Pool, r_s^{PI}, is found to be 16.23 MNOK. However, due to the stochastic nature of the wind it is not possible to obtain precise information on the power production one day in advance, nether a perfect wind power forecast. The imbalance penalty term, r^{IMB}, reflects the price tag of the deviation between the submitted day-ahead bid and the actual power delivery in real time, when neglecting the intraday market. The total revenue of the power producer, r, is a result of subtracting the imbalance penalty term from the revenue in case of precise information. Note that a negative imbalance term indicates a revenue for the power producer.

Evaluating the effectiveness of the trading strategies based on economic measures reveals that the basic persistence (PE) model provides the highest total revenue for the power producer, despite the high NRMSE. The total revenue of the power producer is 16.42 MNOK for the PE trading strategy, which is a result of the imbalance revenue of 200 kNOK and the revenue in case of precise information at the time the forecast is provided. Note that the displayed imbalance revenues are results of both imbalance revenues and costs through the entire 2020. The NRMSE for the PE model is 37% of installed capacity. The high NRMSE suggest that the naive persistence model is not used as an trading strategy at the day-ahead market. The potential of large imbalance revenues can also be seen as a risk of large imbalance costs, which in the end heavily depends on the balancing price.

Excluding the naive forecasting models, the ISHK model shows to be the most effective trading strategy with an imbalance revenue of 30 kNOK. This result in a total revenue for the power producer of 16.25 MNOK for the entire 2020. This total revenue is also the one that is closes to the revenue in case of precise information.

The imbalance cost of the remaining point forecast models lays within the range 150 kNOK to 250 kNOK for the entire 2020. The corresponding performance ratio (f) lay within the range 98.5% to 99%. These measures indicate low annual imbalance costs and effective trading strategies. It is also observed that the NRMSE of the point forecasting models lays within the range 21% to 24% of installed capacity. Among the point-forecasting models, the NWP-PPC shows the lowest imbalance cost. In other words, the highest total revenue for the power producer. The total revenue is 16.06MNOK, which is 17 kNOK lower than the revenue in case of precise information, due to the imbalance cost. The MLP model, which showed the best statistical performance, displays a total revenue slightly lower than for the NWP-PPC model. The RT model, on the other hand, shows the highest imbalance cost. Thus, the lowest total revenue for the power producer of 15.98 MNOK.

Table 4.8: economic performance measures on the MEPS dataset using wind speed and wind direction as input features

Model	NRMSE	r^{PI} [MNOK]	r^{IMB} [MNOK]	r [MNOK]	f
ISHK	0.223	16.23	-0.03	16.25	1.002
AVG	0.747	16.23	0.55	15.68	0.966
NPE	0.406	16.23	0.40	15.83	0.975
PE	0.370	16.23	-0.20	16.42	1.012
NWP-PPC	0.240	16.23	0.17	16.06	0.990
RT	0.217	16.23	0.24	15.98	0.985
RF	0.220	16.23	0.23	15.99	0.986
MLP	0.214	16.22	0.18	16.04	0.989

4.3.4 Performance measures AROME Arctic

Table 4.9 shows the statistical and economic performance measures for the RT and MLP model applied on the AA dataset. The dataset includes wind speed and wind direction as input features. Comparing the statistical performance measures with the results for the MEPS dataset, we generally observe that the models perform worse on the AA dataset. The NRMSE of the RT and MLP model are 24.2% and 22.2%, respectively. Doing the same for the economic performance measures, we observe that the imbalance cost is smaller on the AA dataset, compared to the MEPS dataset. The imbalance cost of the RT and MLP model are 130 kNOK og 70 kNOK, respectively.

Table 4.9: Statistic and economic performance measures on the AA dataset using wind speed and wind direction as input features

Model	NRMSE	NBIAS	NMAE	EVS	r^{PI} [MNOK]	r^{IMB} [MNOK]	r [MNOK]	f
RT	0.242	-0.095	0.183	0.508	16.16	0.13	16.03	0.992
MLP	0.222	-0.062	0.174	0.548	16.16	0.07	16.09	0.996

4.3.5 Enhancement of Predictive Accuracy

Among the investigated models the MLP model shows the highest predictive accuracy. Table 4.10 displays the statistical and economic performance measures of the MLP model with different input features. This is an attempt to further increase the accuracy of the MLP model by selecting the best input features for the model. Using the wind data at component form shows lower NRMSE and NMAE compared to the wind data at vector form. Adding air temperature as the third input feature do not enhance the accuracy compared to using only the wind speed and direction. Including air pressure, on the other hand, increase the accuracy somewhat. In a statistical perspective, it has the best overall performance with a NRMSE of 20.9%. Considering the economic measures, the imbalance cost is reduced by more than 50% by including the temperature as the third input feature, in addition to the wind speed and wind direction. A reduction of the imbalance cost is also seen when including air pressure as an input feature.

Table 4.10: Feature selection for the MEPS dataset based on statistical and economic performance measures

features	NRMSE	NBIAS	NMAE	EVS	r^{PI} [MNOK]	r^{IMB} [MNOK]	r [MNOK]	f
$\hat{U}, \hat{V}$	0.216	0.004	0.165	0.539	16.22	0.27	15.96	0.983
$\hat{ws}, \hat{wd}$	0.214	-0.024	0.161	0.550	16.22	0.18	16.04	0.989
$\hat{ws}, \hat{wd}, \hat{T}$	0.221	-0.073	0.175	0.568	16.22	0.06	16.16	0.996
$\hat{ws}, \hat{wd}, \hat{P}$	0.209	-0.005	0.159	0.568	16.22	0.10	16.12	0.994

5

Discussion

5.1 NWP models and limitations

The input data of the WPF models investigated is limited to four years of data ranging from 2017 to 2020. Note that one year of data was initially tested for the WPF models. A higher accuracy was observed by increasing the amount of data to four years. The data is also limited to two NWP models, the MEPS model and the AROME Artic(AA) model. Both are provided by MET Norway and has a horizontal resolution of 2.5 km. The raw NWP data is down-scaled to the location of Fakken wind power park. For the AA and MEPS dataset, the statistical down-scaling approach differs, as well as the altitude at which the weather is extracted. For the MEPS dataset the weather parameters are extracted at hub height and linear interpolation is used for the statistical down-scaling. For the AA dataset, on the other hand, the weather parameters are extracted from the nearest grid cell and are taken at 10 meters above ground level.

Applying the RT and MLP model on the AA dataset results in a lower accuracy, compared to the MEPS dataset including the same input features. The NRMSE for the MLP model applied on the AA dataset is 22.2%, while for the MEPS data set it is 21.4%. The results may suggest that the statistical down-scaling approach applied is decisive for the final predictive accuracy, as described by Hanifi et al. [2020]. The altitude at which the weather parameters are extracted also appears as a determining factor for the predictive accuracy, as stated by Birkelund et al. [2018]. The resolution of a NWP models is another decisive factor for the accuracy [Birkelund et al., 2018]. By increasing the resolution

it is expected an higher accuracy for the WPF models on both datasets. NWP models in general, requires large resources in form of computational power and power consumption. A drawback of increasing the resolution is that even more computational power and power consumption is needs. In addition, if a higher resolution is not provided operationally, the power trader may be imposed to simulate the a NWP models themselves. In an environmental perspective, utilizing the operational weather forecast for day-ahead trading are more climate and environmentally friendly. In addition, the operational weather services have a specified competence, which the power traders take advantage of.

5.2   Single-step and multi-step models

A single-step model differs from a multi-step model regarding whether the time aspect of the data, or the temporal dependencies in the data, is considered by the model or not. The time aspect is not considered by a single-step model, as in a multi-step model. For the latter it is therefore important that the temporal order of the observations are kept or reserved.

A potential source of error is how the two datasets are prepared to match the framework of day-ahead trading. Daily time series are merged to form one coherent time series for the four years. This preparation works fine for single-step models where one input value for each feature is used to provide the representative power prediction. However, using the multi-step model with the same datasets, a source of error is introduced. At midnight each day there is a "break" between datapoints extracted from two forecasts provided at different days. The first datapoint has a prediction horizon of 36 hours while the second has a prediction horizon of 18 hours. In other words, the multi-step model is trained on abnormal transitions, which in turn might affect the test behaviour of the model. One possible solution is to use complete 06:00 forecasts with a 67-hour lead time as input to the WPF models. Power predictions are then provided for 67 hours at a time and is similar to what Birkelund et al. [2018] did in his study. After the predictions are provided looping over each forecast, the day-ahead power values can be extracted and merged to form a coherent time series. In this way the training on abnormal transitions would be avoided.

There are many ways of handling missing values in a dataset. Which strategy is chosen may introduce a certain source of error. In this study the missing values are filled with the last available value. In NWP data complete forecasts are typically missing. In the perspective of how data are merged to one coherent time series, 24 hours of data are (at least) missing at a time. In other

words, there are introduced a high number of equal samples based on the last available value, which probably is not the most accurate strategy. The result is many equal and temporally stacked values. For time-steps where missing values are handled the accuracy of the forecast is poorer compared to the rest of the dataset. One optional strategy, that might be more appropriate for the day-ahead trading purpose, is to use the 00:00 forecast if the 06:00 forecast is missing. This is possible since the operational forecasts used in this study has a lead time of 67 hours. For the 00:00 forecast the day-ahead values are found for the time-steps between 24 and 48 hours. Another optional strategy, used by Sæther [2021], is to use the slope of the previous forecast, that is the last 67 time steps, to impute the missing values.

Another data pre-processing step that might introduce other types of errors is the detection and handling of anomalies. In this study the bottom stacked anomalies are removed from the training. That is when the wind speed is between the cut-inn and cut-out but the park power output is zero. The purpose is to avoid training at invalid samples and force the predictions closer to the maximum power production of the park. This works fine for single-step models. However, since the temporal order is not preserved it might be problematic for multi-step models. When one observation is removed the previous and next observation are set together. The sliding-window of a multi-step model will then interpret the observation aligned in time, while in reality there is one or more observations in-between.

Another note regarding the physical explanatory model for anomalies is that the interval set for the wind speeds might be too strict. Using the cut-inn and cut-out as the interval for the wind speed might be a too strict rule. From the descriptive statistics displayed in the results, it is observed a significantly lower maximum wind speed for the modelled data compared to the on-site measurements. This suggests that the handling of anomalies remove training data for high wind speed events.

Despite the sources of errors related to the pre-processing steps which may affect the multi-step model, the MLP model shows the overall best results considering the statistical measures. The NRMSE of the best performing MLP model is 20.9%.

5.3   Assumptions and reality

The intraday market is neglected, forcing the day-ahead deviation to be corrected in real time at the balancing market. Neglecting intraday trading of wind

power is a reasonable assumption because it is not often used in the current practice for Fakken wind power park. Eventually the intraday trading may be more attractive due to the volatile prices of the power market.Imagine that the the operation hour is closing up and a new updated NWP forecasts predicts significantly changes in the wind speed for the operation hour. Then the wind power trader should consider correcting the bid at the intraday market with known prices, instead of letting it go to the balancing market. However, manual wind power trading at the intraday market is too time consuming. There is needed automatic prediction systems for this purpose. Further research may investigate automatic prediction systems for intraday-trading of Fakken wind power park. The time horizon of intraday trading highly differs from the day-ahead treading, making this a quite different research. Nevertheless, the overall objective of maximizing the profit of the power trader is still present.

Suppose a deviation between power contracted at the day-ahead market and power delivered in real time. The imbalance cost is just one of the risks a wind power producer is exposed to in the balancing market. The imbalances are also charged with imbalance fees, added on top of the imbalance cost. These fees are currently at a high level, which gives the producers an incentive to reduce the power imbalance. The imbalance fees, together with incentives are neglected for the calculation of the total revenue of the wind power trader. Further research may investigate how much these fees amounts to the total revenue of the wind power producer. Assuming that the market prices are known in advance is a strong assumption. In reality this assumption is relaxed [Mazzi and Pinson, 2017]. In this study the market prices are only used to evaluate the effectiveness of the trading strategies after the operation hour. Thus, they are not used as input to decision-making before the operation hour takes place.

5.4   Performance measures

Performance evaluation in wind power forecasting are commonly based on statistical measures that indicates the deviation between predicted and measured wind power [Hanifi et al., 2020]. However, error rates are not the only criteria that affects our decisions in model selection. Other criteria are for instance the interpretability and the run-time of the models [Alpaydin, 2014]. Among the point-forecasting models investigated, the NWP-PPC model and RT model are highly interpretable. The former trough the site-specific park power curve and the RT model through sets of IF-THEN statements. At the other end of the interpretability scale, we find the MLP model. For large datasets it is difficult to interpret due to many neurons and fully connected layers. However, using

Keras sequential API can make it easier to explain the architecture design and in turn interpret the model behavior.

The application of wind power forecasting models for day-ahead wind power trading based on NWP data, limits the accepted runtime of the models. The gate of the day-ahead market closes at 12:00 each day, and the submission of the bid is assumed to be at 11:00. The operational 06:00 forecasts are made available for the power trader some hours later depending on the simulation time of the NWP model. Typically, the 06:00 forecast is made available before 09:00. In other words, a run-time of circa two hours is reasonable.

In a market trading perspective where the WPF models are to be used as trading strategies, the economic value is a deciding criteria. Revenues and imbalance costs can be used to determine which WPF model that provide an optimal trading strategy for the wind power park. In the liberalized power market wind power traders participates under the same rules as the traders of conventional power plants. Neglecting the intraday market, the deviation between day-ahead contracts and real time delivery of wind power are corrected at the balancing market and priced according to the implemented pricing system. In other words, the imbalance between predicted and measured wind power has a price tag. The new pricing system at the Nordic power market makes it possible for the power trader to receive revenues or costs related to their imbalance in real time. The economic value of different WPF models highly depends on the pricing system and the relative market prices. Excluding the naive forecasting models, the selection of WPF model seems to have little impact on the economic value, compared to the impact of the pricing system and the market prices. The point-forecasting models shows similar statistical performance measures, and similar economic value in form of total revenue for the power producer.

Considering the statistical performance measures, the point-forecasting models shows much better performance compared to the naive forecasting models. The NRMSE of the naive forecasting models lay around 40% to 70% of installed capacity, while the point-forecasting model has a NRMSE around 20 %. Considering the economic measures, the naive forecasting models shows the highest imbalance revenues and costs, compared to the point-forecasting models. The imbalance cost of the AVG and NPE model lay in the range 40 kNOK to 50 kNOK, while the imbalance revenue of the PE model is 20 kNOK. The point-forecasting models, on the other hand, show similar and more stable imbalance costs with slight variations. According to the performance ratio, f, the imbalance cost is small compared to the total revenue. If the power trader uses one of the naive forecasting models as trading strategy, he or she incur a higher risk of large imbalances. This implies a high reward potential but also cost potential. However, choosing one of the point-forecasting models the risk

of large imbalances are highly reduced, but then also the revenue and cost potential. The wind power traders are restricted to submit their best point-forecast of the wind power park and not to speculate in market prices. This brings us back to the importance of the statistical performance measures. It also suggest that the total revenue for the power producer is as close as possible to the revenue in case of precise information at the time when the forecast is provided. Analyzing the absolute value of the imbalance penalty term, the ISHK model has the smallest imbalance term, which gives the performance ratio closest to one.

5.5   Validity of results

The paper of Giebel et al. [2011] states that typical numbers of NRMSE for a 36-hour prediction horizon lays within 10-15% of the installed wind power capacity. The accuracy of the best performing model in this study is not within the proposed range. The complex terrain surrounding Fakken wind power park might be one explanation. A complex terrain may make it more challenging for the NWP model to make accurate weather forecasts. Alessandrini and Sperati [2017] states that the power prediction accuracy highly depends on the complexity of the terrain. The RIX is set in context with the park power prediction accuracy for a 12-hour prediction horizon where NWP data is the input. The paper show that complex terrain gives rise to a much higher NMAE compared to flat terrain.

The models investigated by Birkelund et al. [2018] and the respective results are comparable with the results of this study. The design of the NWP-PPC investigated in this study is inspired by the direction specific park power curve provided by Weir [2014], which also is the basis for the PNWP model investigated by Birkelund et al. [2018]. From a plot in this paper it is observed an NRMSE around 25% of nominal power, which is in the same area (24%) as the NWP-PPC of this study. The Analog Ensembles (AnEn) model resembles the RF model which is an ensemble of RT models. These models are all based on the assumption that similar inputs have similar outputs, and may have similar behaviour.

6

Conclusion

In this study we have answered how well a set of wind power forecasting (WPF) models works as day-ahead trading strategies for Fakken wind power park. Applying the WPF models on MEPS data, using wind speed and wind direction as input features, and evaluating the statistical performance measures over a test period (Chapter 4.3.2), reveals that the MLP model built as an multi-step model provides the highest accuracy. The NRMSE is 21.4% of installed capacity. Setting the basic persistence (PE) model as reference the MLP model shows an improvement of 42%. Similarly, setting the ISHK model i.e., the operational power forecast, as reference the MLP model shows an improvement of 4.0%.

Further enhancement of the predictive accuracy of the MLP model is attained by adding air pressure as the third input feature to the model. The resulting NRMSE is 20.9% of installed capacity, which is the overall best statistical performance of this study. This result corresponds to a 6.3% improvement compared to the ISHK model, which verifies that the MLP model can compete with the operational power forecast. From these findings it is evident that model selection is not the only way to enhance the predictive accuracy. Selection of input parameters also plays a big part.

Evaluating the effectiveness of the trading strategies through economical measures over the same test period (Chapter 4.3.3), reveals that the naive persistence (PE) model provides the highest total annual revenue for the power producer despite the high NRMSE. The total revenue is 16.42 MNOK, where the imbalance revenue accounts for 200 kNOK. The NRMSE is 37% of installed

capacity, which suggests that the PE model is not used as an day-ahead trading strategy. The potential of large imbalance revenues also contributes a risk of large imbalance costs, which in a single price system heavily depends on the balancing price. The risk of large imbalances are highly reduced choosing one of the point-forecasting models. Choosing one of the point-forecasting models the risk of large imbalances are highly reduced, but then also the revenue and cost potential.

Excluding the PE model, the ISHK model is the most effective trading strategy. It provides a total annual revenue for the power producer of 16.25 MNOK, where the imbalance revenue accounts for 30 kNOK. This is also the total revenue which is closest to the revenue in case of precise information. The latter is suggested as an important criterion for an effective trading strategy in this study. In a single price system, the aim should be to maximize the total revenue to a certain level. The level is set by the revenue in case of precise information. This leads us back to the importance of the statistical measures, which examines the error in the wind power forecasts.

Inspecting the behaviour of the WPF models for a high wind speed event (Chapter 4.3.1), reveals that all models are sensitive to wind speed variations close to cut-out. Focusing on a wind ramp with observed wind speeds higher than the cut-out, a notable feature is how a small increase in the wind speed forecast, reduces the power production forecast by 50%. While this single high wind event limits the generalizability of the findings, it provides insight into the important role of the inputs. When the NWP models are unable to resolve the high winds, nether the WPF models succeed in predicting high production. This highlights the importance of NWP-data pre-processing.

6.1   Future research

Based on the conclusions of this study, the suggestions for future research is as follows. With the aim of achieving a higher accuracy for the MLP model, a natural first step could be to obtain an optimal combination of the two NWP datasets. Expanding the study to include weather data from an additional operational NWP model outside MET Norway could also be interesting. In that case consider using a common statistical down-scaling approach for the different NWP models. The results of this study suggests linear interpolation for this purpose. Beyond the different weather parameters investigated in this study, it could be interesting to include the turbulent kinetic energy as an input feature to the MLP model, as proposed by Optis and Perr-Sauer [2019]. When

applying WPF models on NWP data, one of the major challenges is the errors in the NWP data [Alessandrini and Sperati, 2017]. A potential future path could be to focus at data pre-processing steps to reduce these errors. Kalman filter is suggested in the literature by Hanifi et al. [2020] and Jung and Broadwater [2014] to reduce systematic errors in the NWP data. Another path is to focus on physical explanatory models for anomalies in the data, which is especially important in wind power curve modelling [Wang et al., 2019]. In the economic perspective, a step for further research can be to investigate how much the imbalance fees, on top of the imbalance cost, affects the total revenue for the wind producer. Another economic direction could be to study the intraday trading stage, which is neglected in this study, and work on automatic prediction systems.

6.2   Concluding remarks

In this study we found that the MLP model and the ISHK model shows the overall best results taking both statistical and economical considerations, respectively. Nevertheless, the MLP model only provides a small decrease in NRMSE compared to the ISHK model. The applied pre-processing steps affects the temporal order of the datasets, and might have an impact on the MLP results. NWP data is used as input for both models. Improvements through data pre-processing steps may provide better results for both models. Note that the ISHK model already use weighting between different weather service providers, as suggested in the further research section.

Bibliography

ACER. Acer's preliminary assessment of europe's high energy prices and the current wholesale electricity market design. `https://extranet.acer.europa.eu//Official_documents/Acts_of_the_Agency/Publication/ACER's%20Preliminary%20Assessment%20of%20Europe's%20high%20energy%20prices%20and%20the%20current%20wholesale%20electricity%20market%20design.pdf`, 2021. Accessed: 2021-02-02.

S. Alessandrini and S. Sperati. Characterization of forecast errors and benchmarking of renewable energy forecasts. In *Renewable Energy Forecasting*, pages 235–256. Elsevier, 2017.

E. Alpaydin. *Introduction to Machine Learning Third edition*. Massachusetts Institute of Technology, Reading, Massachusetts, 2014.

B. Babar, L. T. Luppino, T. Boström, and S. N. Anfinsen. Random forest regression for improved mapping of solar irradiance at high latitudes. *Solar Energy*, 198:81–92, 2020.

N. Belyakov. *Sustainable Power Generation: Current Status, Future Challenges, and Perspectives*. Academic Press, 2019.

J. Benesty, J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer, 2009.

H. Birkelund, F. Arnesen, J. Hole, D. Spilde, S. Jelsness, F. H. Aulie, and I. E. Haukel. Langsiktig kraftmarkedsanalyse 2021 – 2040 forsterket klimapolitikk påvirker kraftprisene. `https://publikasjoner.nve.no/rapport/2021/rapport2021_29.pdf`, 2021. Accessed: 2022-04-22.

Y. Birkelund, S. Alessandrini, Ø. Byrkjedal, and L. D. Monache. Wind power prediction in complex terrain using analog ensembles. *Journal of Physics: Conference Series*, 1102:012008, oct 2018. doi: 10.1088/1742-6596/1102/1/012008.

J. Brownlee. Time series forecasting as supervised learning. *Machine Learning Mastery [online]*, 5, 2016. Accessed: 2021-05-30.

J. Brownlee. *Long short-term memory networks with python: develop sequence prediction models with deep learning*. Machine Learning Mastery, 2017.

C. Donald Ahrens and R. Henson. Meteorology today: an introduction to weather, climate and the environment. 2015.

O. F. Eikeland, F. D. Hovem, T. E. Olsen, M. Chiesa, and F. M. Bianchi. Probabilistic forecasts of wind power generation in regions with complex topography using deep learning methods: An arctic case. *arXiv preprint arXiv:2203.07080*, 2022.

A. A. Fossem. Short-term wind power prediction models in complex terrain based on statistical time series analysis. Master's thesis, UiT The Arctic University of Norway, 2019.

I.-L. Frogner, A. T. Singleton, M. Køltzow, and U. Andrae. Convection-permitting ensembles: Challenges related to their design and use. *Quarterly Journal of the Royal Meteorological Society*, 145(S1):90–106, 2019. doi: https://doi.org/10.1002/qj.3525. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3525`.

M. W. Gardner and S. Dorling. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15):2627–2636, 1998.

A. Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", 2019.

G. Giebel and G. Kariniotakis. Wind power forecasting—a review of the state of the art. *Renewable energy forecasting*, pages 59–109, 2017.

G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. The state of the art in short-term prediction of wind power a literature overview, 2nd edition. 01 2011. doi: 10.13140/RG.2.1.2581.4485.

S. Hanifi, X. Liu, Z. Lin, and S. Lotfian. A critical review of wind power forecasting methods—past, present and future. *Energies*, 13(15):3764, 2020.

S. E. Haupt, P. A. Jiménez, J. A. Lee, and B. Kosović. Principles of meteorology and numerical weather prediction. In *Renewable Energy Forecasting*, pages

3–28. Elsevier, 2017.

N. R. Hoot, L. J. LeBlanc, I. Jones, S. R. Levin, C. Zhou, C. S. Gadd, and D. Aronsky. Forecasting emergency department crowding: a discrete event simulation. *Annals of emergency medicine*, 52(2):116–125, 2008.

K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. ISSN 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2006.03.001. URL `https://www.sciencedirect.com/science/article/pii/S0169207006000239`.

IPCC. Summary for policymakers [h.-o. pörtner, d.c. roberts, e.s. poloczanska, k. mintenbeck, m. tignor, a. alegría, m. craig, s. langsdorf, s. löschke, v. möller, a. okem (eds.)]. Technical report, 2022a. In: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press.

IPCC. Summary for policymakers. Technical report, 2022b. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA. doi: 10.1017/9781009157926.001.

P. L. Jackson, G. Mayr, and S. Vosper. Dynamically-driven winds. In *Mountain weather research and forecasting*, pages 121–218. Springer, 2013.

M. Jacobsen. Short-term wind power prediction based on markov chain and numerical weather prediction models: A case study of fakken wind far. Master's thesis, UiT The Arctic University of Norway, 2014.

J. Jung and R. P. Broadwater. Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews*, 31: 762–777, 2014.

G. Kariniotakis. *Renewable energy forecasting: from models to applications*. Woodhead Publishing, 2017.

G. Kariniotakis, I. Martí, D. Casas, P. Pinson, T. S. Nielsen, H. Madsen, G. Giebel, J. Usaola, I. Sanchez, A. Palomares, et al. What performance can be expected by short-term wind power prediction models depending on site characteristics. In *CD-Rom proceedings of the European wind energy conference EWEC*, 2004.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

A. Kusiak, H. Zheng, and Z. Song. Wind farm power prediction: a data-mining approach. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 12(3):275–293, 2009.

T. Langset and H. Nielsen. Rme rapport – national report 2021. 2021.

H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. S. Nielsen. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering*, 29(6):475–489, 2005. doi: 10.1260/030952405776234599. URL https://doi.org/10.1260/030952405776234599.

P. Markowski and Y. Richardson. *Mesoscale meteorology in midlatitudes*, volume 2. John Wiley & Sons, 2011.

N. Mazzi and P. Pinson. 10 - wind power in electricity markets and the value of forecasting. In G. Kariniotakis, editor, *Renewable Energy Forecasting*, Woodhead Publishing Series in Energy, pages 259–278. Woodhead Publishing, 2017. ISBN 978-0-08-100504-0. doi: https://doi.org/10.1016/B978-0-08-100504-0.00010-X. URL https://www.sciencedirect.com/science/article/pii/B978008100504000010X.

C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, et al. Wind power forecasting: State-of-the-art 2009. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States), 2009.

J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno. *Integrating renewables in electricity markets: operational problems*, volume 205. Springer Science & Business Media, 2013.

P. Moriarty and D. Honnery. The limits of renewable energy. *AIMS Energy*, 9 (4):812–829, 2021.

M. Müller, Y. Batrak, J. Kristiansen, M. A. Køltzow, G. Noer, and A. Korosov. Characteristics of a convective-scale weather forecasting system for the european arctic. *Monthly Weather Review*, 145(12):4771–4787, 2017.

NBS. Nordic imbalance settlement handbook. `https://www.esett.com/app/uploads/2021/11/NBS-Handbook-v3.2.pdf`, 2021. Accessed: 2021-12-12.

T. Nielsen, H. Madsen, H. A. Nielsen, P. Pinson, G. Kariniotakis, N. Siebert, I. Marti, M. Lange, U. Focken, L. V. Bremen, et al. Short-term wind power forecasting using advanced statistical methods. In *The European Wind Energy Conference, EWEC 2006*, pages 9–pages, 2006.

T. S. Nielsen, A. Joensen, H. Madsen, L. Landberg, and G. Giebel. A new reference for wind power forecasting. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 1(1):29–34, 1998.

NordPool. Bidding areas. `https://www.nordpoolgroup.com/the-power-market/Bidding-areas/`, n.d. Accessed: 2021-02-01.

NVE. Wholesale market: Timeframes. `https://2021.nve.no/norwegian-energy-regulatory-authority/wholesale-market/wholesale-market-timeframes/`, 2021. Accessed: 2021-11-08.

M. Optis and J. Perr-Sauer. The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production. *Renewable and Sustainable Energy Reviews*, 112:27–41, 2019.

O. Ramadan, S. Omer, M. Jradi, H. Sabir, and S. Riffat. Analysis of compressed air energy storage for large-scale wind energy in Suez, Egypt. *International Journal of Low-Carbon Technologies*, 11(4):476–488, 11 2016.

H. Ramchoun, Y. Ghanou, M. Ettaouil, and M. A. Janati Idrissi. Multilayer perceptron: Architecture optimization and training. 2016.

F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

D. E. Rumelhart, R. Durbin, R. Golden, and Y. Chauvin. Backpropagation: The basic theory. *Backpropagation: Theory, architectures and applications*, pages 1–34, 1995.

B. B. Sæther. Wind power prediction with machine learning methods in complex terrain areas. Master's thesis, UiT The Arctic University of Norway, 2021.

M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, et al. A survey of forecast error measures. *World Applied Sciences Journal*, 24(24):171–176, 2013.

L. Söder and M. Amelin. Efficient operation and planning of power systems. 2011.

K. Solbakken, Y. Birkelund, and E. M. Samuelsen. Evaluation of surface wind using wrf in complex terrain: Atmospheric input data and grid spacing. *Environmental Modelling & Software*, 145:105182, 2021.

Statnett. Langsiktig markedsanalyse 2020-2050 - oppdatering våren 2021, 2021. Accessed: 2022-04-22.

J. T. Svane. Wind power prediction, p. unpublished, 2021. EOM-3010 Project thesis in Energy, Climate and Environment, The Arctic University of Norway.

E. Tan, S. S. Mentes, E. Unal, Y. Unal, B. Efe, B. Barutcu, B. Onol, H. S. Topcu, and S. Incecik. Short term wind energy resource prediction using wrf model for a location in western part of turkey. *Journal of Renewable and Sustainable Energy*, 13(1):013303, 2021.

TromsKraft. Fakken vindpark. `https://www.tromskraft.no/ produksjon/ kraftverk/fakken-vindpark`, n.d. Accessed: 2021-11-08.

Vestas. V90-3.0 mw manual, n.d.a. Accessed: 2021-12-13.

Vestas. Powerplus case study: Extended cut out. `http://nozebra.ipapercms. dk/Vestas/Communication/Productbrochure/ProductImprovements/eco- case-study/`, n.d.b. Accessed: 2022-01-31.

X. Wang, P. Guo, and X. Huang. A review of wind power forecasting models. *Energy Procedia*, 12:770–778, 2011. ISSN 1876-6102. doi: https: //doi.org/10.1016/j.egypro.2011.10.103. URL `https://www.sciencedirect. com/science/article/pii/S1876610211019291`. The Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE 2011.

Y. Wang, Q. Hu, L. Li, A. M. Foley, and D. Srinivasan. Approaches to wind power curve modeling: A review and discussion. *Renewable and Sustainable Energy Reviews*, 116:109422, 2019. ISSN 1364-0321. doi: https://doi.org/10.1016/j. rser.2019.109422. URL `https://www.sciencedirect.com/science/article/ pii/S1364032119306306`.

D. Weir. Vindkraft -produksjon i 2013. nve. `https://webfileservice.nve.no/ API/PublishedFiles/Download/201307790/1878182`, 2014. Accessed: 2021-11- 28.

T. Wizelius and J. Earnest. *Wind power plants and project development*. PHI Learning Pvt. Ltd., 2011.