Hybrid Wind Power Forecast

Combining Physical and Statistical Approaches to Better Predict Wind Power Production

Mathias Knudsen Sollie

Thesis submitted for the degree of
Master in Programming and network
30 credits

Department of Informatics
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

Spring 2018

Hybrid Wind Power Forecast

Combining Physical and Statistical Approaches to Better Predict Wind Power Production

Mathias Knudsen Sollie

Hybrid Wind Power Forecast

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Acknowledgements

First, I would like to thank my two supervisors, Yan Zhang and Arne Gravdahl of WindSim AS for making this thesis possible and for their guidance throughout the process.

In addition, I also wish to thank the team at WindSim AS for their technical assistance and free lunches, it was great working with you.

Thanks for the love and support from my closest family, the Knudstads, as well as the rest of my family. Forever grateful to the GOAT, Jan the IV. Shoutout to the boys.

Lastly, a shoutout to my boy Bjørn Fesche for cleaning this mess. We out here, we made it.

Contents

List of Figures

List of Tables

Abstract

Wind turbines are increasing in popularity as a power source around the world, and are among the cheapest sources of energy. They introduce problems related to the volatile nature of the wind, and the uncertainty of the timing and amount of energy production. Current methods generally use a physical or statistical approach for predicting power generated from wind turbines.

Our thesis explores combining these approaches for a hybrid approach, using computational fluid dynamics as the physical approach, and machine learning models for the statistical approach. Six different machine learning models have been tested, resulting in six different hybrid approaches. The data consisted of historical wind condition forecasts and the historical power production, for five different wind farms in Norway. The output was the predicted power production for the next 24 hours, given in hourly measures of megawatt hours. The hybrid approaches are tested against the standard physical and statistical approach, and their performance is measured and compared.

The results show that for four of the five wind farms, the hybrid approach outperformed both the statistical and physical approaches.

Summary

Wind turbines are increasing in popularity as a power source around the world, and are among the cheapest sources of energy. However, unlike traditional power sources, they introduce problems related to the volatile nature of the wind, and the uncertainty of the timing and amount of energy production. Statnett, the Norwegian power grid operator, states that renewable sources of energy, especially wind, are a problem when operating the grid and balancing its supply. There are methods for forecasting wind power production, but there are still much room for improvement.

This thesis looks at a combination of the two most common wind power production forecasting approaches, physical and statistical, and combine these into a hybrid approach. We will then test this approach for predicting the future power generated from wind turbine sources. Current methods generally use one of the two approaches, with varying results. There have been attempts at the hybrid approach, with some success but limited further exploration. Our suggested approach uses computational fluid dynamics as the physical approach, and machine learning models for the statistical approach. Six different machine learning models have been tested, resulting in six different hybrid approaches. The data for the thesis consisted of historical wind condition forecasts and the historical power production, for five different wind farms in Norway. The historical wind condition forecasts were supplied by Yr weather forecast service and the historical power production were supplied by the Norwegian Water Resources and Energy Directorate.

The hybrid solutions start with simulating the wind flowing from the geographical point of the wind condition forecast, to each wind turbine in the wind farm. This is accomplished using computational fluid dynamics, and provides a per-turbine wind speed and wind direction forecast. The average wind speed and average wind direction for the wind farm were calculated, based on the new per-turbine forecasts. The average wind speed and average wind direction, along with the time of the day, is the input for the machine learning models. The models tested were artificial neural network, k-nearest neighbors, decision tree, random forest regression, linear regression and support vector regression. The output was the predicted power production for the next 24 hours, given in hourly measures of megawatt hours.

The physical, statistical, and hybrid approaches are tested, and their performance is measured and compared. The results show that there are

many uncertainties in wind power prediction, but that for four of the five wind farms, the hybrid approach outperformed both the statistical and physical approaches. This indicates that in the future, the accuracy of the predictions may be increased, and the hybrid approach can be further explored.

Chapter 1

Introduction

The world is moving towards a greener future. Our old way of living has come to an end, and we need to change as a global society. A great challenge has long been where to get our power from, once we throw out the fossil fuel power production. No concrete answer or solution has been provided, but collectively we have already started on a path to greener energy.

Renewable energy sources, like wind, hydro and solar power, are slowly becoming the new standard for producing electrical power. That said, with new solutions come new problems. For renewable energy sources, these problems are mainly caused by their volatile nature. Especially for wind, this is a difficult problem. Wind turbines are used for generating electrical power, and the use of wind turbines as a source of electrical power has increased in the later years. The International Renewable Energy Agency reports a mean growth of installed wind capacity of 15% annually [26]. This makes wind one of the fastest growing sources of energy. Wind turbines are cheap to build and maintain, resulting in cheap energy for the consumers, and is expected to become the world's cheapest source of energy in the years to come [26].

The main problem introduced by wind power lies in the balancing of the power grid. Using a volatile energy source makes it hard to know when when power will be generated and how much power is provided to the grid. Statnett, responsible for the energy grid in Norway, reports renewable power sources as an issue in balancing the grid, and points especially towards wind farms. The volatility of wind disturbs the balance of the grid [50]. There are two primary approaches for forecasting wind power production. The physical approach simulates the wind using physical principles. The statistical approach uses historical data and statistical methods to make future predictions. These are two very different approaches, but both try to answer the same question of future wind power production.

In my thesis, I will look at how we can combine physical and statistical approaches for improved wind power forecasting. The expected outcomes are prediction methods which outperform the two basic approaches.

Chapter 2

Wind Power

In this chapter, we will explain what a wind farm is and how we use them to generate electrical power.

2.1 Wind

The temperature of the earth's atmosphere varies from place to place. Air pressure is determined by temperature, which results in varying pressures around the globe. Wind is air moving air. The reason air moves, is difference in pressure. Air will move from a place of high pressure to a place of low pressure [10], creating wind. Due to the constantly changing temperature and pressure, wind is quite volatile, changing direction and speed on a second's notice. It is not easy to predict the winds behavior in advance, and even today the meteorological institutes can be wrong. To keep track of the wind and explain its behavior locally, we need specialized methods.

There are many ways to describe the wind, but the two most important parameters are wind speed and wind direction. Weather forecasters use terms like "gentle breeze", "north-east" and "light air". A more accurate way to present wind speed is by common speed units, like knots or meters per second. This is more relatable, and much easier for a computer to work with! An admiral in the 1800's named Francis Beaufort created a wind scale model for sailing ships, which later has been built on to become the international wind scale. Table 2.1 shows the common wind speed measuring units: Beaufort's scale, the wind descriptive terms, knots and meters per second [10].

The common unit of measure for wind direction are the cardinal directions (north, south, west, east), and their intermediate points. A more precise unit used is the angle of the wind, given in degrees (°). Figure 2.1 shows the different cardinal directions, and the corresponding angle in a marine rose compass. For this thesis, we will use the wind speed unit meters per second (m/s), and the wind angle unit degrees (°).

Norway can generally be divided into two parts based on wind conditions: These are the coastline and the mainland. In the mainland, the average wind speed is around 3-5 m/s. The mountain ranges slow

Beauforts scale	Term	Knots	Meters per second
0	Calm	<1	<0,3
1	Light air	1 - 3	0,3 - 1,5
2	Light breeze	4 - 6	1,6 - 3,3
3	Gentle breeze	7 - 10	3,4 - 5,4
4	Moderate breeze	11 - 16	5,5, - 7,9
5	Fresh breeze	17 - 21	8,0 - 10,7
6	Strong breeze	22 - 27	10,8 - 13,8
7	Near gale	28 - 33	13,9 - 17,1
8	Gale	34 - 40	17,2 - 20,7
9	Strong gale	41- 4 7	20,8 - 24,4
10	Storm	48 - 55	24,5 - 28,4
11	Violent storm	56 - 63	28,5 - 32,6
12	Hurricane	>63	>32,6

Table 2.1: The international wind scale [10]

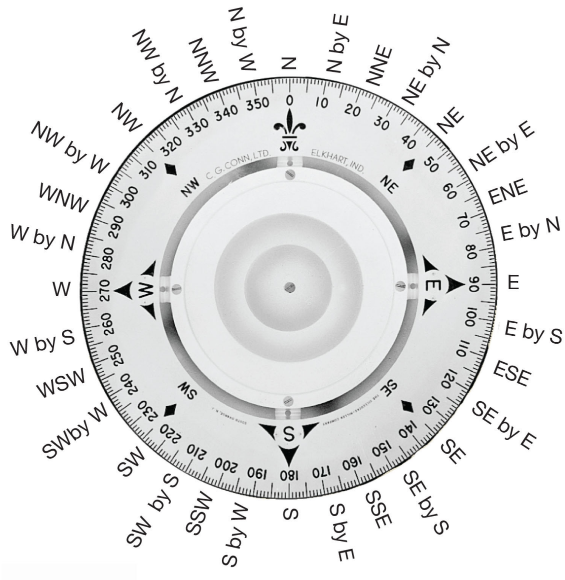

Figure 2.1: The compass rose showing both cardinal directions and degrees [37]

much of the wind in this area, and the conditions are mostly calm. Along our coast, we have an average wind speed between 7-9 m/s. According to the wind map report produced by the Norwegian Water Resources and Energy Directorate (NWE), the offshore wind blows at even higher speeds, of around 11 m/s [38].

The Norwegian coast still has one of the highest average wind speeds in Europe. Still, the area in which such high wind speeds are measured in Norway consists of a slim line that stretches up our coast, with difficult terrain like fjords. In Denmark, the average wind speed is between 5-7.5 m/s for the whole country. The same goes for Germany, which has an area up north equivalent in size to Denmark with average wind speeds of 5-6 m/s. In the middle of Germany, the average wind speed is 4.5-5 m/s. These countries are both quite flat, and the wind meets little resistance when the wind is blowing. Compared to Norway, Denmark and Germany have more area with good wind conditions. Norway's nature makes it difficult to build wind turbines, and prevents the high winds from reaching flatter grounds in Norway's mainland [55].

2.2 Humans and Wind

Humans have been using wind for their advantage for centuries. From grinding crops to moving ships, it is a well-known source of power for human kind. The earliest signs of any land wind-powered contraption were found in the lands of what is now Sri Lanka. The excavations and studies done there found evidence of a wind-based furnace for melting iron and steel. Built on a hilltop, the furnaces would have an opening facing the wind direction, so the wind would blow directly into the fires of the furnace. The evidence suggests these furnaces dates all the way back to the third century [28].

Roughly four centuries years later, in ancient Egypt, there was a man named Heron. Little is known about his life other than his inventions, and one is very special. Heron drew the earliest, undisputed record of an idea using the wind to power an axle. This was for a pump to provide air to the pipes of an organ, hence the nickname "Heron's wind-powered organ". His work dates to around first century CE. Although there are discussions of whether Heron's contraption was the first windmill or not, it is the earliest, clearest evidence of the thought and concept of using axle-driven windmill [12].

Since then, our knowledge of the wind and its advantages have increased significantly. This includes windmills, which from Heron's days have only increased in use and importance. The windmill technology, alongside the watermill, was the leading source of kinetic, natural energy for many centuries. The use of windmills peaked right before the industrial revolution, with the emergence of the combustion engine providing a more powerful and compact energy provider with many more uses than grinding corn and pumping water [46].

Still, windmill technology persisted, and found new use in the late

1900's. The American Charles Brush was conducting experiments on how to use the power of the wind to generate electricity, which coincidentally James Blyth was also doing across the Atlantic in Scotland. Brush is often credited with being the first of the two with a complete, automatically-controlled wind turbine generator, which was first used in the winter of 1887 [42].

Some years later, the Dane Poul La Cour of Askov Folk High School was granted money to erect an "experimental" windmill to study his keen interest in utilization of wind power. The result was a wind turbine that eventually powered the town of Askov and the high school with electricity and in 1890, Poul La Cour submitted his design of wind power plants for supplying electricity to villages and agriculture [56].

2.3 What is a Wind Turbine?

The traditional wind turbine today consists of a tower, a machine house and a rotor. The rotor is fitted on a horizontal axle, which goes into the machine house. The machine house consists of a gearbox (though some models are without), brakes, yaw motor, a controller, and a generator. The axle connected to the rotor has a low rotation speed, so it is connected to a gearbox. A new axle is connected from the gearbox to the generator. A brake is connected to the high-speed axle to regulate the rotation speed [43]. A generator is an electric machine that converts mechanical energy to

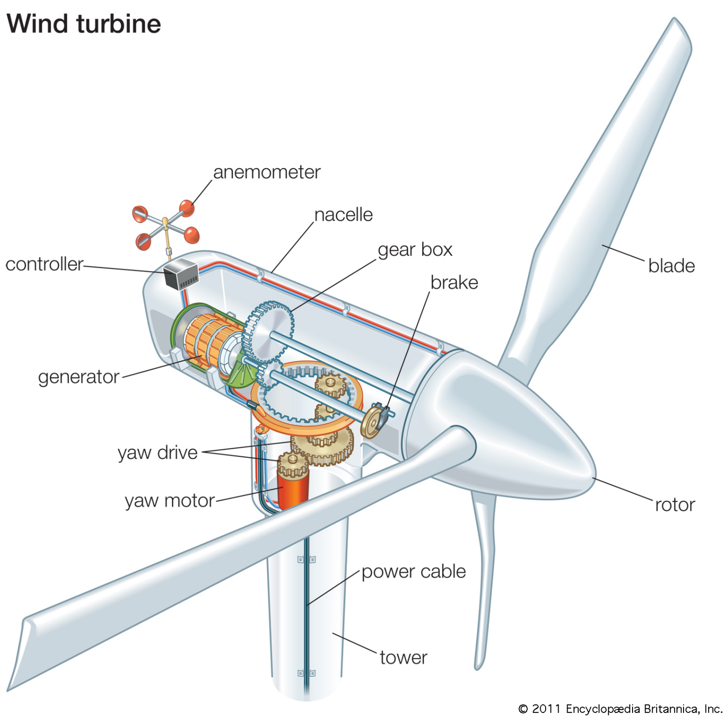

Figure 2.2: Components of a wind turbine [16]

electric energy [23], so when the high-speed axle starts rotating, electrical power is produced. The whole machine house can rotate on top of the tower, and the yaw machine is what powers this rotation. This is so the machine house can turn and face the wind in every direction, providing an opportunity at each breath of wind to generate an optimal amount of electricity. The controller is connected to the essential parts and coordinates this whole operation [43]. In total, when the wind blows, the rotor blades catch the wind and the rotor spins, generating electricity. When many wind turbines are set up in an area and part of the same network, they are called a wind farm.

2.3.1 What is a Wind Farm?

There are two types of wind farm, onshore and offshore. An onshore wind farm is the most common wind farm, and is built on land. This is the cheapest and easiest solution, but the wind conditions are less stable than offshore. Other downsides are the impact on the wildlife, especially birds. Since the spinning blades occupy much space in the air, birds tend to fly into the blades or see the turbine as a threat in the air, which disturbs their natural habitats.

An offshore wind farm is built out at sea. This can be done by attaching the tower of the turbine to the ocean floor, or provide a floating platform on which the tower can stand. This is more expensive to construct, but has many advantages after the initial cost. Offshore wind farms are out of sight for people, do not disturb wildlife on land, and have more stable wind conditions. However, other factors must be considered, like marine wildlife and safety for boats [43].

Figure 2.3: The wind farm at Smøla, Norway [15].

2.4 Wind Power Production Forecast

Wind power production forecasting is the science of predicting how much power a wind turbine or wind farm will generate in a given time frame. This has proven to be a difficult task, which includes many smaller problems. An important notion is that each wind farm is in different area, which means different nature and surroundings, maybe even climate and weather. A wind power forecast method may work great on one wind farm, and not work at all on another. Finding the aspects of a forecasting method which applies to all wind farms, and not only to a specific wind farm, is important for further development.

2.4.1 Wind Power Curve

There is a strong relation between the wind speed and the power production in a wind turbine. The wind power curve illustrates this relationship, and shows how much power a wind turbine will produce at different wind speeds. The cut-in speed is the minimum speed required for the rotor to start spinning and generate power. The cut-out speed is the speed where there is too much wind, and there is a risk of damaging the turbine, resulting in the brakes slowing the rotor down to a standstill. The wind power curve is produced in two ways. The theoretical way, where the wind turbine producers creates the curve based on the amount of power the wind turbine should produce at certain wind speeds, or the empirical way, plotting the observations of wind speed and power production on a graph to see the relationship, which hopefully results in a nice curve. If the data is good, this will produce a functional power curve, which is based on real performances of the wind turbine. Either way, the wind power curve is a

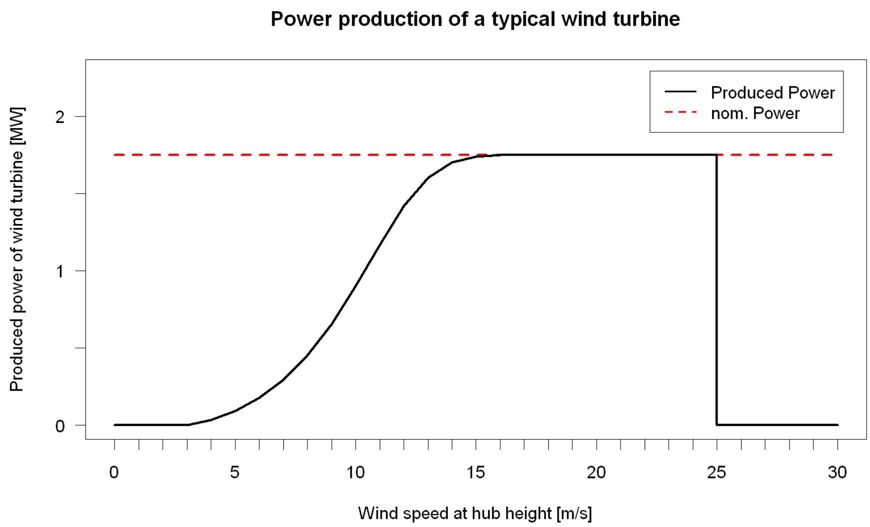

Figure 2.4: Example of wind power curve. Cut-in speed at 3 m/s, cut-out speed at 25 m/s. Maximum production achieved at >15 m/s [19].

useful tool in wind power prediction.

The simplest wind power production model is commonly known as the persistence model. This model only uses a time series of the production, and predicts the same production for the next hour as the previous hour. No estimation or calculation done, and still it produces usable results in a 4-6 hour window. Another basic prediction method is the simple use of the power curve. By using wind speed forecasts for the wind farm and applying the wind power curve, one gets a simple forecast of wind power production. Luckily, wind power forecast methods have advanced and there are more sophisticated approaches being researched. Modern wind power forecast can be divided into two approaches: Physical and statistical.

2.4.2 Physical Approach

The physical wind power forecast approach aims to describe or simulate the flow of the wind towards or within a wind farm. The flow of the wind can be simulated up to each wind turbine, giving the per-turbine wind conditions. These conditions can be applied to the wind power curve, and the output will be the wind power forecast. This is a more advanced and, hopefully, more accurate method than the previous mentioned use of the wind power curve. The physical approach requires input from meteorological forecasting methods like wind speed, wind direction, pressure, temperature, humidity etc. The approach also needs information about nature and landscape to simulate their impact on the wind flow. The physical approach attempts to apply the physical laws of nature and predict the behavior of the wind according to these laws. The physical approach is deterministic, meaning it will always output the same values given the same input. This is because the laws of nature do not change. Computational fluid dynamics is one such physical method for simulating wind flows [58].

2.4.3 Statistical Approach

Statistical wind power forecast models look at historical data from the wind farm, and aim to find the relation between certain data (such as meteorological weather forecasts), and power production. These relations can be used to predict future wind power production, based on wind conditions forecasts (or other data). Predictions can be based on a variety of data, and a combination of different data is often preferred (e.g. using wind direction and wind speed for prediction, instead of just wind speed). In most cases, the hard part is finding the relationship between the data and the produced power. Even if a relationship is found, there is still a problem for future predictions: The fact that wind speed and wind direction are only forecasts. There are no guarantees that the wind condition on the day of production is the same as forecasted. This further contributes to the challenge of accurate wind power prediction. Statistical models are non-deterministic, and their performance vary between each run. A term commonly used for data driven statistical models is machine learning [58].

2.5 Potential of the Wind Farm

The main reason for a society to build wind farms, is their lower environmental impact. Wind turbines are one of the cleanest energy production methods to this date. The energy payback time is the time it takes for the energy source to produce the same amount of total energy which will ultimately be used in its life cycle. A study conducted in Austria in 2011 investigated the pollution from different parts of a wind turbine life cycle (construction, transportation, assembly, maintenance, disassembly etc.). The results showed that wind turbines had the second lowest energy payback time of any energy sources, only beaten by hydropower. The energy payback time for an average 2-MW wind turbine is at 1.13 years [8].

Measuring the average gram of CO2e [1] emission per kWh produced (g/kWh) throughout the life cycle, is also second lowest for wind turbines. Hydro power is at the top, but wind power is second best with 9.73 g/kWh. Compared to other power sources, this is very good. Nuclear power plants show an energy payback time of over 3 years, while the life cycle pollution is around 50 g/kWh. For the coal power plant, the energy payback time is between 2.5 and 3 years, but the life cycle pollution is around 1050 g/kWh

[1]Carbon dioxide equivalent (CO2e), is a unit for measuring carbon footprints. The impact is measured in the amount of CO2 that would create the same warming effect.

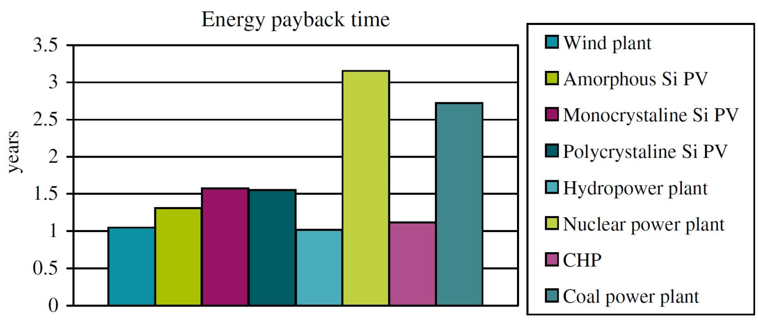

Figure 2.5: The energy payback time for renewable and non-renewable sources [8]

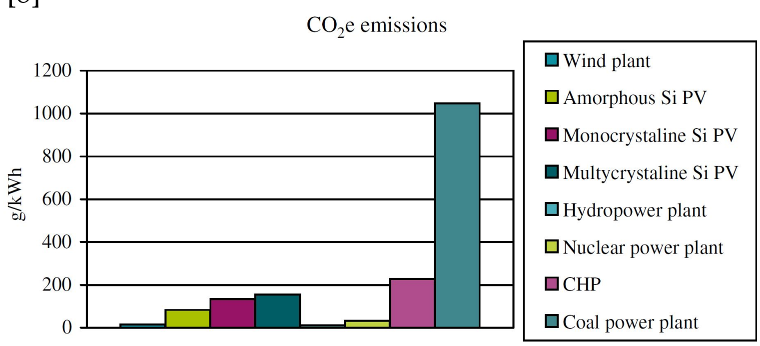

Figure 2.6: Emission per produced kWh for renewable and non-renewable sources [8]

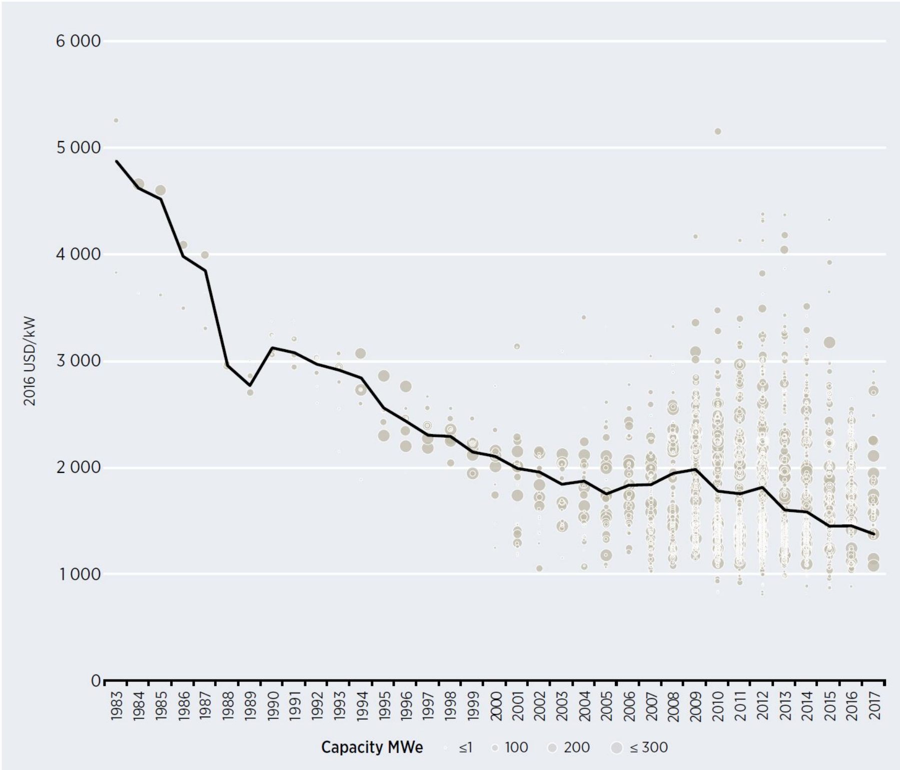

Figure 2.7: Total installed costs of onshore wind projects (gray dots, size indicating the capacity MWe[4]) and global weighted average, 1983-2017 [26].

[8].

One of the reasons behind the slow wind turbines adaption rate is likely the economic cost of wind farms. The latest report from the International Renewable Energy Agency (IRENA) [26] shows that total installation cost[2] of a wind turbine has decreased steadily since 1983. From 1983 to 2017, the global weighted average total installation cost has dropped from 4880 USD/kW to 1477 USD/kW, amounting to a price reduction of 70%. The average capacity of a typical 1985 wind turbine was at 50kW with a rotor diameter of 15 meters. Today, there are turbines installed with 8 MW capacity and a rotor diameter of 164 meters, with 9.5 MW capacity turbines are now available on the market [26]. The incentive for choosing wind power has therefore increased in later years, due to higher profitability.

The levelized cost of electricity (LCOE) is the lifecycle cost of the generating source divided over the number of energy units which it is likely to produce. The IRENA report also shows that the LCOE range for fossil fuel energy sources ranges from around 0.047 - 0.170 USD/kWh. The onshore wind energy source shows a LCOE of 0.068 USD/kWh as of 2016, while the offshore wind source shows 0.155 USD/kWh. It is predicted

[2]Total installed cost consists of five factors: turbine cost, construction works, grid connection, planning and project cost and cost of land

[4]Megawatt electric, the electric power output given in megawatt

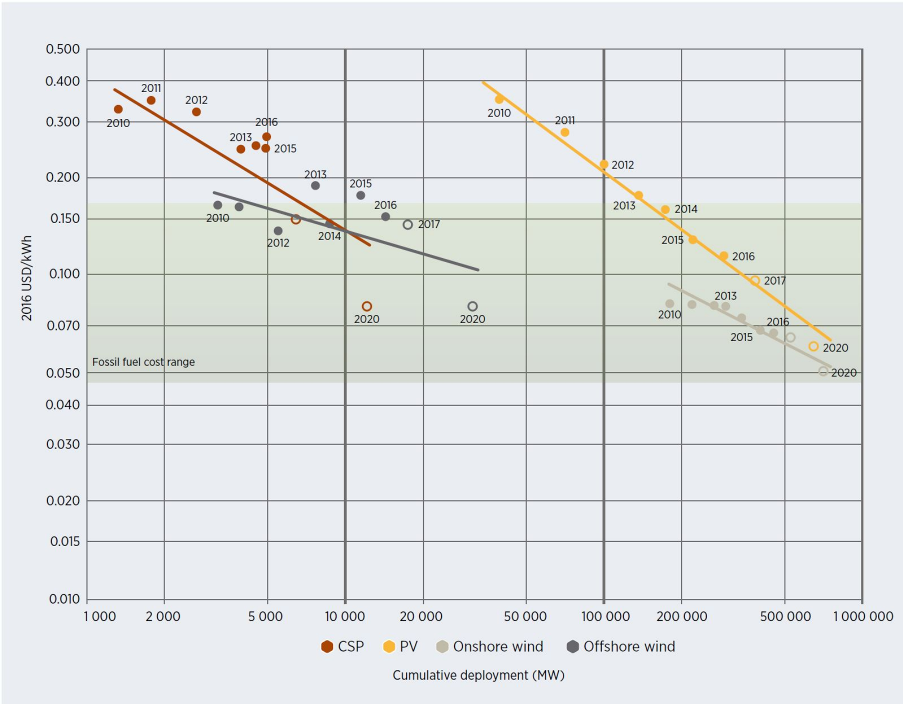

Figure 2.8: Annual global weighted average levelized cost of electricity from consentrated solar power (CSP), solar photovoltaic (PV) and onshore and offshore wind, 2010-2016, and predictions for 2017 and 2020. Learning curves show predicted future directions [26].

that the onshore wind energy source will have a LCOE of around 0.050 USD/kWh in 2020 [26].

2.6 The Power Market

The power market is where actors within the power industry buy and sell power. The amount of energy dealt within the grid system is far too large to be preserved for later use, which magnifies the importance of balance. The balance between production and consumption is the most important criterion for the power grid to stay operational and providing everyone with the power they need.

2.6.1 Statnett, Norways Power Grid Operator

Statnett is Norway's Transmission Supply Operator (TSO), and is responsible for balancing the power grid. Statnett states that there are two ways to balance the grid: Producing or trading power. Often, it is a combination of the two [49].

A crucial tool in the balancing of the power grid, is the prediction of power production in advance. This is a very hard task, and when the prediction does not exactly match, it creates imbalance. In this case, Statnett

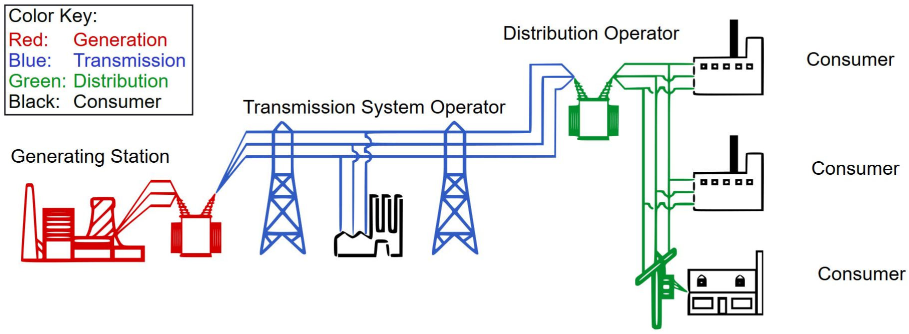

Figure 2.9: The basic operators in the electrical power grid [13]

relies on power reserves. Power reserves are sources of power which are easily regulated, so the production can be increased on demand. These are used for regaining balance when there is a sudden shortage in power. The power reserves are divided into three categories: Primary, secondary and tertiary. Momentary imbalances are covered by the primary reserve. If the imbalance endures for several minutes, the secondary reserve takes over. This frees the primary reserve, which can be used for new occurrences of imbalance. If the imbalance continues, tertiary power reserve is activated, freeing the secondary power reserves. Primary and secondary power reserves are activated automatically, while tertiary power reserves need to be activated manually [51]. Statnett's primary, secondary and tertiary power reserves all consist of power trading.

2.6.2 Nord Pool, the North-European market

In northern Europe, the electrical market is called Nord Pool, and consists of Scandinavia, Baltikum and the UK. The Nord Pool group creates a marked, which lets countries trade energy for grid balance and profit [41]. Norway and Sweden are divided into sectors. The sectors act as individual actors in the market, and can buy and sell from other sectors within the country, or other countries. If a country produces more power than its anticipated consumption, the excess power is sold to any country interested in buying. For the buyer, this may be to cover their own shortage in production, or that the price of power on the market is cheaper than the cost of producing their own power. The power market consists of two separate markets, the Day-Ahead market and the Intraday market.

2.6.3 Day-Ahead and Intraday Market

The Day-Ahead market is when all the participating countries report how much power they anticipate producing, and how much they want to buy for the following day.

The Intraday market supplements the Day-Ahead market and is where actors can trade power on the operation day. The differences between the Day-Ahead planning and operation day realities determine what happens in the Intraday market. Maybe a power plant has stopped working, greatly

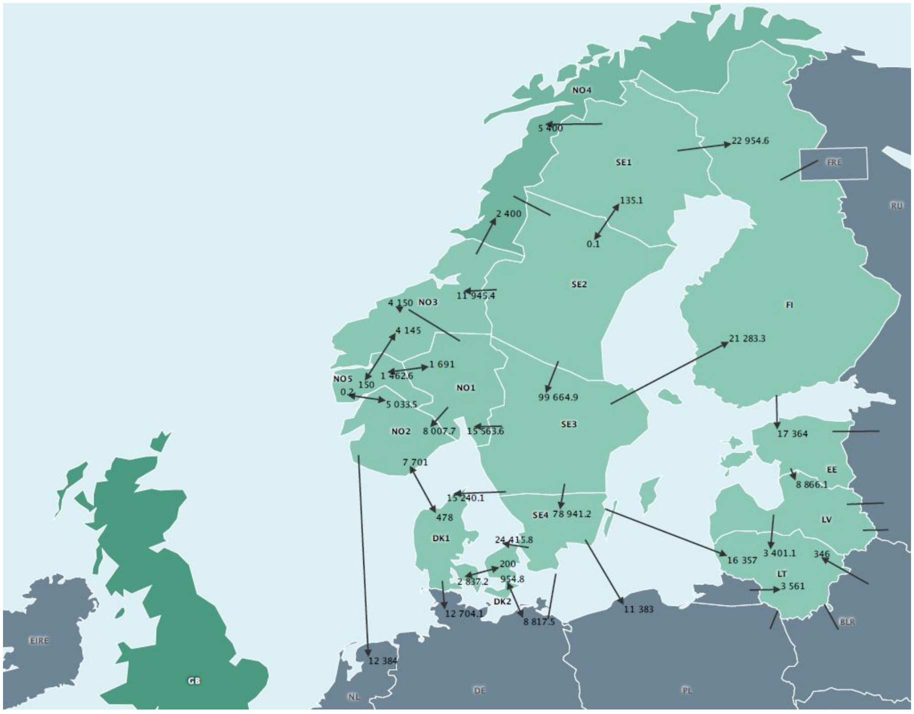

Figure 2.10: The Nord Pool group (green) trading on 10/05/18. The numbers indicate the amount of power bought, in kWh, and the arrow indicate where the power came from [40]

shortening the power production of a country, or stronger winds than anticipated made power generation from wind turbines higher, so the excess power can be sold. On the Intraday market, actors can trade up until one hour before the power is used.

The production of power in Norway, according to the Norwegian center for statistics, is divided into three sources for electricity production: hydropower (96,3%), thermal power (2,3%) and wind power (1,4%) [44]. Hydropower and wind power are renewable sources of energy, while thermal power consist of non-renewable sources such as fossil-fuel. The problem with renewable power sources is that these sources are dependent on the weather and therefore unstable, especially wind [43]. There are many reasons one would want to predict the wind power production, like grid balance as well as environmental and economical perspectives.

2.6.4 Pricing, the Merit Order Effect

Renewable sources like wind turbines are cheap in operation. This helps pushing the power prices down, creating cheaper energy on the market. Electric power does not come from a single source, but a mixture of different sources. These sources have different costs of production, and the prices of electricity follow the production costs. The pricing scheme based on power demand and available power sources is called the merit order effect. A renewable source like wind power has huge advantages in

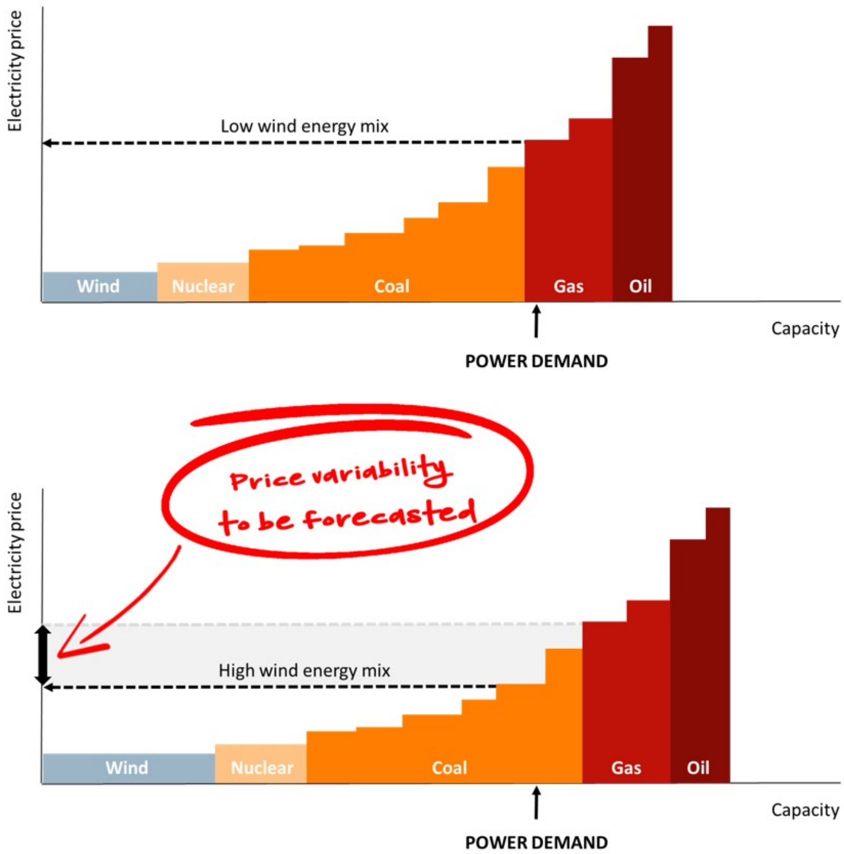

Figure 2.11: The merit order effect, and the variation in price based on the amount of wind power in the mix [20].

the power market due to its cheap operation. Wind is therefore the first source of energy included in the power mix [7].

Non-renewable sources of power can usually be regulated, meaning the amount of production can be adjusted freely. Under such controlled circumstances, the price of the power is easier to determine. In figure 2.11, we can see that the "stair" of non-renewable power sources, and that their step-wise escalating prices are dependent of the amount of wind in the mix. This means that forecasting the wind power production will also give an indication towards the price of electricity. This gives advantages when planning to trade in the Day-Ahead marked.

2.6.5 Power Grid Balance

Imbalance may be created by shortage in power production, or excessive power consumption. When imbalance occurs, it is hard to predict how the situation will develop further. The imbalance may increase or decrease, and can last from a few minutes to many days. Power reserves can remove imbalances, but can take time to be activated. By the time they are running, the imbalance may have changed and the activation of the reserve power can increase the imbalance further [48].

The ability to predict power produced from wind farms can help reducing these instances of imbalance. By knowing how much power the wind farms will produce the day ahead, the trading on the Day-Ahead market can be used to cover the biggest gaps in supply. When it comes to operation day, the wind forecasts or conditions may have changed since the day before. Wind power forecasts can then be used to trade in the Intraday market, covering any new or previously unseen imbalance occurrences. Wind power forecasts will be able to show future imbalances caused by lack of wind power production, giving the TSO the ability to cover the imbalances in advance, preventing the imbalance of actually occurring [50].

2.6.6 Fossil Fuel Power Reserves

The source of reserve power is an important factor. As mentioned, the power reserves need to be easily regulated, and renewable sources are not flexible enough. Therefore, reserves often consist of fossil fuels. Statnett's reserve solution is power trading, and report that the Nordic coal-based power production has been important to compensate for the imbalances caused by the renewable sources [47].

By reducing the occurrences of imbalances caused by wind power, Statnett also reduces the use of fossil fuels based power reserves. Instead, Statnett can in advance prepare for these incoming imbalances by buying greener energy from the market.

2.6.7 Economical Costs of Imbalance

According to eSett, the company handling power balance settlements for Finland, Sweden, and Norway, there are fees to be paid if imbalance occurs [18]. A study from Belgium looked at the cost related to the error of wind power forecasts, with focus on these imbalance costs. The results show that a higher degree of prediction error leads to a higher imbalance cost [39]. Figure 2.12 shows the average amount of primary reserve power bought per hour, each week of 2016, according to Statnett [50]. Primary power reserve is the power reserve most used. The columns show total average hourly amount bought that week.

The price tag of the imbalance in Statnett's power grid is high, and in 2016, the total price of reserve power was 85 000 000 NOK. Statnett reports unregulated power sources as a prominent issue in the balancing of the grid, and points especially towards wind farms. The volatility of wind power disturbs the balance before and within the hour of operations [50]. Table 2.3 shows the different costs of reserve power bought in 2016.

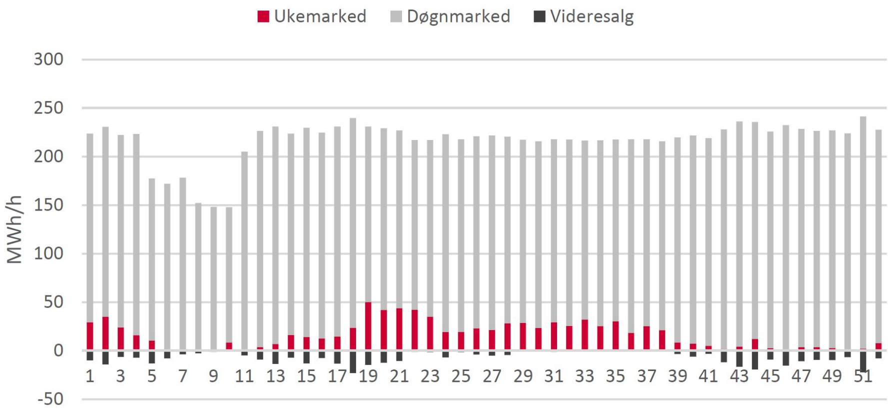

Figure 2.12: Average amount of primary reserve power bought per hour in every week. Power measured in MWh. The pink show the amount of energy bought in the start of the week, while the rest of the column represents the energy bought within a day before time of use [50]. The dark, gray column show sales.

Power source	**Price (in NOK)**
Primary reserve	85 000 000
Secondary reserve	7 000 000
Tertiary reserve	75 000 000
Total	167 000 000

Table 2.2: Statnett imbalance costs 2016 [50]

Chapter 3

Computational Fluid Dynamics

Computational fluid dynamics (CFD) is the use of computers to calculate the dynamics of fluids (gas and liquids). CFD is used, among other things, to calculate the flow of wind in a wind tunnel and how water flows through a pipe."Computational Fluid Dynamics, an introduction" by John Anderson [57] provides a good introduction to the subject.

3.1 Aerodynamics

The study of wind is called aerodynamics, which stems from Greek and means air (aero) and power (dynamic). Movements in fluids are called fluid dynamics, and aerodynamics is a sub-area of fluid dynamics [24]. Specifically we can say that aerodynamics concerns behavior of wind with objects in their way, and how the air reacts to an object moving through it [53].

The application of aerodynamics, both in theory and in practice, has for many years mainly been connected to the airplane. One of the main tools of practical aerodynamic research has been the wind tunnel. It produces good data and is a showcase for the science in use. A wind tunnel is a tunnel where artificial wind blows through. An object is placed in the tunnel and the air movement around the object is measured and observed [52]. The use of aerodynamics to optimize performance includes cars, trains, boats, athletes and even the power sector with the wind turbines.

The observation, measurement and simulation of wind has become so complex that many institutions and researchers has begun using computers to solve their problems. Combined with the decrease in the price of computational power, this has opened for possibilities in use of computers. The use of computers to solve tasks regarding fluids is its own field, called Computational Fluid Dynamics [53].

3.2 CFD's Big Breakthrough

In the 1950's and early 1960's, researchers faced one of the most important aerodynamic problems. When rockets and spaceships were entering the

Earth's atmosphere in speeds of kilometers per second, they generated very high amounts of aerodynamic heat. A report from National Advisory Committee for Aeronautics from 1958 showed that a blunt tip of the aircraft produced significantly less aerodynamic heat than a pointed tip [3]. The problem was that there were no known theoretical methods in aerodynamic to predict the airflows over the blunt tip body shape. Many years and resources were allocated to this problem, and as late as the mid-1960's it was still theoretically unsolvable.

In 1966, two scientists named Abbett and Moretti solved the blunt tip body problem using computational fluid dynamics [34]. It marked the first solution for this problem, and the industries and governments struggling with the same problem simply applied the computational technique. Computational fluid dynamics showed its potential and quickly became a third approach to solving fluid dynamic problems, in addtion to the classical approaches of theory and experimentation [57].

3.3 Physical Modelling

Computational fluid dynamics uses three fundamental principles in its fluid simulation:

- Mass is conserved
- Force = mass * acceleration
- Energy is conserved

Assuming these physical principles, we need a model to apply them to. For the moving fluids, there are two common models.

3.3.1 Finite Volume Control

This model focuses on a finite region within the flow. This region is defined by the control volume V and control surface S which bounds the control volume. Now that we have defined which portion of the flow we want to focus on, the control volume, there are two options. The control volume can be fixed in space and look at how the fluids behave moving through the control volume, or the control volume is moving with the flow, always containing the same fluid particles. The three physical, fundamental principles are applied to the fluid particles within the control volume, and the advantages of this approach is that we do not need to model the whole fluid flow, only the fluid particles within the finite region of the control volume and the control surface (if fixed in space) [57]. See figure 3.1a for illustration.

3.3.2 Infinitesimal Fluid Element

This model focuses on a fluid element in the flow which is infinitesimally small. The volume of the element is dV, and the element is a continuous

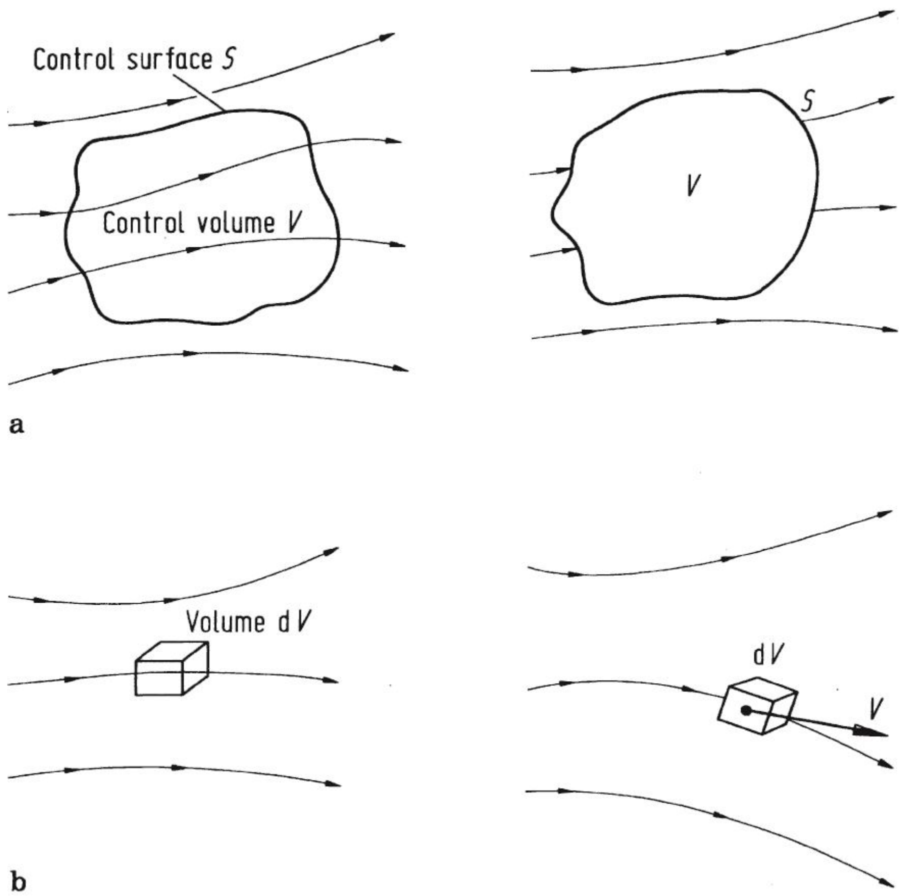

Figure 3.1: a) The finite Volume Control model b) The Infinitesimal Fluid Element method [57]. The flow is represented by streamlines.

medium. As with finite control volume, it can be fixed in space, with the fluid flowing through the element. Otherwise, it can move along with the flow, with the vector V being set to the flows velocity at any given point. This is another model where we do not look at the whole flow field at once, and we only apply and calculate the physical principles for the fluid element [57]. See figure 3.1b for illustration.

3.4 Mathematical Approach

At the core of the CFD approach lies the complex Navier-Stokes equations. The answers to the Navier-Stokes equations is everything we want in fluid dynamics: The motion and behavior of the fluids. Computational fluid dynamics provides a way of solving these equations. These are complex equations, and getting the answers is not easy. Luckily, we have the help of computers.

3.4.1 Analytical vs. Numerical Methods

Mathematical equations can be used to express the three fundamental principles of fluid flow, and partial differential equations are the most general form to use [57]. The analytical approach for solving these equations are hard, but techniques for solving the linear and "almost" linear partial differential equations do exist [17]. When the problem is a non-linear partial differential equation, there is no known general analytical method that can be applied. Therefore, numerical methods are more appropriate. Numerical methods use both an algorithm and a set of numbers to solve the equations. They iteratively try the numbers given and calculate whether the numbers solve the equations. If not, they change the number and try again [22]. This can lead to thousands or millions of calculations, which for a human would be impossible to finish. In this manner, it tests its way toward a solution, trying out all numbers given until one set of numbers solves the equation. The CFD's elevated level of detail and sophistication can only be credited to the computers, and this is also where the future of CFD lies. Powerful computers with more storage can increase the complexity and detail of the calculations, which again can be applied to more complex problems [57].

3.4.2 Navier-Stokes Equations

The Navier-Stokes equations are a set of non-linear partial differential equations which calculate the behavior of fluids. The solutions to these equations reveal the motion and behavior of the fluid. The equations originate from Leonhard Euler, and his attempt to calculate the flow of incompressible and frictionless fluids. Later, the Frenchman Claude-Luis Navier added the element of friction to the equation, so it could be applied to viscous (friction) fluids. Finally, the British Sir Georg Gabriel Stokes further improved the equation. The complexity of modeling fluids in three dimensions has proven not manageable for any other method than through numerical analysis [25].

Initially the Navier-Stokes equations were connected only to the equations for momentum in viscous flows. Later, the Navier-Stokes equations came to refer to the principles of continuity and energy, thus representing the entire system of equations for flows in viscous fluids [57]. When solved, Navier-Stokes describes the density, velocity, temperature and pressure of moving fluid.

3.5 WindSim CFD Program

The program we used for computational fluid dynamic simulations was provided by WindSim AS. The WindSim program is used to model the physical conditions of an area and simulate the flow of wind as it blows over the landscape. This can be done by solving the Navier-Stokes equations, and creating a numerical simulation of the flow of wind.

Figure 3.2: The real world and the digital terrain model [21]

Computational fluid dynamics and the simulation of fluids can be complex. For guidance, we can divide the simulation into four stages: Real world, mathematical model, discrete model and discrete solution.

Real world

First we need to find the landscape we want to run the simulations on. For our purposes, this means the wind farm and its surrounding landscape. The complexity of the region and the size of the area are things to consider when determining the scope. Using computational power to simulate high-complexity terrain far away from the area in focus is generally wasteful.

Mathematical Model

The mathematical models to be solved in the simulation generally consist of the Naiver-Stokes equations. As we know, CFD can be used to simulate all fluids, from honey to gas, so we insert into the formula the parameters corresponding to air. Specifically, we need to input the characteristics of the atmospheric boundary layer (ABL). The ABL is the lowest region of the atmosphere, and temperature, moisture and wind so low are heavily influenced by the earth's surface.

Discrete Model

Finite-volume method is (3.1) is the discrete model used in the WindSim simulation. We take the region, which we defined in the first step and make a grid for the whole area. Now our geographical terrain is parted into 3D cells, and the size of these cells are defined by how granular we want our simulation.

Discrete Solution

Finally the Navier-Stokes equation we made ready in step two is applied and solved for each of the cells. Each cell is a simulation of a small area,

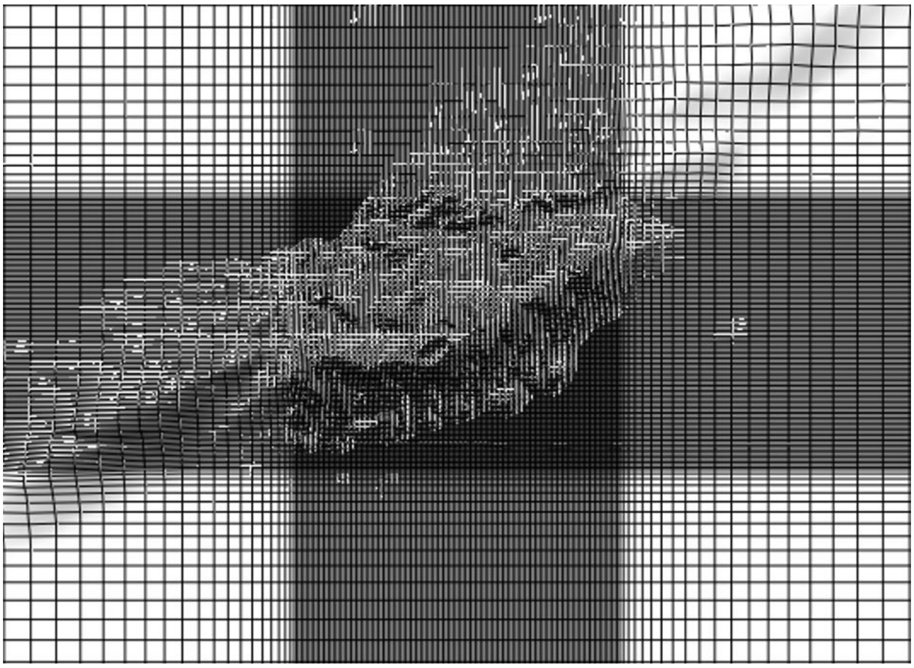

Figure 3.3: The grid mesh of the area [21]

and all the cells together show the big picture of the wind's movement over the terrain.

Output

The output is the numerical representation of the wind. The numbers can be used as input to further models or used to graphically visualize the wind.

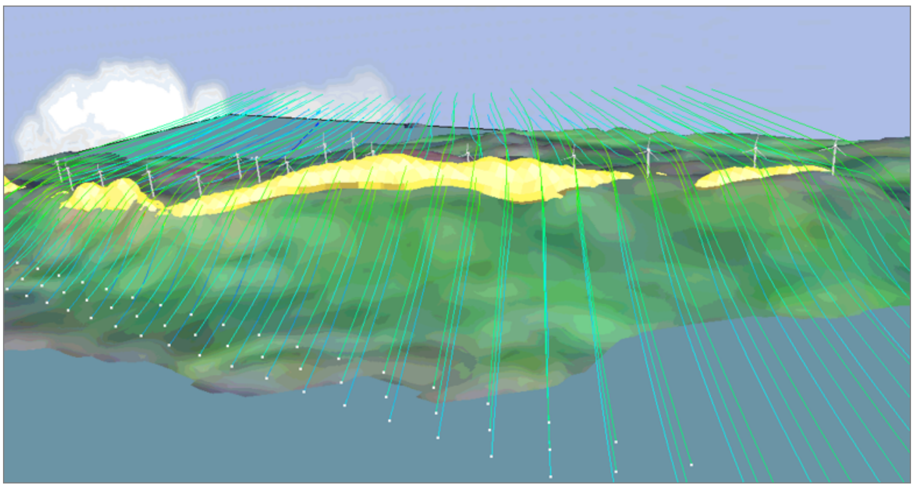

Figure 3.4: The wind simulation [21]

3.6 CFD Today

Today, computational fluid dynamics is widely used in industry. The Norwegian company CFD Marine AS uses CFD in the marine sector. They simulate the movement of different marine vessels through the ocean, and how the ocean responds to these objects. They focus on lifeboats, and started using CFD to model lifeboats when they are dropped from oil platforms, breaching the water surface at high speeds. The CFD method of modelling supplements or replaces traditional experiments and can be used to model conditions which would be difficult to replicate in experiment, like extreme weather and big storms [30].

British company BMT Fluid Mechanics write they are experts in the use of computational fluid dynamics in the fields of oil and gas. Flow assurance, as they call it, focuses on the transport of different liquids in the oil and gas industry, such as crude oil, gas, water and sand. By modelling the fluids behavior in pipes, they can optimize pipes and remove unwanted conditions or problems which may appear in such a process. CFD is a key component in their work, and their models use CFD [33].

WindSim AS is a Norwegian company which specializes in simulation of wind. Their focus has been on optimizing the planning of wind farms, as well as wind resource and energy assessment in general. WindSim AS uses computational fluid dynamics to model the wind and how it flows and reacts to landscape and objects [5].

Chapter 4

Machine Learning

Machine learning is a field in computer science, which uses methods from statistics to give computer algorithms the ability to learn from data and use the learned knowledge to perform tasks (prediction, recognition etc.). The algorithm tries to find a pattern in the data and use this to improve its performance of the task. This is in contrast with most algorithms, which perform tasks based on pre-programmed rules [14]. Data-driven models attempt to learn these rules by themselves. Machine learning is therefore useful when problems become too complex to define through a set of rules. With the power of modern computers, highly complex and advanced data-driven models can be created at home [14].

4.1 Categories of Machine Learning

There are three main categories of machine learning: Supervised learning, unsupervised learning and reinforcement learning.

4.1.1 Supervised Learning

Supervised learning is an approach to machine learning where the model tries to approximate an unknown function given examples of input and corresponding output. The model makes a prediction from the provided input, and adjusts its parameters (weights) according to the difference between the prediction and the target output [14].

Consider the following example: We want an algorithm to classify pictures into two classes, pictures with cats and pictures without cats. The model starts with a random set of weights. We then sequentially feed the model the provided images, and compare them to the corresponding labels of 'CAT' or 'NO CAT'. If the model's predictions are correct, we change nothing. If the model is wrong, we change its weights in such a way that it is more likely to be correct given the same image again. Note that this requires a way of translating the error to a desired change in weights. After showing the model enough labeled pictures, we hope that it is able to generalize and recognize cats, even in pictures which it has never seen before.

This is an example of classification, where the outcome is a member of pre-defined classes. If we want to work with numbers and get an estimate or prediction in a numerical value, it is called regression [14].

4.1.2 Unsupervised Learning

In unsupervised learning, we supply the model with unlabeled data and leave the model to find relations and structures in the data. Perhaps the most common objective is to cluster the data in groups of similar characteristics, or finding unknown patterns in the data [14].

Keeping with the above example, consider providing a model with a set of pictures. This time the pictures consist only of black cats and white doves. Hopefully, if we supply these to an unsupervised learning model with clustering objective, it will cluster the data into two groups, one containing the cats and one containing the doves.

Humans are very good at finding patterns and structures, and often outperform computers on low dimensional data, or data which has a sensible visual representation. However, since humans depend on visual information, they are useless when dealing with higher dimensional data, and this is where machine learning can be applied [14].

4.1.3 Reinforcement Learning

In reinforcement learning, we think of the model as an agent in an environment. The environment is constructed to provide the agent with feedback in forms of punishment and reward. By choosing appropriate punishments and rewards, we want the agent to find a solution to a task by simply by interacting with the environment. Significantly, the agent is never told what it did right or wrong, or how it can improve its solution. The agent is simply punished or rewarded for its actions and must learn from this experience to improve its solution [14].

Consider designing a chess-playing computer program. Chess is exponential in complexity, and is thus impossible to solve using a set of rules. It is also impossible to label a single move as good or bad, meaning the program cannot be trained in a supervised fashion. Instead, we let the program play many games of chess and apply rewards for capturing pieces or winning, and punishments for losing pieces or losing. In this fashion, we hope that the program learns to models the value of a move according to the set of likely outcomes weighted by score and immediacy. This kind of learning is applicable to problems where the feedback is delayed, and the result of many consecutive actions.

4.2 Machine Learning, Step by Step

Teaching a supervised machine learning model to perform its task follows some general steps: Data acquisition and cleaning, training, and testing. In the following example, we try to estimate housing prices in Oslo using a supervised regression machine learning model.

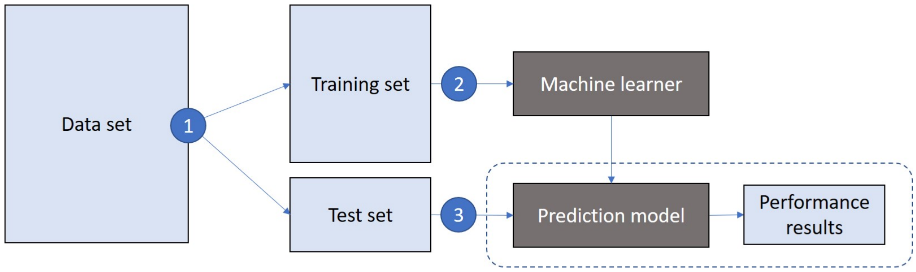

Figure 4.1: The typical process of the supervised method

4.2.1 Data Acquisition

When we talk about the dataset, we refer to the labeled data provided to the model. In the case of regression models, the input data can consist of multiple values, while the output is usually a single value. In our example of housing prices, the input data can be size (in square meters) and location (in postal code). The output data is the value of the house. We divide the dataset into two parts, one for training and one for testing. Generally, we want the training set to be as large as possible, while keeping our test set large enough to be representative. This fraction is up to the programmer to decide based on the data and prediction task. In our example, we choose a 70-30 split [31].

4.2.2 Training

Before the training step, the model is completely untouched, and does not know anything about the data or its task. We supply the model with the training data.

The model tries to approximate the unknown function between the input data and the targets of the training set. For each entry in the input dataset, house size and location, the model tries to estimate the expected value of the house. When the model is wrong, it adjusts its function slightly and tries again on the next entry. There are other methods for training as well, for example where the model only adjusts after a certain amount of entries (batch). It is also common to let the model work its way through the training set several times during the training step [31].

4.2.3 Testing

After the learning step is done, we have a trained predictor. In the testing step, we run our model on the test data, and record the performance. This is done to see how well the model performs on previously unseen data. The data still consists of house size and location as input, and value as output, but it is unknown in the sense that the model has likely not seen these exact combinations of size, location and value before. Hopefully, the model has generalized enough to give a good estimate of the value of these unseen houses, and has learned a deeper relationship between house size, location, and house value, and not just the specifics of the training data [31].

4.3 Different Machine Learning Models

We will here explain the general workings of the models we use in this thesis. The models are:

- Artificial Neural Network
- K-Nearest Neighbors
- Decision Trees
- Random Forest Regression
- Linear Regression
- Support Vector Regression

4.3.1 Artificial Neural Network

The artificial neural network (ANN) is a model which is inspired by the neurons and synapses of the animal brain, and how they interact to learn and create representations of the world. In animals, neurons communicate through signals of electricity. The neurons are arranged hierarchically, with the first neurons responding to external stimuli, and subsequent levels of neurons creating increasingly complex abstractions based on the signals they receive from other neurons. The ANN is an attempt to model this structure. It consists of an input layer of neurons, a number of hidden layers, and an output layer. The neurons each have an activation function, which calculates the signals fired from the neurons in the previous layer and decides whether it should fire a signal to the next layer or not. All neurons are connected to every neuron in the subsequent layer, and weights are introduced on each connection to adjust the importance of a signal from the sending to the receiving neuron [31].

When training the neural network, the weights of the connections are what we adjust. Training usually consists of two phases, forward-pass and back-propagation. Training starts with the forward-pass, where input is fed to the network and an output is produced. The output is then compared to the target, what we wanted the network to produce given that input. A loss is calculated based on the difference between the output and the target. Choosing an appropriate loss function is important for the model to learn effectively. During the back-propagation phase, the weights of the model are adjusted based on the loss. We calculate the gradient of the loss function with regards to the weights of each layer of the model, and adjust them in the opposite direction the gradient.

Now we feed the network the next input, and we are back to the forward-pass phase. This cycle goes on until the neural network no longer improves and the training stops.

4.3.2 K-Nearest Neighbors

K-nearest neighbors (KNN) represents a simple supervised approach to machine learning. It predicts the value of an unseen data point by remembering the value of all training data, and interpolating between the k training data points which are closest to the new data point. Note that this requires a distance measure.

Consider again the prediction of housing prices, this time based on the age and size of the house. We train a KNN model on a dataset of houses with their age, size, and price, and choose k=2. We then give the model a previously unseen data point consisting of a 30 year old house with a size of 100 m^2. Since we chose k=2, the model searches through the training data and finds the two houses closest to this new house by Euclidean distance (this is also a design choice). The two closest houses have sizes of 95m^2 and 110m^2 respectively, and they are 28 and 29 years old. Their prices are 1MNOK and 1.4MNOK, meaning that the model predicts a price for the new house of $(1MNOK + 1.4MNOK)/2 = 1.2MNOK$.

In some cases, there can be an advantage that the closer neighbor's contributions are more important than neighbors further away. We introduce weights for this purpose. They can be set as uniform or adjusted according to distance. Normal approaches for calculating distances are brute force search or KD-tree (a binary tree search approach). The computational cost of brute force is O(N^2) while KD-tree is O(N log N), but there can be instances where KD-tree does not terminate, while brute force will always find the nearest neighbors. Distance is normally measured in Manhattan or Euclidean distance. An issue with k-nearest neighbors is dimensionality. When the dimensions of the data increases, so does the computational cost [31].

4.3.3 Decision Tree

A decision tree constructs a tree graph where each leaf is an output value, and the input values determine how we iterate down the tree. The tree forms a control sequence, starting at the root node, and discriminating the input at every subsequent level, until it reaches a leaf. The tree is constructed by choosing, at each node, the discriminator which maximizes the entropy of the result. The entropy in this case means the amount of information gained by knowing the value of an input.

Consider again houses on the market. We want to determine, based on a house's price, age, and size, whether or not it will be sold within a week of being listed. This is a classification problem, with classes 'SOLD' and 'NOT SOLD'. After training, we can imagine that the price of a house is the most determining factor, and so the first discrimination is based on price. Next, the size is determined to be most important, and so the data is split based on this factor. Finally, the data is split based on age.

When working with continuous output values (regression), the model functions in an equivalent way, but returning a weighted average of the data partitions [31].

4.3.4 Random Forest

A random forest is an ensemble learning method. Ensemble learning is the use of many machine learning model working cooperatively.

The way ensemble learning methods work is by combining many machine learning algorithms, each providing a slightly different answer. The models catch distinctive characteristics or traits of the data, and by combining all the answers we can get a significantly better answer than what a single algorithm could provide.

A random forest builds on the principle of wisdom of the crowd, and applies it by constructing a forest of decision trees. Random forests use the bagging method. Bagging is done by sampling the original dataset with repetition, and using this for training. By doing this for each tree in the forest, we ensure that each tree gets a slightly different training set, and so focuses on different aspects of the data. To include even more randomness, the random forest regressor reduces the available choices the decision trees can make at each split, only making a random subset available instead of the complete set of choices.

The forest now consists of lots of unique trees, all focusing on different aspects of the data. When input is received, it is given to all the trees. The trees then each have a vote, and either the answer which has the most votes or an average is chosen as the output from the forest [31].

4.3.5 Linear Regression

Consider data as points in a two-dimensional plane. Linear regression is a method for constructing the straight line which best represents the data, but can of course be applied to data of any dimension. This is done by minimizing the Euclidean distance between the line and each point. This can be solved analytically, through the normal equations:

$$(X^T X)^{-1} X^T y \tag{4.1}$$

Where X is a matrix, where each row is a data point, and y is a vector of responses. However, as the size of the data set grows, calculating $(X^T X)^{-1}$ becomes infeasible, and numerical methods are required.

To make predictions, linear regression finds the point on the line closest to the input [31].

4.3.6 Support Vector Regression

Support vector regression (SVR) is a modified version of the support vector classifier, which has increased much in popularity in recent years due to impressive performance on medium sized datasets.

SVR attempts to find a hyperplane which best describes the data. The hyperplane is found using a quadratic programming solver, optimizing with regards to the distance to the points falling outside the area defined by the points within a distance ϵ from the hyperplane. Prediction is done by finding the point on the hyperplane closest to the new data point.

Parameters in the support vector regression are ϵ (the radius distance in the tube), C (the penalty parameter for the error function e-incentive error function) and the so-called kernel function [31].

4.4 The Use of Machine Learning

The use of machine learning algorithms has increased much in the later years, and chances that you interact with one on a daily basis is quite high. Algorithms for spam filters, voice recognition, translation, and driving route planning all typically use a flavor of machine learning [31].

Tesla

The car manufacturer Tesla are known for their highly advanced electrical cars. In their later models, they have implemented an autopilot system which gives the car the ability to drive all on its own. They have fitted the car with eight cameras, twelve sensors and a radar to get a 360-degree view of its surroundings up to a range of 250 meters. This equipment produces a lot of data, which is stored in the car. At the core of the computational system sits a so-called deep ANN. The ANN analyzes incoming data in real-time and uses this to create and maintain vision of the car's surroundings. The ANN can then control the car based on this information. Tesla claims that the car can match the traffic speed, keep or switch lanes, exit the freeway, park itself and be summoned to and from its parking space. All without any directions from the driver [54].

Facebook

Facebook has created a face recognition program called DeepFace. DeepFace is a nine-layer deep ANN which was trained on the largest facial dataset to that date. Over four million labeled pictures were used, belonging to over 4,000 different people. They also tested on other, smaller datasets. Their method reached an accuracy of 97.53% on one of the smaller datasets. This is starting to reach human level of accuracy. Facebook wants this program to be generalized enough that it can be applied to different populations without any modification. If this goal is reached, Facebook has created a tool which can find all photos you appear in [59].

Chapter 5

Our Work

Our focus has been on prediction of wind power generated, given the wind conditions. We have used machine learning and computational fluid dynamics. Since our knowledge of the field of CFD is limited, the project has been conducted with help of WindSim AS. WindSim AS is a company based in Tønsberg, Norway. They specialize in the simulation of wind and have been the leading provider of computational fluid dynamic solutions in this area for years [5]. It was WindSim AS who wanted this project conducted, and after contacting the University of Oslo they managed together to create a fitting master's thesis project. WindSim AS has been responsible for the computational fluid dynamics part of this project and been supplying their algorithms and knowledge.

5.1 Related Work

Much research has been done on wind power prediction. The following section summarizes some of the most relevant projects.

5.1.1 CFD Approach

In 2013, WindSim AS has conducted research regarding the combination of CFD and neural networks. They trained a neural network to predict the wind conditions at the wind farm given the wind conditions from a meteorology mast. These wind conditions predicted by the neural network is then used by the CFD to predict the power output on a per-turbine level. They found that using CFD improved accuracy for high production periods and calm periods [32].

Castellani et. al. proposed a hybrid combination of an ANN and CFD, similar to WindSim's approach. They applied two main approaches. One pure ANN approach directly predicted the power. The other approach trained the ANN to forecast wind speed and wind direction from a reference meteorology mast towards the wind farm, and CFD to forecast the power output on a per-turbine level. They conclude that due to the roughness of the local terrain, the CFD did not perform as well as hoped and limited the hybrid solution accuracy [1].

5.1.2 Machine Learning Approaches

Among the machine learning algorithms used for prediction of wind power, the most popular in literature is the ANN [9] [2] [36].

Bhaskar et. al. looks at both an adaptive wavelet neural network and feed-forward neural network for their wind power forecast, with a time-period up to 30 hours. Their approach is interesting because instead of using wind speed forecasts from a meteorological institute, they generate their own wind speed forecasts first. This was done by decomposing historical wind data to wavelets, and train an adaptive wavelet neural network (AWNN) to predict the wind speeds for the next 30 hours. For the wind power forecast, a standard ANN was used and trained on historical wind speed and historical power output. The ANN is then fed the 30 hour wind speed predictions from the AWNN, to make a per hour day-ahead power forecast. The results were measured in root mean squared error (RMSE), and ranged between 2.849 MWh and 16.090 MWh from the first to the last prediction hour. Average hourly RMSE of their method was 10.221 MWh, with a wind farm capacity of 176.8 MW [9].

Men et. al. produced an ensemble of mixture density neural networks for short-term wind speed and power prediction. The results were however unpromising, with a RMSE of 176.72 kWh, and a wind turbine effect of 650 kW [2].

Olaofe increased the prediction window and used a layer recurrent neural network to predict the wind power output for 30 days at two different wind farms, Paarl and Vredenburg in South Africa. The layer recurrent neural network (LRNN) consisted of an input layer, two hidden layers and an output layer. There was a connection between the output and the input of the first hidden layer, creating a recurrent layer. The recurrent layer used only its own previous output as additional input. Two models with different numbers of neurons were used, one with a recurrent hidden layer of 24 neurons, the other using 15. For both models, the second hidden layer and output layer was a single neuron. The data consisted of a one month time series including wind speed, wind direction, wind gust, humidity, air temperature and atmospheric pressure. The entries of the data were sampled at a 5 and 10 minute intervals. The results show that a 30 day power forecast for the Paarl farm using 24 neuron model and 5-min interval data resulted in a average symmetric mean absolute percentage error (sMAPE) of 0,096%. The 30 day power forecast for the Vredenburg farm using 15 neuron model and 10-min data interval resulted in a average sMAPE of 0,069% [36].

Support vector regression is another machine learning algorithm which has been widely used in the field [4] [60].

Kusiak et. al. tests several machine learning models. The models tested were SVR, ANN, decision tree and random forest regression. The input data was wind speed, wind direction, average wind speed in the lower atmosphere, average wind direction in the lower atmosphere, average air density in the lower atmosphere and potential temperature difference in the lower spaces. This data originally consisted of 10-minute intervals, but was converted to hourly measurements. The models forecasted 6 hours ahead. The results were measured in Mean Absolute Percentage Error (MAPE). The ANN was best, with 10,94% MAPE, follow by SVR with 16,98% MAPE. Random forest regressor got 22,19% MAPE and decision trees landed on 25,43% MAPE [4]. The ANN was chosen for further testing, but did not improve substantially.

Zeng and Qiao used SVR for wind power prediction, and tested prediction performance between 2 and 24 hours. The data consisted of historical wind data, measured at 10-minute intervals over two years. The model was trained on wind speed, power output, and a statistical estimation of power output called SCORE-lite power. For predictions longer than 6 hours, the data interval was changed from 10-minute entries to 2-hour entries, using the average. The SVR was used to first predict the wind speed for the current hour, using the historical wind data from earlier hours. Based on the predicted wind speed, the model then predicts the wind power production. The error rate increased with the time from prediction to real-time. The SVR for one hour forecast had 1,07% MAPE, two hour forecast 2,87% MAPE, three hour forecast 5,36% MAPE and for 8 hours the MAPE rises to 12.11%. The paper concludes that forecasts over 24 hours ahead should be complemented by extra meteorological variables as temperature and pressure [60].

Marĉiukaitis et. al. used a non-linear regression with only wind speed as input to predict one wind turbines power curve. Such a power curve can be equally correct for predicting a few hours to many days, all dependent on the quality of the wind speed forecasts. In their test set, they acheived a MAPE of 8.4% [45].

Support vector regression and K-nearest neighbors are implemented to look at single turbine and total wind farm power prediction by Treiber et. al. The dataset was the same as Zenf and Qiao use in their SVR project [60]. Their model predicted power output of a wind farm consisting of 25 turbines. This was done by building one aggregated time series using the summation of the power outputs of all individual turbines. The SVR outperforms the KNN on both single turbine prediction and total park prediction, but on the latter prediction, KNN performs much better than on a single turbine.[27].

Kusiak and Zhang looked at ANN, SVR, KNN, boosting trees and random forest regression in their attempt to predict wind turbine power output. The time frame of the prediction was very short, from 10 seconds up to one minute. The ANN performed best, with a MAPE of 0.026%. KNN was close, with a MAPE of 0.034%. The rest of the models had MAPE of over 0.12%. They used wind speeds as input for their power forecasts, but also two controllable parameters: the generator torque and the blade pitch angle of the wind turbine [29].

The time-period of the predictions in the projects discussed varies from ten minutes [4] to many weeks [6]. In the paper "Current status of wind energy forecast and a hybrid method for hourly predictions" from 2017, Okumus and Dinler looks at over 80 different articles containing the latest efforts in wind prediction tools and methods. They write that the typical error rate in 12-24 hours predictions, measured in MAPE, is in the range of 15-20%, while for 24-72 hours, it rises to 25% [35].

5.2 Data

The data for this project has been supplied by Yr and the Norwegian Water Resources and Energy Directorate (NWE). The data consists of hourly entries.

5.2.1 Yr Forecasting Service and Wind Data

Yr is a weather forecasting service, and the Norwegian Meteorological Institute supply their data. Most of the weather data are free for the public to use and can be downloaded at their site Yr.no.

Wind Data

The weather data is not the actual recorded wind conditions at the different farms, but forecasts by Yr. The calibration points used for the forecasts has been put in the middle of each of the wind farms. In other words, there are no recorded measurements of the wind conditions for these wind farms. Yr's forecasts are the wind data for this project. The calibration points were put in the middle of these wind farms: Hitra, Bessaker, Smøla, Valsneset and Ytre Vikna. The data consists of hourly entries between 21/12/2016 00:00 and 24/01/2018 23:00.

The data is represented in a comma separated file (CSV). The CSV consists of seven columns: "year", "mon", "date", "hour", "min", "dir[degree]" and "speed[m/s]". "year", "mon" and "date" are the year, month and date of the entry. "hour" and "min" was the hour and minute that entry was recorded. The data was recorded with an hour interval, at each hour shift, so the value "min" is "0" for all entries. "dir[degree]" represents which angle the wind is coming from measured in degrees (°), and "speed[m/s]" is how powerful the wind is, measured in meters per second (m/s).

5.2.2 Norwegian Water Resources and Energy Directorate

The power output data from the different wind farms was provided by NWE, who are responsible for managing Norway's water and energy sources. NWE publishes their data, but due to the sensitivity of recent power production data, they are legally obligated to wait for three months. A deal was made between WindSim AS and NWE, where we could get the power production data on the terms that this thesis was not published before the three-month time frame was over (this ended 01/04/2017).

year:	mon:	date:	hour:	min:	dir[degree]:	speed[m/s]:
2017	1	1	1	0	339.1	5.8
2017	1	1	2	0	349.2	4.1
2017	1	1	3	0	348.1	4.1
2017	1	1	4	0	344.5	4.2
2017	1	1	5	0	17.6	12.4
2017	1	1	6	0	15	13.6
2017	1	1	7	0	8.9	14.8
2017	1	1	8	0	3.3	12.4
2017	1	1	9	0	359.6	9.7
2017	1	1	10	0	345.7	10

Figure 5.1: 10 first rows of Yr wind data, for Hitra wind farm

Date/Time	Smøla	Bessakerfjellet	Valsneset	Hitra	Ytre Vikna
01/01/2017 01:00	90.91	31.861	8.39	36.572	12.48
01/01/2017 02:00	72.31	28.345	8.68	32.934	17.184
01/01/2017 03:00	60.52	27.551	8.12	19.928	16.8
01/01/2017 04:00	71.36	17.679	7.54	25.497	24.576
01/01/2017 05:00	122.79	17.898	7.84	36.462	32.64
01/01/2017 06:00	142.6	4.105	9.07	53.608	36.384
01/01/2017 07:00	140.01	35.274	11.32	54.311	36.384
01/01/2017 08:00	139.91	42.086	11.3	46.158	33.792
01/01/2017 09:00	140.35	40.088	11.29	52.841	31.2
01/01/2017 10:00	140.63	34.501	11.26	51.808	35.136

Figure 5.2: 10 first rows of NWE power data

Power Data

The data is the recorded power produced for the five different wind farms, Hitra, Bessaker, Smøla, Valsneset and Ytre Vikna. The data consists of hourly entries between 01/01/2017 01:00 and 01/01/2018 00:00.

The data is represented by a comma separated file (CSV). The CSV consists of six columns, "Date/Time", "Smøla", "Bessakerfjellet", "Valsneset", "Hitra" and "Ytre Vikna". "Date/Time" was the date, month, year, hour and minute for the entry, represented as this "01/01/2017 01:00". "Smøla", "Bessakerfjellet", "Valsneset", "Hitra" and "Ytre Vikna" each represented the power produced, measured in megawatt-hours (MWh), for the wind farm with the same name as the column.

5.2.3 The Use of the Data

We had to remove the excess data from the weather data set provided by Yr. This ranged from 21/12/2016 00:00 and 24/01/2018 23:00 and consisted of 9456 entries. The power data from NWE only ranged from 01/01/2017 01:00 to 01/01/2018 00:00, containing 8760 entries. After the data cut, all data sets ranged from 01/01/2017 01:00 to 31/12/2017 23:00. The weather data set from Yr were missing whole days of data, these days being 07/04, 06/06, 04/07 and 31/10. These had to be removed from the power data set from NWE. From the power data set, there was one hour missing: 26/03 02:00. This had to be removed from the weather data set.

After cleaning the data, we split the power data set into five different files, each file containing the power output for a single wind farm.

5.3 Methods for Predicting

In this chapter we will go through the models we have used for prediction, their settings and use. We will also compare the models. The models used for prediction are from the computational fluid dynamics field and the machine learning field. The computational fluid dynamics algorithm was provided by WindSim AS. The following models were used:

- Physical approach
  - WindSim CFD program
- Statistical approaches
  - Artificial Neural Network
  - K-Nearest Neighbors
  - Decision Trees
  - Random Forest Regression
  - Linear Regression
  - Support Vector Machine

Testing

When measuring the performance of prediction methods, we need to make sure the numbers are measured correctly and that they represent the true performance of the models. The first problem to think about is differences within the dataset. There may be parts of the dataset that are good, some that are bad, some that are easy to understand, some that are harder to understand etc.. If we train the model on a dataset that is easy to understand for the model, and test on a hard part of the dataset, the performance may not be that good. This has nothing to do with the capabilities and performance of the actual model, and it is only as good as the data we provide it with. Imagine that we build a model for predicting waves at the sea, and train it only with the summer months. If we test it on the winter months, and use that as our performance measure, it would probably not do well. The sea is possibly frozen too. Again, the model is only as good as our data. To assure that test results are representative, we have used the k-fold cross validation technique when testing our models.

In k-fold cross validation, the dataset is sliced into k slices. In the first run, the first slice is used for testing, and the rest is used for training. The model is trained and tested, and the test results are saved. For the next run, the second slice is picked as testing set, while the first slice is now a part of the training set with the rest of the slices. The model trains, this time testing on the second slice, and results are stored. The process is repeated until all slices have been tested on. The final result is the average of all runs [11].

Another issue in testing performance is non-deterministic models. These models always try to find the best solution, but their training is influenced by randomness in adjustment, thus the models are never identical, even when given the same data as before. A non-deterministic model can be the best performing model in one run, and the worst in the next. By training and testing the models a number of times, we can use the average performance of these runs as the performance of the model. Therefore, all the models have predicted each month once in each run, by using the k-fold cross validation technique with k = 12. The numbers presented are the mean average performance of the k-fold cross validation run 100 times. The variance in the mean average performance between two such runs becomes less than 1/1000.

Error Measurement

We use the Root Mean Squared Error as our error measure. For each prediction a method makes, y_i, there is a target value t_i. The difference between the target and prediction value, $(t_i - y_i)$, is the measure of error for that prediction. To calculate the error for the model as a whole, we sum up all these individual prediction errors. Since we only care about the magnitude of the error, and not the sign, we square the errors before summing. After summing the errors together, we divide the result by the number of predictions made, to get the mean average for each prediction. This is called the Mean Squared Error (MSE).

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (t_i - y_i)^2 \tag{5.1}$$

Since all the values are squared, they are much bigger than the values of the data we are working with. Therefore, we take the square root of the Mean Squared Error. This is called the Root Mean Squared Error (RMSE).

$$RMSE = \sqrt{MSE} \tag{5.2}$$

The RMSE error measurement is good for working with values that can be less or more than the target value. The downside is that it becomes hard to compare different solutions using RMSE, since the RMSE is proportional to the magnitudes of the values in the dataset.

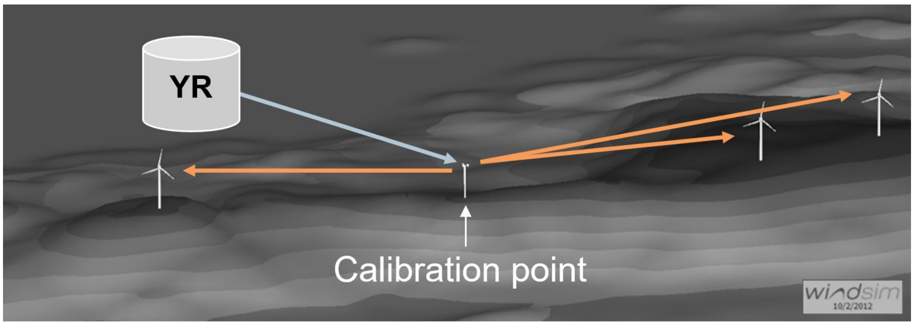

Figure 5.3: Wind conditions are given for the calibration point from Yr. From the calibration point to each turbine, CFD is used to simulate the flow of the wind. The orange arrows show the connection between the calibration point and the turbines. After the simulation, the relative change in wind conditions between the calibration point and a turbine can be calculated based on the results of the CFD simulation. [20]

5.3.1 Physical Approach: WindSim CFD

We have used CFD for simulating the wind from the Yr calibration point and to each of the turbines. The simulations have been done using the WindSim 8.0.0, which is WindSim AS' own wind simulating software. To run such a simulation for each of the different wind inputs would have been far too time consuming. A solution for this was that the WindSim team created a lookup table.

The lookup table consists of a sub-set of the numerical values produced after running the CFD. This simulation was run on certain wind conditions and would not be directly applicable for a scenario of different wind conditions. By calculating the relative change in wind conditions from the calibration point and to the turbines, they could create a lookup table. This lookup table was sent to us.

With the lookup table, we had the relative relationship in wind conditions between the calibration point and the wind turbines. New per turbine wind conditions and forecasted power output could therefore be calculated by running new wind conditions in a program using the lookup table. The output was three files: forecasted individual wind speeds for all turbines (5.4), forecasted individual wind degrees for all turbines (5.5) and individual forecasted power output for all turbines and total for the whole farm (5.6).

The snippets figure 5.4, figure 5.5 and figure 5.6 show the three different output files from the CFD program, run on Valsneset wind farm, which has five wind turbines. Notice the forecasted power output file, figure 5.6, which has one more column than the others, showing total forecasted power output for the whole farm.

Date/Time	P_001	P_002	P_003	P_004	P_005	
01/01/2017 01:00	10.8374	10.9425	10.8149	10.6212	9.5251	
01/01/2017 02:00	5.6364	5.68433	5.64341	5.54697	5.6241	
01/01/2017 03:00	3.27268	3.28	2.56546	3.2578	3.2521	

Figure 5.4: Snippet of the three first entries in the CFD forecasted wind speed output file for Valsneset wind farm.

Date/Time	P_001	P_002	P_003	P_004	P_005	
01/01/2017 01:00	34.8374	34.7597	34.8726	35.0026	34.9071	
01/01/2017 02:00	23.1058	23.1233	23.0603	23.1946	23.2148	
01/01/2017 03:00	350.015	350.18	349.778	349.905	350.079	

Figure 5.5: Snippet of the three first entries in the wind CFD forecasted degree output file for Valsneset wind farm.

Date/Time	P_001	P_002	P_003	P_004	P_005	P_ALL:
01/01/2017 01:00	1530.31	1568.91	1522.07	1450.96	1065.81	7138.07
01/01/2017 02:00	198.913	204.329	199.705	188.807	197.523	989.278
01/01/2017 03:00	28.3619	28.6398	0	27.7963	27.5797	112.378

Figure 5.6: Snippet of the three first entries in the CFD forecasted power output file for Valsneset wind farm.

CFD Power Prediction

The average error for the forecasted power production by the CFD alone can be seen in table 5.1. This is for the whole year of 2017. The average error rate shows significant differences in the prediction results between the wind farms. The two most similar farms based on production capacity are Bessaker and Hitra. Bessaker has max production 2MW higher than Hitra, but still has a lower RMSE. The error rate of Ytre Vikna is almost identical to Hitras, even though Ytre Vikna has a max production 15.53MW lower than Hitra. Following is a single day prediction for 01/12/17 for Bessaker (figure 5.7), Hitra (figure 5.8), Smøla (figure 5.9), Valsneset (figure 5.10) and Ytre Vikna (figure 5.11).

Wind farm	Max capacity in MW	RMSE	%
Bessaker	56.446	12.912	22.86%
Hitra	54.410	14.935	27.44%
Smøla	148.45	32.509	21.89%
Valsneset	11.71	2.925	24.97%
Ytre Vikna	38.880	14.911	38.35%

Table 5.1: CFD error for all wind farms. Prediction for whole 2017.

The predictions for Bessaker and Hitra are quite similar. The first 10 hours are very good. After this, they overestimate the rate of increasing power production, and have a steeper curve than actual power produced. They both also suddenly dive at around 17-18 hours.

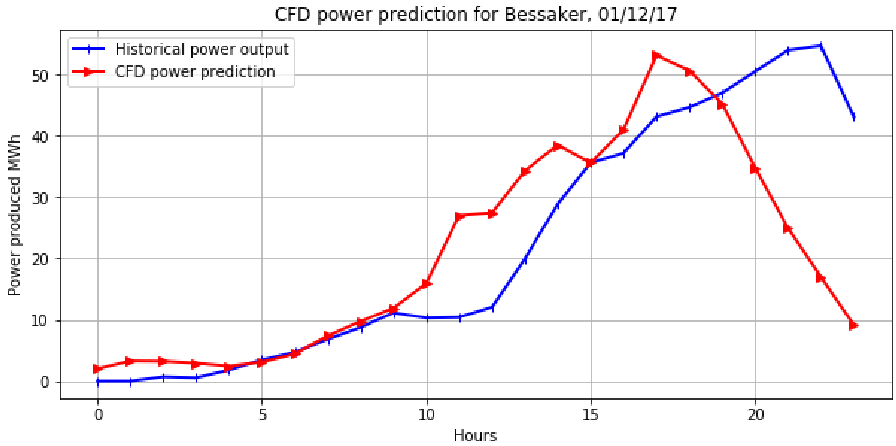

Figure 5.7: CFD single day prediction of power production for Bessaker

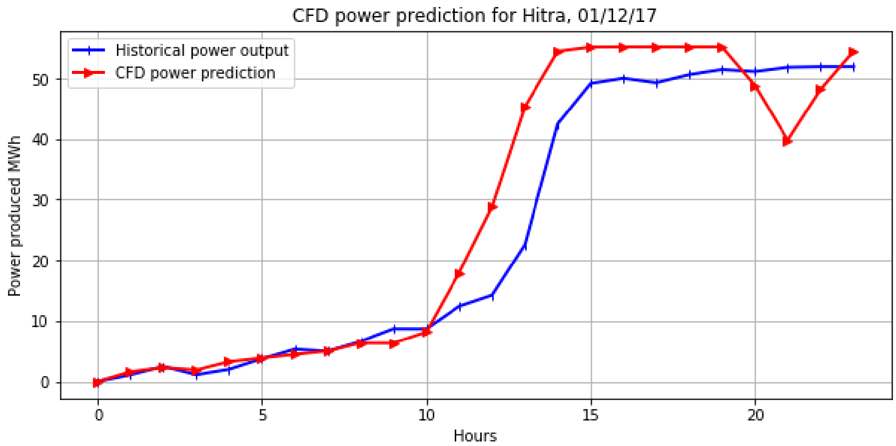

Figure 5.8: CFD single day prediction of production power for Hitra

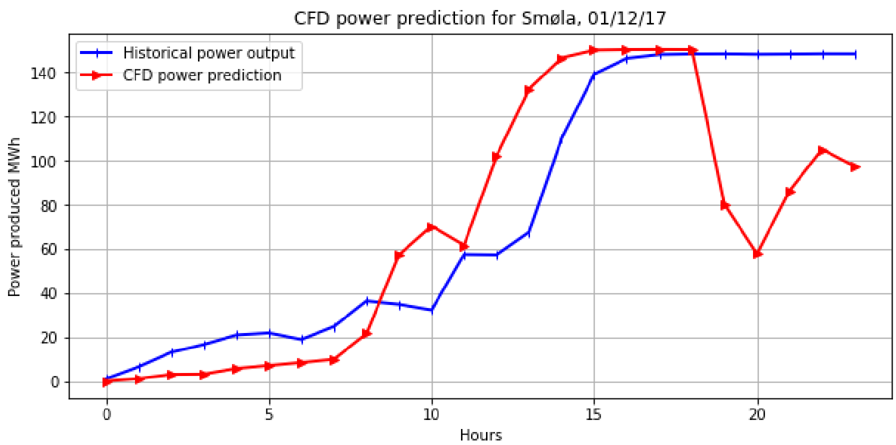

Figure 5.9: CFD single day prediction of production power for Smøla

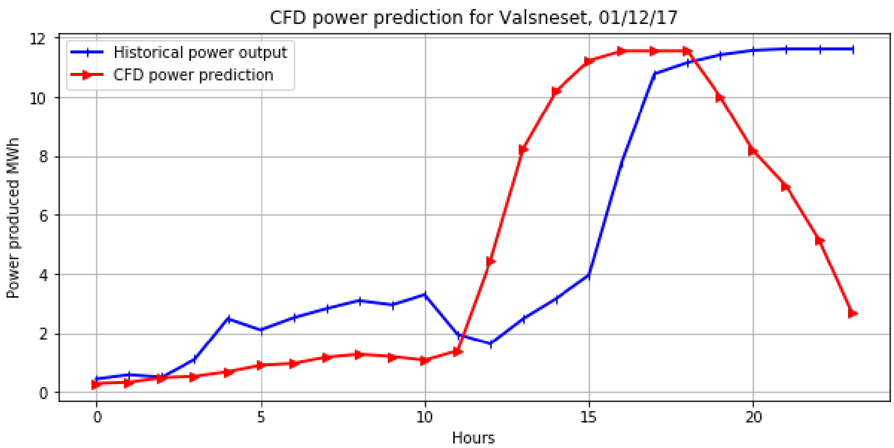

Figure 5.10: CFD single day prediction of production power for Valsneset

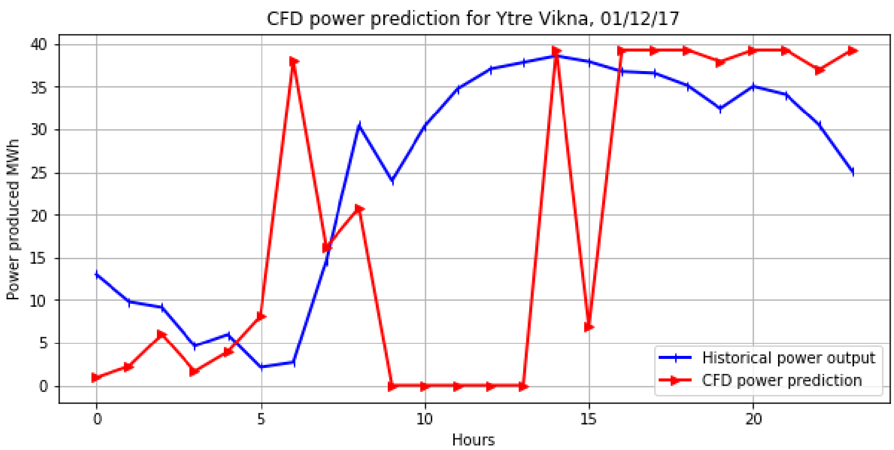

Figure 5.11: CFD single day prediction of production power for Ytre Vikna

At Smøla and Valsneset, the model clearly underestimates the power produced for the first 8-10 hours. After this, they overestimate the power produced for the next hours. When the historical power output flattens at the top, the predictions dive at 17-18 hours.

Ytre Viknas CFD power prediction for this day is a good example of the model failing to provide good estimates. The simulation does not appear to get a good understanding of the wind behavior in this case. The terrain may be complex making it difficult to simulate the wind, or the forecast data from Yr is deviating from the actual wind conditions.

5.3.2 Statistical Approach: Machine Learning Models

All the machine learning methods were provided by the Python package Scikit-Learn. All machine learning algorithm parameters was optimized by using the Scikit Gridsearch tool. The Gridsearch tool allows for the user to input ranges of different parameter settings, and the Gridsearch will run the model once per unique combination of the different settings. The error rate of each model run, using a unique combination of the parameter settings, is stored. After all runs, the error rates of the models are compared. The best performing parameter combination is presented.

Artificial Neural Network

The ANN consisted of an input layer, two hidden layers and an output layer. The input layer consisted of three neurons, one for each input value. The first hidden layer consisted of 250 neurons, the second of 180 neurons. The output layer consisted of a single neuron. The ouput was the estimated power production in kWh. The activation function for the neurons was the logistic Sigmoid function $f(x) = \frac{1}{1 + exp(-x)}$. The weight optimization was done with the Adam optimizer. Batch-sizes were set to 40, and learning rate was set to 0.0001.

K-nearest Neighbor Regression

The KNN model used the 22 nearest neighbors when predicting the outcome (k=22). All neighbors contributed equally to the outcome of the prediction. The method for calculating the distance to each neighbor was done using the binary tree search inspired method KD-tree.

Decision Tree Regression

The strategy for choosing the data splits was based on best split strategy. This costs more computation time than the using the random best split, which offer some random splits and chooses the of these, but ensures better splits by checking for the best available split at each point. When adjusting the splits afterwards, the model used the Mean Squared Error.

Random Forest Regression

The number of decision trees in our random forest regressor was set to 45. The criterion for the sum-of-squares error optimization were based on the MSE. The max depth of the individual trees was set to 4.

Linear Regression

The linear regressions only parameter was if the model calculates the intercepts in the data to use in its calculations. This performed best when set to true. No regularization was applied.

Support Vector Regression

The SVR's important parameter is which kernel type to be used in the algorithm. The kernel is the core of the calculations when fitting a SVR. The kernel which performed best for my data was the radial basis kernel (RBF). The penalty parameter for the error function for adjusting the tube, C, was set to 15.0. The ϵ had to be set manually for each farm. This is because epsilon is based on the value of what we are predicting.

Wind farm	ϵ
Bessaker	5.0
Hitra	5.6
Smøla	10.0
Valsneset	1.4
Ytre Vikna	3.5

Table 5.2: The SVR ϵ parameter for each wind farm.

5.3.3 Machine Learning Prediction Results

Here we present the performance results for each machine learning model for each site, and forecast a day-ahead power production.

Bessaker Wind Farm

From the statistical prediction methods, the artificial neural network is the best one for Bessaker wind farm, as shown in table 5.4. The SVR is a close second, and the random forest regression is the third best. When we look at the single day prediction for Bessaker in figure 5.12, we see that all the models perform quite good. The ANN, shown in a blue line, performs best.

Machine learning model	RMSE	%
Artificial Neural Network	9.824	17.39%
K-Nearest Neighbors	11.208	19.85%
Decision tree	9.988	17.69%
Random Forest Regression	9.871	17.48%
Linear Regression	10.545	18.68%
Support Vector Regression	9.842	17.43%

Table 5.3: Bessaker machine learning wind power prediction performance. Max capacity is 56.446 MW

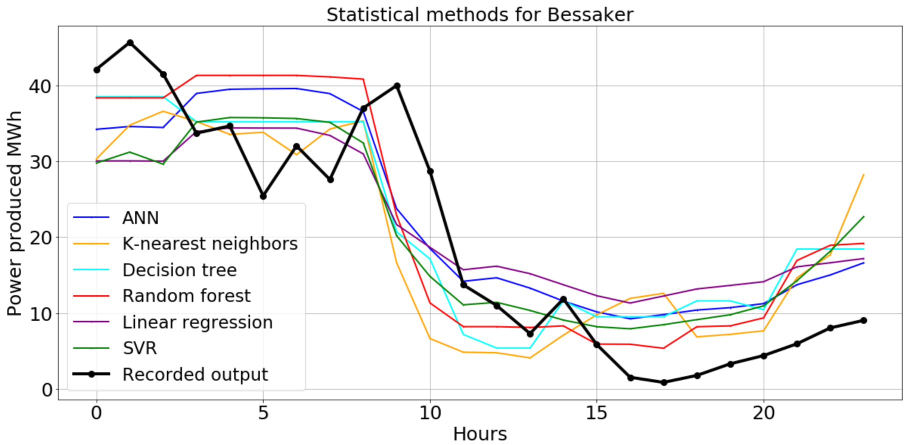

Figure 5.12: Machine learning single day prediction of power production for Bessaker.

Hitra Wind Farm

The best model for Hitra wind farm is the ANN. The random forest is the second best, with the SVR in third place, all seen in table 5.4. For the single day power prediction in Hitra wind farm in figure 5.13, the models' prediction curves are similar. The first 10 hours, with a declining power production, we see the predictions have some distance from the recorded output. Once it flattens, the predictions follow. The ANN in blue is again the best predictor.

Machine learning model	RMSE	%
Artificial Neural Network	11.488	21.11%
K-Nearest Neighbors	12.753	23.43%
Decision tree	11.802	21.69%
Random Forest Regression	11.593	21.30%
Linear Regression	12.051	22.14%
Support Vector Regression	11.706	21.51%

Table 5.4: Hitra machine learning wind power prediction performance. Max capacity is 54.410 MW

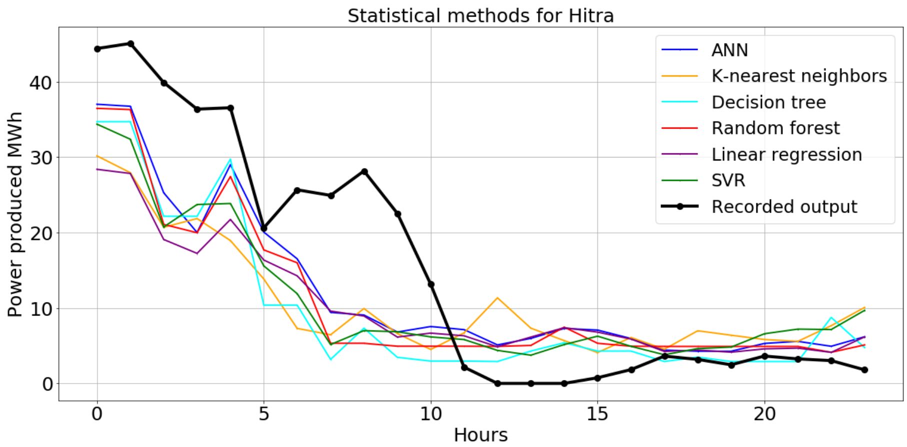

Figure 5.13: Machine learning single day prediction of production power for Hitra.

Smøla Wind Farm

For the biggest wind farm in this project, Smøla, the random forest performs best of the statistical approaches. The second best is the decision tree. The third best was the ANN. The start of the predictions in figure 5.14 are very similar, but they start to spread after the first five hours. They all follow the behavior of the recorded output. After 20 hours, all models overestimate the production slightly at the end.

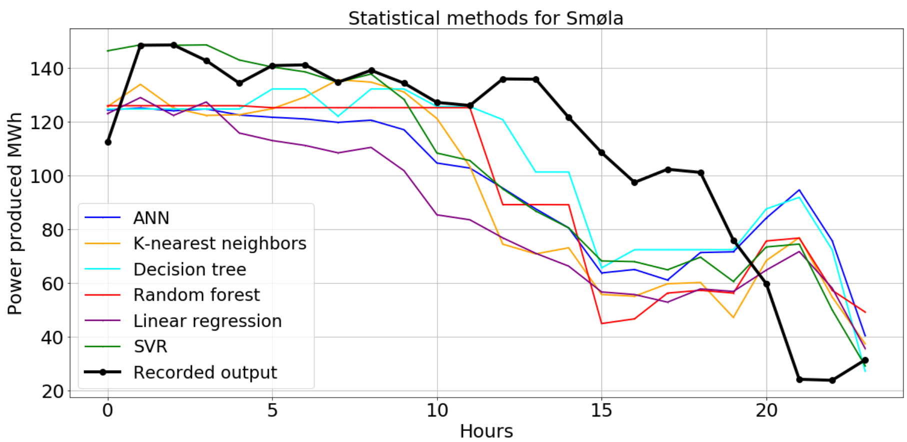

Figure 5.14: Machine learning single day prediction of production power for Smøla

Machine learning model	RMSE	%
Artificial Neural Network	28.206	19.00%
K-Nearest Neighbors	30.676	20.66%
Decision tree	28.102	18.93%
Random Forest Regression	28.030	18.88%
Linear Regression	29.281	19.72%
Support Vector Regression	28.681	19.32%

Table 5.5: Smøla machine learning wind power prediction performance. Max capacity is 148.450 MW.

Valsneset Wind Farm

The smallest wind farm, Valsneset, has the ANN as the top performing statistical model. After this comes the random forest, and third best was the SVR, all shown in table 5.6.

Machine learning model	RMSE	%
Artificial Neural Network	2.453	20.94%
K-Nearest Neighbors	2.705	23.05%
Decision tree	2.547	21.75%
Random Forest Regression	2.466	21.05%
Linear Regression	2.567	21.92%
Support Vector Regression	2.476	21.14%

Table 5.6: Valsneset machine learning wind power prediction performance. Max capacity is 11.71 MW.

.

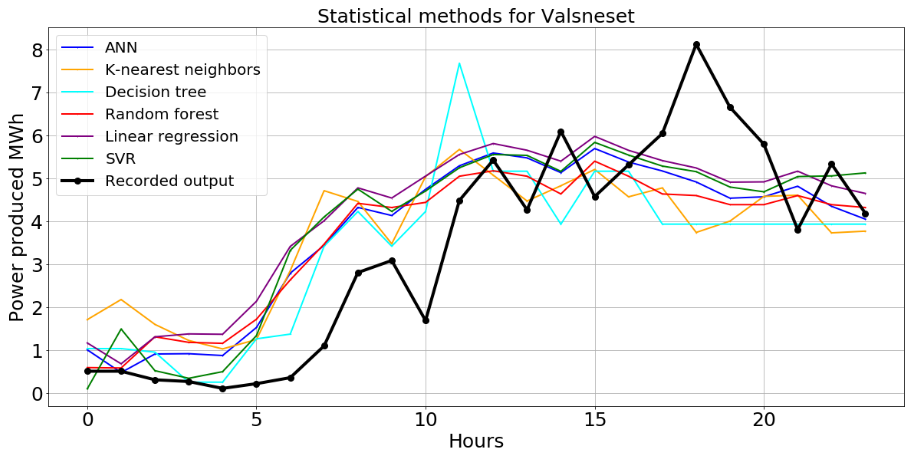

Figure 5.15: Machine learning single day prediction of production power for Valsneset

The single day prediction for Valsneset show a more varying curve than the other farms, but still the models follow the general behavior of the recorded power output well. None of the models capture the spike of power production at around 18 hours. We think this displays the volatility of wind well. The ANN shown in blue is the best performing model on this day 5.15.

Ytre Vikna Wind Farm

For the last wind farm, Ytre Vikna, the SVR was the best statistical model. This was closely followed by the random forest and the ANN, as seen in 5.7. The day ahead power predictions in Ytre Vikna show that the models follow the recorded output during the increase in production. The sudden spike at 15 hours is not predicted, but the steep rise at 20 hours is. The best model is the SVR, as shown in cyan in figure 5.16.

Machine learning model	RMSE	%
Artificial Neural Network	7.400	19.03%
K-Nearest Neighbors	8.251	21.22%
Decision tree	7.426	19.09%
Random Forest Regression	7.375	18.96%
Linear Regression	7.761	19.96%
Support Vector Regression	7.248	18.64%

Table 5.7: Ytre Vikna machine learning wind power prediction performance. Mac capacity is 38.88 MW

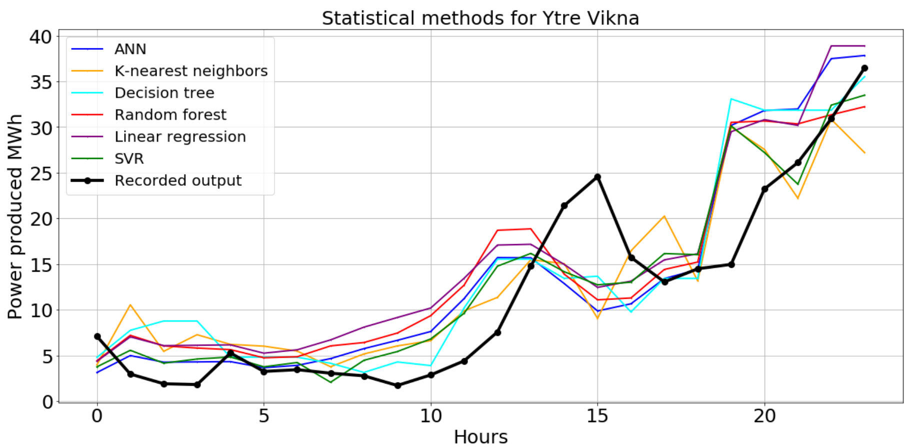

Figure 5.16: Machine learning single day prediction of production power for Ytre Vikna

5.4 The Hybrid Approach Setup

The hybrid program was set up by running the CFD program on wind conditions from Yr, and using the wind speed and wind degree simulation prediction from the CFD as input to a machine learning model.

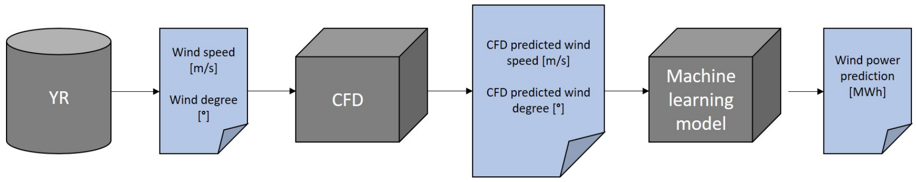

Figure 5.17: Setup for hybrid of physical and statistical wind power forecast approach

5.4.1 Physical Approach: CFD

The CFD used only wind speed and wind degree as input. It was used to simulate the difference in wind between the calibration point (in the middle of the farm) and each individual wind turbine in the farm. With these simulated, per-turbine wind conditions, the CFD would estimate the power production per turbine and for the whole farm. The output which the CFD algorithm produced were three files: forecasted power, forecasted wind speed and forecasted wind direction. The forecasted power was the estimated power production for each individual turbine, and for the whole farm. The forecasted wind speed was the estimated wind speed for each turbine, and forecasted wind direction was the estimated wind degree for each individual turbine. All files gave a per hour estimation.

5.4.2 Between the Physical and Statistical Approaches

We estimated wind speed and wind degree for the whole farm based on the per-turbine forecasts. This was done by calculating the average of all the per-turbine forecasts. Now we had wind speed estimations and wind degree estimations for the whole farm, based on simulations.

5.4.3 Statistical Approach: Machine Learning Model

The machine learning models used the CFD wind speed and wind direction estimate whole-farm averages as input. The machine learning models were also given the power production data from NWE. The output of all the models were hourly forecasts of power production for the wind farm.

Date/Time	P_001	P_002	P_003	P_004	P_005	P_ALL
01/01/2017 01:00	34.8374248	34.759718	34.872598	35.002594	34.90707	34.87588
01/01/2017 02:00	23.1058104	23.123259	23.060304	23.194639	23.21484	23.13977
01/01/2017 03:00	350.014983	350.1796	349.77779	349.90457	350.0791	349.9912
01/01/2017 04:00	6.07189783	6.2140412	5.8855275	6.040179	6.202299	6.082789
01/01/2017 05:00	31.9081881	31.859874	31.93314	32.058355	32.00265	31.95244
01/01/2017 06:00	26.4517575	26.444712	26.43392	26.564265	26.55659	26.49025
01/01/2017 07:00	26.2489728	26.243412	26.229459	26.360045	26.35406	26.28719
01/01/2017 08:00	7.38999821	7.5224925	7.214528	7.3676074	7.518745	7.402674
01/01/2017 09:00	358.804821	358.98115	358.57475	358.72866	358.927	358.8033
01/01/2017 10:00	10.2289836	10.340695	10.076991	10.226684	10.35417	10.2455

Figure 5.18: Wind direction file after mean average of wind direction for whole farm

Date/Time	P_001	P_002	P_003	P_004	P_005	P_ALL
01/01/2017 01:00	10.83737	10.94254	10.81491	10.62116	9.525097	10.54822
01/01/2017 02:00	5.636396	5.684331	5.643411	5.546967	5.6241	5.627041
01/01/2017 03:00	3.272682	3.279996	2.565462	3.257799	3.252096	3.125607
01/01/2017 04:00	16.50257	16.5824	16.59482	16.36978	16.42647	16.49521
01/01/2017 05:00	10.41836	10.52336	10.40908	10.21584	10.40465	10.39426
01/01/2017 06:00	18.27378	18.44142	18.28202	17.95775	18.24202	18.2394
01/01/2017 07:00	19.16752	19.34259	19.17708	18.83768	19.1337	19.13171
01/01/2017 08:00	19.8519	19.95389	19.95581	19.67944	19.7643	19.84107
01/01/2017 09:00	17.33452	17.39298	16.47831	16.73051	17.23791	17.03485
01/01/2017 10:00	15.16793	15.25552	15.23588	15.01557	15.1074	15.15646

Figure 5.19: Wind speed file after mean average of wind speed for whole farm

Chapter 6

Results

The overall performance of the machine learning models were generated using eleven months of training data, and predicting the remaining month. This was achieved using the k-fold cross validation technique.

6.1 Bessaker Wind Farm

The maximum recorded power output in a single hour for Bessaker wind farm in 2017 was 56.446 MWh. Prediction values above 56.446 were cut to 56.446 and negative prediction values were set to 0. The prediction error rate is presented in table 6.1 in column "Hybrid", measured in RMSE.

Hybrid Performance

The best performing hybrid model for Bessaker was the SVR, while the worst performing hybrid model was the linear regressor. K-nearest neighbors is the second worst performing model of the hybrid approaches. The bagging technique in random forest appears to work well, considering the decision tree model had a higher error rate. Still, they all had quite similar performance, at around 10 RMSE.

Prediction model	Pure ML	Hybrid	Improvement
Artificial Neural Network	9.824	10.037	-2.16%
K-Nearest Neighbors	11.208	10.354	8.20%
Decision tree	9.988	10.149	-1.61%
Random Forest Regression	9.871	9.892	-0.21%
Linear Regression	10.545	10.963	-3.96%
Support Vector Regression	9.842	**9.788**	0.55%
CFD	12.912	-	-

Table 6.1: Bessaker wind farm. Wind power prediction models and their error rate, given in RMSE. Best performance in bold numbers. Improvement is how much the solutions average RMSE improved going from pure machine learning to hybrid in percentage (%).

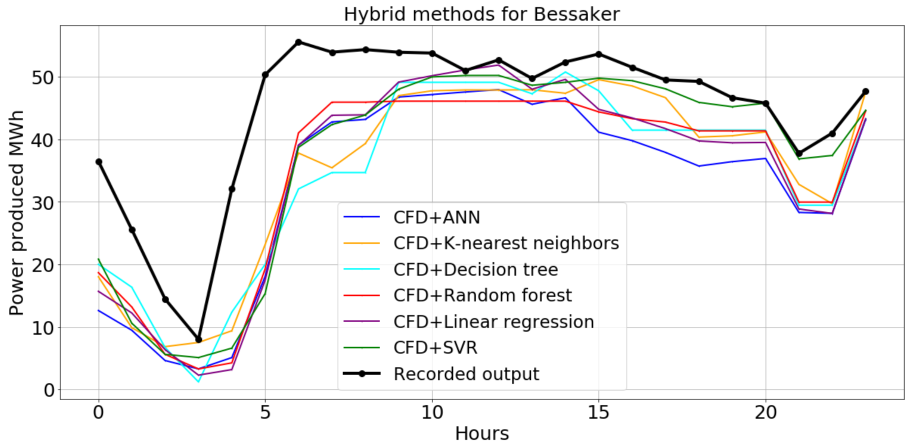

Figure 6.1: The variance in error rate between the models is low. The best model is the SVR, in green

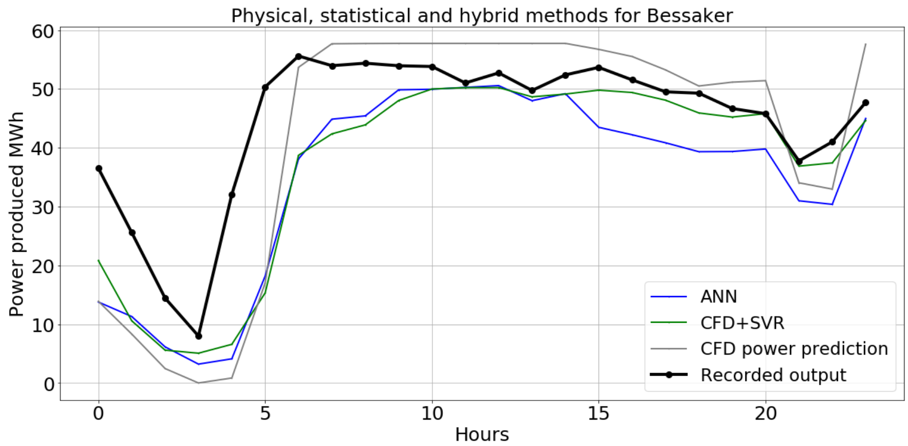

Figure 6.2: Performance of CFD, ANN, and the CFD + SVR hybrid model. The hybrid model outperforms both.

Overall Performance

Table 6.1 presents the performance of the physical model, statistical models and the hybrid models. The worst model for power prediction was the purely physical approach, the CFD model, with around 1.7 RMSE more than the second worst. Next, we compare the hybrid physical and statistical approaches with the purely statistical ones. For most of the models, adding the CFD before the statistical models had a negative impact on prediction results. Hybrid models which used ANN, decision trees, random forest regression and linear regression all performed worse than the purely statistical models. Only the hybrid solutions using k-nearest neighbors and support vector regression increase in performance compared to the purely statistical approach. Note that the two best performing statistical models were ANN with RMSE 9.819 and SVR with RMSE 9.850. When included in a hybrid solution, we see that the ANN performance decreases, and the RMSE rises to 10.036. For the hybrid solution using SVR, the performance improves, and the RMSE is reduced to 9.797. This makes the hybrid solution using SVR the best wind power prediction model for Bessaker wind farm.

6.2 Hitra Wind Farm

The highest recorded power production in a single hour in 2017 for Hitra wind farm was 54.652 MWh. All prediction values above 54.652 were adjusted to 54.652 and negative prediction values set to 0. The prediction error rate is presented in table 6.2 "Hybrid" column, measured in RMSE.

Prediction model	Standalone	Hybrid	Improvement
Artificial Neural Network	11.488	11.151	3.02%
K-Nearest Neighbors	12.753	11.611	9.83%
Decision tree	11.802	11.049	6.81%
Random Forest Regression	11.593	**10.884**	6.51%
Linear Regression	12.051	11.786	2.24%
Support Vector Regression	11.706	10.898	7.41%
CFD	14.935	-	-

Table 6.2: Hitra wind farm. Wind power prediction models and their error rate, given in RMSE. Best performance in bold numbers. Improvement is how much the solutions average RMSE improved going from standalone to hybrid in percentage (%).

Hybrid Performance

For the Hitra wind farm, the best hybrid prediction model was the random forest. The hybrid SVR had a slightly higher error rate, making it the second best hybrid model. In the day-ahead prediction, the models are less similar in their predictions, but still follow the same general pattern.

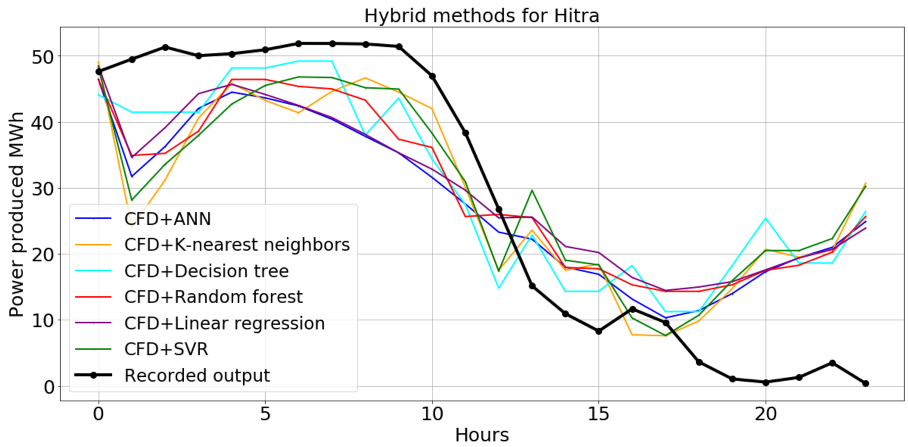

Figure 6.3: Hybrid approach performance for Hitra

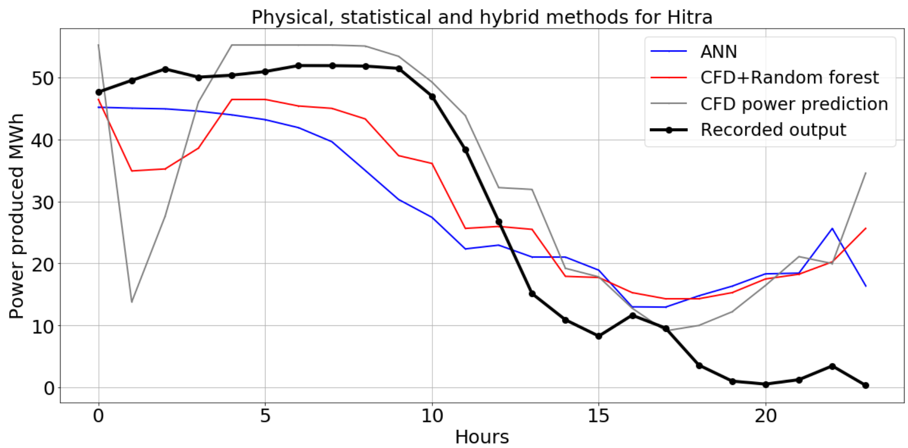

Figure 6.4: The CFD, the ANN, and the CFD + SVR hybrid in the day-ahead prediction. Except the CFD's dip at 3 hours, the hybrid model stays between the physical and statistical predictions, for the first 16 hours. The hybrid approach was the best approach for Hitra.

None of them follow the recorded output out to the "hill" at 10 hours, instead cutting down to the end of the slope at 15 hours. After this, the models predict a rise in power production, which never occurs. The best performing hybrid in figure 6.3 is the random forest.

Overall Performance

The CFD had the overall worst performance. Comparing the performance of the statistical and hybrid models, we see that all models improve their performance when moving from the statistical to the hybrid model. Therefore, the best performing overall model was also the best performing hybrid model, the random forest. However, the differences between the models are small.

6.3 Smøla Wind Farm

For Smøla wind farm, the highest power production in a single hour was 148.45 MWh. Prediction values above 148.45 has been adjusted to 148.45 and negative values have been set to 0. The models' prediction error rate is presented in table 6.3 "Hybrid" column, measured in RMSE.

Prediction model	Standalone	Hybrid	Improvement
Artificial Neural Network	28.206	28.288	-0.29%
K-Nearest Neighbors	30.676	29.597	3.64%
Decision tree	28.102	28.236	-0.47%
Random Forest Regression	28.030	**27.941**	0.31%
Linear Regression	29.281	29.271	0.03%
Support Vector Regression	28.681	28.025	2.34%
CFD	32.509	-	-

Table 6.3: Smøla wind farm. Wind power prediction models and their error rate, given in RMSE. Best performance in bold numbers. Improvement is how much the solutions average RMSE improved going from standalone to hybrid in percentage (%).

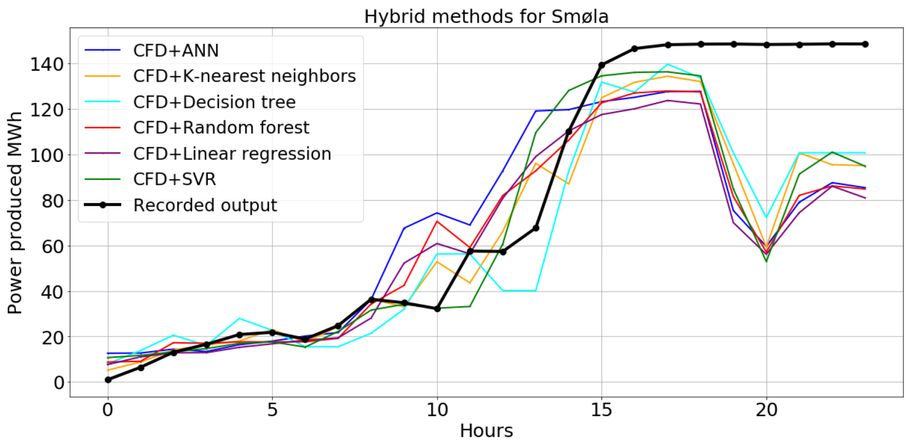

Figure 6.5: Hybrid approach performance for Smøla

Hybrid Performance

Smøla is the largest farm of them all, and has the largest production. The best hybrid model was the random forest, followed by SVR. ANN is third best, and k-nearest neighbors has the worst performance. The hybrid models' performance for a day-ahead prediction in Smøla wind farm is shown in figure 6.5. All models start out well, but after 10 hours they disperse. All models follow the general behavior of the recorded output, until around 18 hours, where there is a sudden steep slope in the curves for all the models. Errors in the Yr wind forecast may be the cause of the discrepancy. The random forest in green is the best performing model.

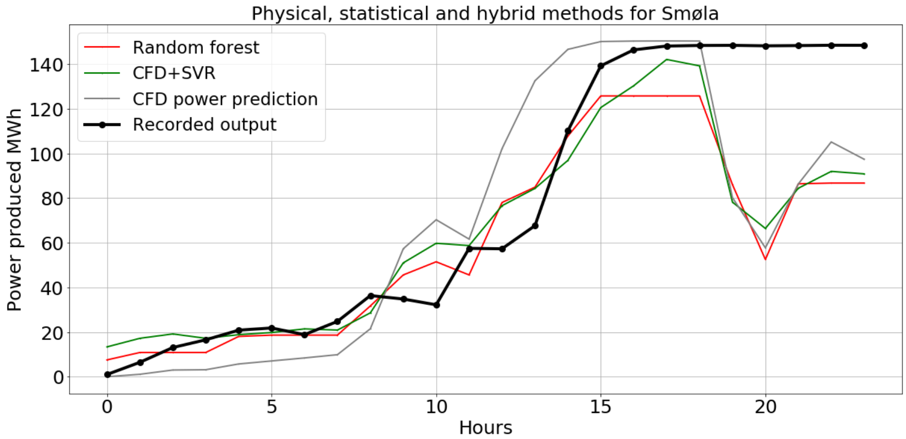

Figure 6.6: The CFD, the random forest, and the CFD + random forest hybrid, compared in a day-ahead prediction. Both of the random forests perform well, but the hybrid is better. CFD is the worst, and the hybrid approach is the best.

Overall Performance

Artificial neural network and decision tree were the only two models which did not improve their performance between being a standalone and as part of the hybrid solution. The rest had a positive change, and for support vector regression, it went from being the fourth best standalone model to become the second best hybrid model. Random forest regression is clearly the best, having the hybrid model as the overall best model, and the standalone model as the overall second best model. The CFD model again, having no good performance in the wind power prediction task.

6.4 Valsneset Wind Farm

Valsneset wind farm's highest power production hour was 11.71 MWh. Prediction values above 11.71 were adjusted to 11.71, and negative values were set to 0. The models' prediction error rate is presented in table 6.4 "Hybrid" column, measured in RMSE.

Hybrid Performance

The best hybrid solution for Valsneset was the random forest. Second best was the ANN, followed by the decision tree. The SVR is the worst by some margin. The predictions for this wind farm fluctuate more than for the others. The models vary more in their predictions, although they converge at some points like 4 and 16 hours. None of the models follow the produced power when it increases between 10 and 20 hours. The best performing hybrid model was the hybrid random forest.

Prediction model	Standalone	Hybrid	Improvement
Artificial Neural Network	**2.453**	2.477	-0.97%
K-Nearest Neighbors	2.705	2.608	3.71%
Decision tree	2.547	2.550	-0.11%
Random Forest Regression	2.466	2.476	-0.40%
Linear Regression	2.567	2.606	-1.51%
Support Vector Regression	2.476	2.483	-0.28%
CFD	2.925	-	-

Table 6.4: Valsneset wind farm. Wind power prediction models and their error rate, given in RMSE. Best performance in bold numbers. Improvement is how much the solutions average RMSE improved going from standalone to hybrid in percentage (%).

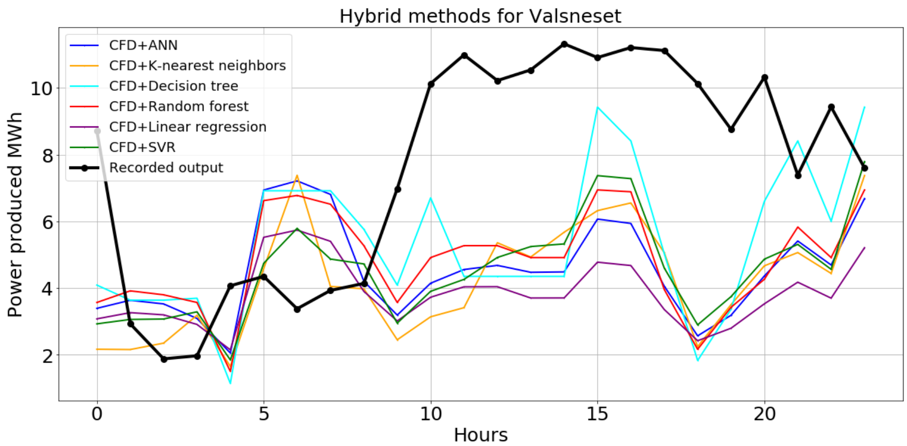

Figure 6.7: Hybrid approach performance for Valsneset

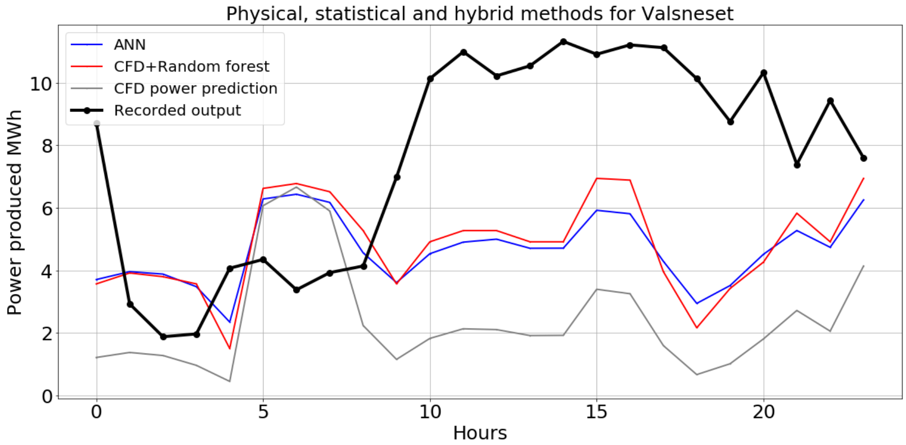

Figure 6.8: The CFD, ANN, and CFD + random forest hybrid, compared in the day-ahead prediction. Valsneset is the only wind farm where a statistical model is the best overall model for the farm. The pure ANN predicts slightly higher values than the hybrid solution, but all models underestimate. This makes the pure ANN the best performing model in the day-ahead prediction for this day, and is also the best overall model for Valsneset.

Overall Performance

The k-nearest neighbor was the only statistical model which improved its performance when combined with the CFD. The remaining models performed worse as part of a hybrid model than as a purely statistical model. High terrain complexity may have contributed to the CFD's error. Valsneset was the only wind farm in this project where the best performing solution was not a hybrid. Pure ANN was the overall best performing model for this farm, closely followed by the pure random forest.

6.5 Ytre Vikna Wind Farm

Ytre Vikna wind farm's highest power production hour was 38.88 MWh. Prediction values above 38.88 were adjusted to 38.88, and negative values were set to 0. The models' prediction error rate is presented in table 6.5 "Hybrid" column, measured in RMSE.

Hybrid Performance

The SVR was the best performing model of all the hybrid models on Ytre Vikna wind farm. K-nearest neighbors was the second best, and decision tree third. The hybrid models day-ahead prediction, as illustrated in figure 6.9, shows that the models predict very well for this day. They follow the recorded output all the way up, as well as the dive at the end. The best performing hybrid model for Ytre Vikna was the hybrid SVR.

Prediction model	Standalone	Hybrid	Improvement
Artificial Neural Network	7.400	8.128	-9.83%
K-Nearest Neighbors	8.251	7.537	9.47%
Decision tree	7.426	7.576	-2.01%
Random Forest Regression	7.375	8.021	-8.75%
Linear Regression	7.761	9.429	-21.49%
Support Vector Regression	7.248	**7.129**	1.14%
CFD	14.911	-	-

Table 6.5: Ytre Vikna wind farm. Wind power prediction models and their error rate, given in RMSE. Best performance in bold numbers. Percentage (%) improvement in RMSE going from standalone to hybrid model.

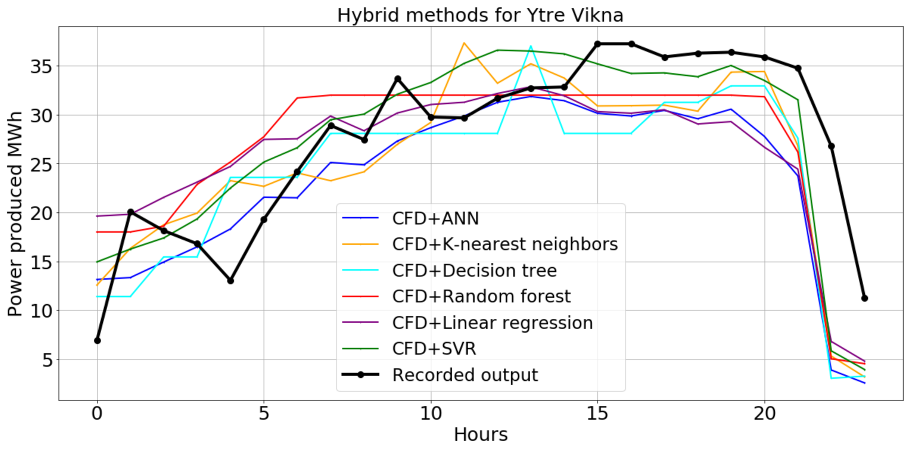

Figure 6.9: Hybrid approach performance for Ytre Vikna

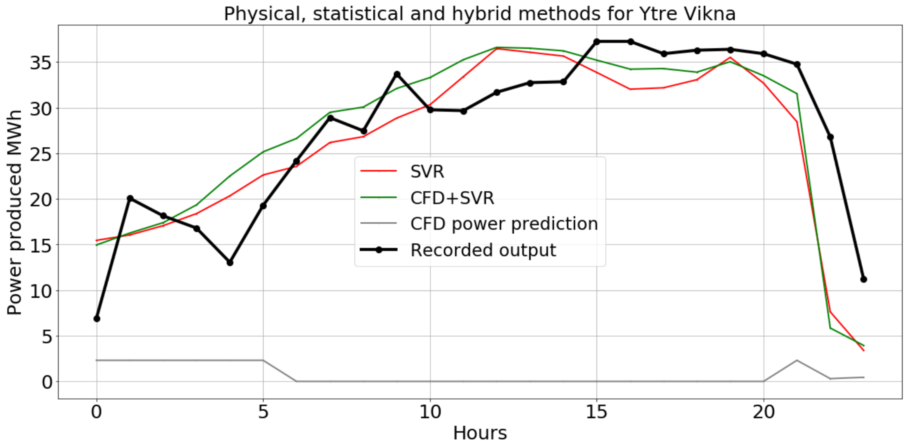

Figure 6.10: The CFD, the random forest, and CFD + SVR hybrid, compared in the day-ahead prediction. The CFD did not perform well, predicting almost zero production for this day. The hybrid SVR is the best overall model for Ytre Vikna.

Overall Performance

The only two models which improved their performance in the hybrid solution was k-nearest neighbor and SVR. The differences in performance between the purely statistical and the hybrid models were bigger for Ytre Vikna than for the other wind farms. The difference in the performance between purely statistical and hybrid model for the linear regressor was the highest of any model on any farm, with a -21.49% difference. CFD is the overall worst wind power predictor, and SVR hybrid model was the overall best performing model.

6.5.1 Overall Results for all the Wind Farms

Considering the performance results across all wind farms, we see that a hybrid approach was the best performing approach on four of the five different wind farms, only Valsneset did not improve its performance. The four best hybrid models consisted of two hybrid SVR models, and two hybrid random forests. The best models are hybrid SVR and hybrid random forest. In table 6.7, we display the difference in performance when applying the CFD to the pure statistical models.

Wind farm	Best model
Bessaker	Hybrid support vector regression
Hitra	Hybrid random forest regression
Smøla	Hybrid random forest regression
Valsneset	Pure artificial neural network
Ytre Vikna	Hybrid support vector regression

Table 6.6: The wind farms and the best performing model.

Model	Bessaker	Hitra	Smøla	Valsneset	Ytre Vikna	Avg.
ANN	-2.16%	3.02%	-0.29%	-0.97%	-9.83%	-2.04%
KNN	8.20%	9.83%	3.64%	3.71%	9.47%	6.97%
DT	-1.61%	6.81%	-0.47%	-0.11%	-2.01%	0.52%
RFR	-0.21%	6.51%	0.31%	-0.40%	-8.75%	-0.50%
LR	-3.96%	2.24%	0.03%	-1.51%	-21.49%	-4.93%
SVR	0.55%	7.41%	2.34%	-0.28%	1.14%	2.68%
Avg.	0.13%	5.97%	0.92%	0.07%	-5.24 %	-

Table 6.7: The percentage (%) improvement going from standalone to hybrid solution, for all models from all the wind farms. -% means the hybrid solution was less good than the standalone. Avg. is the mean average improvement for that rows model.

The hybrid solution did not uniformly improve or decrease the performance of the purely statistical models. Looking at the average improvement for the farms, Hitra appears to respond positively to the hybrid solution. For this farm, all hybrid models performed better than the pure statistical models, and the average improvement was almost 6%. It appears that the CFD was able to provide precise forecasts for this farm. At Ytre Vikna, we see the opposite. Ytre Vikna does have a positive improvement when using the hybrid KNN and hybrid SVR, but the average improvement was -5.24%. The hybrid linear regression had a significant influence on the mean average for this farm, with an improvement of -21.49%. At Valsneset, five out of six models perform worse as part of a hybrid solution. The average improvement at Valsneset is positive, 0.07%, but this is due to the k-nearest neighbors model which was the only model showing improvement.

Considering the models individually, SVR improves at most farms, except for at Valseneset, and has a average improvement of 2.68%. Since the pure SVR provided consistently good predictions, this is a significant improvement. Linear regression is the model where hybrid has the worst impact on average.

An interesting finding is how k-nearest neighbors improves on all farms when going from purely statistical to a hybrid solution, with an average improvement of 6.97%. This is significant due to the simplicity of k-nearest neighbors. If the closest surrounding datapoints to a value we want to predict, has become on average 6.97% more correct than before, that can mean that for all farms, wind speed and direction correlates better with the power produced for each farm than it did before. For all farms, the wind condition forecasts from CFD are therefore more correct than the Yr forecasts. Due to the lack of historical wind measurements from the farms, this is difficult to verify.

Chapter 7

Discussion

In this chapter we discuss the different aspects of the project that we believe have had an impact on our results. Then we give a final conclusion to our work.

7.1 The Data

The biggest source of uncertainty in our project was the wind forecast data from Yr. This data consisted not of historical recorded wind conditions on or near the wind farms, but rather forecasts. Finding the relationship between the produced power and the wind conditions may have been challenging for the models, since the accuracy of the forecasts may have varied greatly between hours. Hence, an unknown part of our prediction errors can be contributed to weather forecast errors. Further, the wind condition data consists of only two parameters: wind speed and wind degree. This may not be rich enough for the algorithms to model the complexity and volatility of wind, and tie this to power production. Factors like pressure in the atmosphere and the temperature in the area both influence the behavior of the wind. The weather conditions may also be a factor for the mechanical wind turbine, where some conditions are better suited for mechanical performance than others.

7.1.1 The Timescale of the Data

The weather data from Yr shows an estimation of wind speed and wind degree once every hour. At this resolution, the variability of wind within a single hour is lost. If the forecast predicts wind of 5 m/s at 20:00, the wind may change at 20:05 to 10 m/s and stay at this speed for the rest of the hour. The power production data from NWE display the combined power produced throughout the whole hour, while the wind condition data only show the wind conditions at time of measurement. If the wind were to change much within the hour, the power produced will change accordingly, while no new wind condition forecast will be given until the next measurement. There is a possibility that the forecasted wind conditions at the forecast point for a given hour does not represent the wind conditions of the whole hour very well. The relationship between the forecasted wind conditions and the produced power may have been influenced by this.

7.2 Tuning of Machine Learning Models

The tuning of the hyper parameters in the machine learning models could have been further adjusted, to reach even more accurate predictions. The time scope of this project did not allow for individually tuned models for each wind farm, since this would be too time consuming. The models' predictions may have suffered from this.

7.3 Conclusion

In this thesis, we have looked at the combination of the statistical and the physical wind power prediction approach. This was done using computational fluid dynamics for physical simulation of the wind from wind forecast location and to the wind turbines, and machine learning models for statistical prediction of wind farm power output using the simulated wind. The initial wind conditions used for the physical simulation were forecasts, and not historical recordings. This was tested on five different Norwegian wind farms, using six different hybrid solutions. From these experiments, we found that in four out of five wind farms, the hybrid solution outperformed the pure statistical and pure physical approaches. The two best hybrid solutions were computational fluid dynamics and random forest regression, and computational fluid dynamics and support vector regression. With these findings, we can conclude that a hybrid solution is likely a promising approach for some, but not necessarily all, wind farms.

Chapter 8

Future Work

In this chapter, we will address some approaches or improvements which future research could look at. The topics presented was not included or did not fit the scope of my project, but has potential to contribute to a better understanding of this subject.

8.1 Improvements in Data

In our project, the historical wind data consisted of Yr wind forecasts for the local area, and not actual measurements. While working on this project, we found weather stations not far from some of the wind farms, which record wind data. These data could be used instead. Historical recorded weather data may provide more realistic behavior of the wind, and could contribute to better understanding the relationship between the wind conditions and power production for that local area. This removes systematic wind forecasting errors as sources of error in the models.

The amount of data for our project was a year worth of hourly forecasts. Significantly, this covers many days and nights, which have a change in temperature. This does not cover the cycles of the seasons, which have a significant impact of the temperature in the areas included. The wintertime also include other challenges such as icing on the blades and equipment of the turbine. If provided with many years of data, the season cycle would repeat itself, and the models could include this in their predictions.

The wind forecasts in our project consisted of two parameters, wind speed and wind degree. With a richer data set, these predictions may become more accurate. While looking up the local areas in my project, we found that Yr also provide temperature, humidity, pressure, rainfall and other weather measurements on an hourly basis. There are also other sources of data which are possible to include, such as mechanical and technical measurements. Turbine blade angle and turbine torque are examples, as used by Kusiak and Zhang [29].

8.1.1 Imbalance in the Power Grid

Much of the reward from wind power production forecast lies in maintaining balance in the power grid. The prediction of imbalance in the power grid could be an interesting approach for future work to consider. This certainly has value for the transmission supply operator, since this can decrease imbalance-related issues.

8.1.2 Pricing of Power

Another useful idea is the prediction of power pricing. Knowing the power price in advance gives benefits planning a bidding strategy on the market. This can result in cheaper power, which again gives the consumer cheaper prices. In many countries, renewable sources in the energy mix are many, and most have an unpredictable nature. By predicting the influx of renewable power in the mix, one could predict power prices with higher accuracy.

Bibliography

[1] Francesco Castellania et. al. 'Wind energy forecast in complex sites with a hybrid neural network and CFD based method'. In: *Energy Procedia* 45 (2014), pp. 188–197.

[2] Zhongxian Men et. al. 'Short-term wind speed and power forecasting using ensemble of mixture density neural networks'. In: *Renewable Energy* 87 (2015), pp. 203–211.

[3] Harry J. Allen and jr. Alfred J. Eggers. *A Study of The Motion and Aerodynamic Heating of Ballistic Missiles Entering the Earths Atmosphere*. California, 1958.

[4] Haiyang Zheng Andrew Kusiak and Zhe Song. 'Short-Term Prediction of Wind Farm Power: A Data Mining Approach'. In: *IEEE* 24.1 (2009), pp. 125–137.

[5] WindSim AS. *History*. URL: https://windsim.com/about/history.aspx. (accessed: 20.04.2018).

[6] M. Hashem Nehrir Ashraf Ul Haque and Paras Mandal. 'A Hybrid Intelligent Model for Deterministic and Quantile Regression Approach for Probabilistic Wind Power Forecasting'. In: *IEEE Transactions on Power Systems* 29.4 (2014), pp. 1663–1673.

[7] European Wind Energy Association. *Wind Energy and Electricity Prices. Exploring the 'merit order effect'*. 2010.

[8] Rudolf Zauner Begoña Guezuraga and Werner Pölz. 'Life cycle assessment of two different 2 MW class wind turbines'. In: *Renewable Energy* 37.1 (2011), pp. 37–44.

[9] Kanna Bhaskar and S. N. Singh. 'AWNN-Assisted Wind Power Forecasting Using Feed-Foreward Neural Network'. In: *IEEE Transactions on Sustainable Energy* 3.2 (2012), pp. 306–316.

[10] Petter Dannevig and Knut Harstveit. *Vind, i Store Norske Leksikon*. URL: https://snl.no/vind. (accessed: 12.04.2018).

[11] Sci-kit Developers. $sklearn.model_selection.KFold$. URL: http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html. (accessed: 06.05.2018).

[12] Aage Gerhardt Drachmann. 'Heron's Windmill'. In: *Centaurus* 7.2 (1961), pp. 145–151.

[13] *Electricity grid simple- North America.svg*. URL: https://en.wikipedia.org/wiki/File:Electricity_grid_simple-_North_America.svg. (accessed: 12.04.2018).

[14] Anne Cathrine Elster and Axel Tidemann. *Maskinlæring, i Store Norske Leksikon*. URL: https://snl.no/maskinl%C3%A6ring. (accessed: 23.04.2018).

[15] Emelysjosasen. *Smøla vindpark*. URL: https://commons.wikimedia.org/wiki/File:Sm%5C%C3%5C%B8la.jpg. (accessed: 12.04.2018).

[16] Inc. Encyclopaedia Britannica. *Components of a wind turbine*. URL: https://www.britannica.com/media/full/645101/125134. (accessed: 07.05.2018).

[17] The Editors of Encyclopaedia Britannica. *Partial differential equation*. URL: https://www.britannica.com/science/partial-differential-equation. (accessed: 19.04.2018).

[18] eSett. *eSett - Handbook*. URL: https://www.esett.com/handbook. (accessed: 18.04.2018).

[19] *File:Powercurve.png*. URL: https://commons.wikimedia.org/wiki/File:Powercurve.png. (accessed: 11.05.2018).

[20] Arne Gravdahl. *Power Forecasting - Hybrid Approach*. Presentation. 2018.

[21] Arne Gravdahl. *WindSim - The basics*. Presentation. 2018.

[22] Sigbjørn Hervik. *Numerisk metode, i Store Norske Leksikon*. URL: https://snl.no/numerisk_metode. (accessed: 19.04.2018).

[23] Knut Hofstad. *Generator - energiformer, i Store Norske Leksikon*. URL: https://snl.no/generator_-_energiomformer. (accessed: 13.04.2018).

[24] Trygve Holtebekk. *Fysikk, i Store Norske Leksikon*. URL: https://snl.no/fysikk. (accessed: 18.04.2018).

[25] William L. Hosch. *Navier-Stokes equation*. URL: https://www.britannica.com/science/Navier-Stokes-equation. (accessed: 20.04.2018).

[26] IRENA. *Renewable Power Generation Costs in 2017*. Abu Dhabi, 2018.

[27] Kristian Kersting Jörg Lässig and Katharina Morik. *Computational Sustainability*. Springer internatinoal Publishing, 2016. Chap. 2. Wind Power Prediction with Machine Learning. ISBN: 978-3-319-31856-1.

[28] Gill Juleff. 'An ancient wind-powered iron melting technology in Sri Lanka'. In: *Nature* 379 (1996), pp. 60–63.

[29] Andrew Kusiak and Zijun Zhang. 'Short-Horizon Prediction of Wind Power: A Data-Driven Approach'. In: *IEEE Transaction on Energy Conversions* 25.4 (2010), pp. 1112–1122.

[30] CFD Marine. *Om oss*. URL: http://www.cfdmarine.com/historie/. (accessed: 20.04.2018).

[31] Stephen Marsland. *Machine Learning, An Algorithm Perspective*. CRC Press, 2015. ISBN: 9781466583283.

[32] Flavio Corbo Matteo Mana and Dr. Cathérine Meißner. 'Short-Term Forecasting of Energy Production Using CFD Simulations'. In: *EWEA* (2013), p. 1.

[33] BMT Fluid Mechanics. *Flow Assurance*. URL: http://www.bmtfm.com/sectors/oil-and-gas/flow-assurance. (accessed: 20.04.2018).

[34] Gino Moretti and Michael Abbetti. 'A time-dependent computational method for blunt body flows'. In: *AIAA* 4.12 (1966), pp. 2136–2141.

[35] Inci Okumus and Ali Dinler. 'Current status of wind energy forecasting and a hybrid method for hourly predictions'. In: *Energy Conversion and Management* 123 (2016), pp. 362–371.

[36] Zaccheus O. Olaofe. 'A 5-day wind speed and power forecasts using a layer recurrent neural network (LRNN)'. In: *Sustainable Energy Technologies and Assessments* 6 (2014), pp. 1–24.

[37] Encyclopædia Britannica Online. *Mariner's compass card*. URL: https://www.britannica.com/technology/compass-navigational-instrument?oasmId=148655. (accessed: 07.05.2018).

[38] Kjeller Vindteknikk Øyvind Byrkjedal Espen Åkervik. *Vindkart for Norge*. Norges vassdrags- og energidirektorat, 2009.

[39] Erik D.Delarue Patrick J.Luickx and William D.D'haeseleer. 'Considerations on the backup of wind power: Operational backup'. In: *Applied Energy* 85 (2008), pp. 787–799.

[40] Nord Pool. *Day-ahead overview*. URL: https://www.nordpoolgroup.com/maps/#/nordic. (accessed: 11.05.2018).

[41] Nord Pool. *The Power Market*. URL: https://www.nordpoolgroup.com. (accessed: 17.04.2018).

[42] Trevor J. Price. 'James Blythe - Britain's first modern wind power pioneer'. In: *Wind Engineering* 29.3 (2005), pp. 191–200.

[43] Knut A. Rosvold and Knut Hofstad. *Vindkraftverk, i Store Norske Leksikon*. URL: https://snl.no/vindkraftverk. (accessed: 13.04.2018).

[44] Statistisk Sentralbyrå. *Energi og Industri - Elektrisitet*. URL: https://www.ssb.no/energi-og-industri/statistikker/elektrisitet/aar. (accessed: 16.04.2018).

[45] Mantas Marĉiukaitis. Inga Žutautaite. Linas Martišauskas. Benas Jokšas. Giedrius Gecevičius. Athanasios Sfetsos. 'Non-linear regression model for wind turbine power curve'. In: *Renewable Energy* 113 (2017), pp. 732–741.

[46] Dennis G. Sheppard. *Wind Turbine Technology: Fundamental Concepts in Wind Turbine Engineering, Second Edition*. American Society of Mechanical Engineers,U.S., 2009. Chap. 1. Historical Development of the Windmill. ISBN: 9780791802601.

[47] Statnett. *Behov for fleksibilitet*. URL: http://www.statnett.no/Samfunnsoppdrag/hvordan-na-klimamalene/behov-for-fleksibilitet. (accessed: 17.04.2018).

[48] Statnett. *Impala*. URL: http://www.statnett.no/Samfunnsoppdrag/Forskning-og-utvikling/Vare-satsningsomrader/smarte-nett/Impala. (accessed: 17.04.2018).

[49] Statnett. *Operating the grid is all about balance*. URL: http://www.statnett.no/en/Sustainability/Our-social-mandate/Operating-the-grid-is-all-about-balance. (accessed: 17.04.2018).

[50] Statnett. *Rapport fra Systemansvarlig. Om kraftsystemet i Norge 2016*. 2016.

[51] Statnett. *Reserve Markets*. URL: http://www.statnett.no/en/Market-and-operations/reserve-markets/. (accessed: 11.05.2018).

[52] Erik Tandberg. *Vindtunnel, i Store Norske Leksikon*. URL: https://snl.no/vindtunnel. (accessed: 18.04.2018).

[53] Erik Tandberg and Tor Halvor Johansen. *Aerodynamikk, i Store Norske Leksikon*. URL: https://snl.no/aerodynamikk. (accessed: 18.04.2018).

[54] Tesla. *Autopilot*. URL: https://www.tesla.com/autopilot. (accessed: 24.04.2018).

[55] lb Troen and Erik Lundtang Petersen. *European Wind Atlas*. Roskilde, Denmark, 1989.

[56] UNESCO. *Wind and Solar Energy, Proceedings of the New Delphi Symposium*. Imprimerie Chaix, Paris, 1956.

[57] John F. Wendt. *Computational Fluid Dynamics, An Introduction*. Springer-Verlag Berlin Heidelberg, 2009. ISBN: 978-3-540-85055-7.

[58] Xiaobin Huang Xiaochen Wang Peng Guoc. 'A Review of Wind Power Forecasting Models'. In: *Energy Procedia* 12 (2011), pp. 770–778.

[59] Marc'Aurelio Ranzato Yaniv Taigman Ming Yang and Lior Wolf. *DeepFace: Closing the Gap of Human-Level Performance in Face Verification*. California, 2014.

[60] Jianwu Zeng and Wei Qiao. 'Support Vector Machine-Based Short-Term Wind Power Forecasting'. In: *IEEE* 11 (2011), pp. 1–8.