MSc Thesis

Powering the Orkney Cloud

Modeling and Understanding
Curtailment in Orkney

Niels Ørbæk Christensen, niec@itu.dk

Advisor: Philippe Bonnet, phbo@itu.dk

IT University of Copenhagen

May 2019

Abstract

In this thesis we explore the idea of an Orkney Cloud data center, a decentralized data center running on the surplus renewable energy otherwise curtailed in the power grid of Orkney, Scotland. The concrete issue addressed is the unclarity of the logic governing the *Active Network Management* (ANM) curtailment system in Orkney. The focus is to gain the needed understanding of the ANM to design and implement systems that can adapt their power usage to the state of the grid. This is done by inferring the context from available data sources, namely data made public by the grid provider and weather data.

We construct a dataset by scraping publicly available data at 10-minute intervals for three months. We then model our assumptions about the ANM in terms of generation and demand, which we evaluate against our dataset. Inspired by the method of *Exploratory Data Analysis* (EDA), we investigate our assumptions and look for patterns in the data through a series of visualizations. From this we extract correlations allowing us to represent our initial model in terms of factors external to the Orkney power grid: Wind speeds, time-of-day and day-of-week. To address the curtailment situations not described by our models, we search for anomalies in our data set, both by visual exploration and by applying different anomaly detection rules. When redacting these anomalies, we see the ratio of time with any curtailment in the ANM reduce by over 50%.

As a supplement to the manually constructed models, we also evaluate models based on *Artificial Neural Networks* (ANNs), finding a slight increase in model accuracy, but no increases in tangible understanding of the ANM, indicating that ANN-based models can be useful for deployment, but not for the EDA-process. We analyze the data collected by a single wind turbine and find that the majority (> 90%) of the loss reported does not seem to be caused by ANM-issued curtailment. Using the wind data generated by this turbine as input to our ANN-based model leads to additional accuracy gains. Finally, we combine our correlation-based model with publicly available weather forecasts to obtain a predictive model, allowing us to forecast future curtailment.

The models developed in this project can serve as indicators that a significant amount of the curtailment is caused by something other than the demand/generation balance, supporting the stakeholders feeling of an unclear logic governing the ANM. When supplying the models with the developed anomaly detection rules, they can be used to implement the power controller logic of an Orkney Cloud data center.

Contents

Terms and Clarification

- We use "ANM" and "ANM System" interchangeably
- In regard to power grids, we use "demand" and "load" interchangeably
- We will use "generation" and "production" interchangeably when concerning energy.
- As the focus of this project is the electrical power grid, we will often use "energy" and "power" and "electricity" to mean the same thing. We do not include other types of energy sources, such as fossil-fuels, in our discussion of energy generation or demand.
- When using the word "significant" about results or correlations, we do not refer to a defined statistical significance, but rather to the normal definition of the word: "sufficiently great or important to be worthy of attention; noteworthy".

About the Cover

The cover photo depicts the wind turbines at Hammars Hill on Mainland, Orkney. The five turbines have a combined capacity of 4.5MW, and are subject to ANM-based curtailment.

1 Introduction

In this section we will introduce the project described in the present thesis. We will give a brief overview of the context in which the project is situated, by introducing the problem domain and the project focus. After this, we will more clearly define the problem addressed, and the approach used. Finally we will list the main contributions of the project.

1.1 Context

In order to give a proper introduction to the problem domain of this project, we will in this section briefly describe the issues that have formed its final focus. We describe the issue of centralized cloud computing and its power consumption, the vision of the Orkney Cloud data center, and the issue of curtailment in relation to cloud data services. Finally we will describe how these aspects have formed the project focus.

1.1.1 Cloud Computing and Data Centers

In the past years, we have seen an exponential increase in global IP traffic, estimated to reach 4.8 zettabytes ($zetta = 10^{21}$) annually by 2022, corresponding to 50GB per capita per month [1]. This increase has been ushered in by the rise of cloud computing and *Software as a Service* (and other "X as a Service") business-models, moving a large part of both computation and storage from the computers of end-users to warehouse-scale data centers located strategically around the globe, transferring files and results to the user through the internet on demand.

This brings great flexibility to the end-user, who does not need to worry about software installation or updates, and can switch seamlessly between devices, or even retire old devices, without migrating data or software along. These responsibilities are shifted to the data centers, which are optimized for both performance and robustness.

A stable power connection is crucial for these data centers, as a sufficient and constant energy supply is needed to secure the uptime of the many servers. And a lot of power is needed, as a typical data center may consume as much energy as 25,000 households, of which 50% is used solely for cooling [2]. The combined energy usage of the worlds data centers was estimated to 416.2 terawatt hours (TWh) in 2015, which is 40% more than the entire UK that same year [3].

Denmark has a historically stable power transmission grid, and several of the largest tech companies in the world - including Facebook, Apple and Google - plan on using this grid for powering some of their new data centers, which are predicted to account for 16.7% of Denmark's energy consumption when complete [4].

Whether data centers like this constitute a net increase in global energy usage is up for debate.[1] But regardless of whether large-scale data centers are good for the worlds power consumption or not, it is clear that they have a large energy footprint that should be managed and minimized in order to reduce the emission of greenhouse gasses.

The cloud services running on these warehouse-scale data centers are inherently centralized systems, where the data of each user is constantly transferred back and forth. And for the users located far from these data centers, at the edge of the network, where connectivity is poor and bandwidth is limited, this can create bottlenecks and slow response times.

To combat this issue, and lighten the load on the network and data centers, a lot of research and development in the field of Internet of Things (IoT) have been focused on *edge computing*[2] (e.g. [6, 7]), shifting the computation closer to the leafs of the network tree, making the system as a whole more decentralized.

1.1.2 The Orkney Cloud Data Center

Orkney is an archipelago of around 70 islands situated off the northeastern coast of Scotland, with a total of 20,000 inhabitants. The islands have a rich history, dating back to the neolithic age, including 600 years under Norwegian rule, until being annexed by Scotland in 1472.

These islands are in a unusual situation when it comes to energy, as they have an abundance of renewable energy sources in their environment. The islands are incredibly windy, and home to both large waves and one of the strongest tidal currents in the world. Orkney is host to the world's largest test bed for wave- and tidal-based generators, but the vast majority of the energy being produced on the islands is wind based.

These wind turbines are distributed across the islands, with generators varying from larger-scale turbines, many of which are owned by island communities, to the approximately 700 domestic-scale turbines, owned by individuals. The energy production is highly decentralized, with a relatively large part of the 20,000 orcadians as direct stakeholders. Since 2013 Orkney has been a net exporter of electricity, producing more renewable energy than they consume every year.

The Orkney Cloud research project [8] works with the visions for cloud computing and data services that arise from this abundance of low-emission energy. The research project was originally funded by a one-year grant by

[1] An example of this is video streaming. Based on a life-cycle analysis [5], the energy usage of DVD discs are higher than that of video streaming, but as the majority of internet traffic is video streaming, accounting for over 75% of all IP traffic [1], one could imagine that the constant availability of content that comes with streaming services also increases the amount of video watched, possibly leading to an increase in energy consumption.

[2] Sometimes called *fog computing*, analogous to cloud computing

the Mozilla Foundation, resulting in a magazine presenting the findings of the group [9].

One of the visions of the research group is that of an Orkney Cloud data center [10], a decentralized data center comprised of many autonomous devices, or *pods*, with computing and storage capabilities, wirelessly connected across the islands. All of which are powered directly off renewable energy sources, using built-in energy storage to secure up-time while power generation is low. It is this vision that we will explore in this project.

During the preparation for this project, described in [11], I traveled to Orkney to get an impression of the islands and its inhabitants, and from this trip it became apparent that data services as an abstract category was hard to relate to for many orcadians. On the other hand, the issue of *curtailment*, intentional reduction of the output of generators, was highly relevant and important to the orcadians we talked to.

A major problem for Orkney is that they are often unable to accommodate all of the energy they produce. There is a single 20MW transmission line to mainland Scotland, limiting the amount of power that can be exported from the islands [12]. This results in frequent curtailment of turbines, reducing the production of low-emission energy, while, concurrently, other regions are being powered by fossil fuels. A widening of this transmission bottleneck is expensive, and funding it has proved difficult.

This problem is not unique to Orkney or the UK. China, the biggest producer of wind energy in the world, is tackling curtailment issues of an entirely different scale. In 2012, China lost 20.8 TWh of electricity due to curtailment [13], more than double the total generation of wind energy in Denmark that same year [14]. Again, the main reason was bottlenecks in the power transmission capacity from the northern regions of China to the demand center in the south-east.

Besides the obvious environmental problem, curtailment is also an economical problem, resulting in lost revenue for both the large wind farm operators in China and the smaller island communities in Orkney. This creates an even greater incentive to find ways of accommodating the power that would otherwise be curtailed.

1.1.3 Project Focus

In the above sections, we have described three important aspects for this project: The data centers that are the backbone of modern cloud computing, the Orkney islands and their plentiful renewable energy sources, and the issue of curtailment, forced reduction of renewable power generation.

This leads us to evolve the vision of the Orkney Cloud data center, going beyond simply sourcing the power from renewable sources, and instead making the data center react to the state of the grid, trying to only use power that would otherwise be curtailed.

In order to implement this functionality, we need a system that models some aspects of the Orkney power grid and is able to identify when curtailment is going to be issued. We find this modeling both highly interesting and important in designing a data center with a smarter energy footprint and one that is more relevant for Orkney. It is this modeling that is the focus of our project.

1.2 Problem

In order to design a data center that only uses power that would otherwise be curtailed, we need some way of identifying when curtailment is happening. In other words, we need the data center to be *context-aware*, where the context in this case is the state of the power grid. This context is known to the grid provider, but not to the orcadians whose turbines are being curtailed.

One of the frustrations of the orcadians, with regards to curtailment, is that they get little information about when and to which degree their turbines are going to be curtailed, as this is handled automatically. The grid provider, Scottish and Southern Electricity (SSE), does not inform the owners of curtailed turbines when curtailment is being issued, and they can only find out the amount of curtailment by looking at their actual power output and contrasting that with an estimated uncurtailed power output.

Orkney therefore has a great interest in finding ways of generating this knowledge, without relying on direct SSE involvement. And in the Orkney tradition of getting stuff done themselves [15], they are not keen on waiting for external forces to fix the problems for them. They have started numerous collaborations with universities, of which this is a part, investigating and experimenting with ways to create an energy system of the future.

The problem then becomes one of inferring the context of the system from publicly available data sources, and using that information to identify whether curtailment is happening. This information can then potentially be used to predict whether curtailment will happen at a given time in the future, and to issue actions that could alter the state of the grid to reduce curtailment. And this leads us to the research question we will seek to answer in this project.

Research Question

> *How can we identify curtailment of turbines in the Orkney power grid by inferring the context from publicly available data sources?*

1.3 Approach

The methodology of this project is mainly based on *exploratory data analysis* (EDA) [16, 17], an inherently iterative process for working with a dataset, described by John W. Tukey as "quantitative detective work" [16, p. 21]. The purpose of this method is to extract knowledge from data, not simply by answering the questions we already have, but also raising new questions by exploring the possibilities and limitations of the data. This should be complimented by *confirmatory data analysis*, in which the patterns, or lack thereof, identified through EDA can then be confirmed by evaluation.

There is not a number of defined steps in EDA, but rather a set of techniques that should be used according to the affordance of the data and the accumulated insights. Of these techniques we will primarily be using visualization, querying and computation. As a result we have worked extensively with visual representations of the data, as will figure prominently throughout the report.

So, we need some data to explore. To construct a dataset, we identify the relevant available sources, including the information made public by SSE, and continuously gather it for a set time period. In this dataset we look for patterns and anomalies that might reveal knowledge about the issue of curtailment. We then describe an initial simple model of when we would expect curtailment to occur, and use that as a starting point for our analysis.

1.4 Contribution

The main contributions of this project can be summarized as the following:

- **Data scraping:** We construct a dataset, through continuous web scraping, with which it is possible to evaluate assumptions about the ANM and the Orkney power grid.

- **Simple descriptive model:** We describe and evaluate a simple model of the ANM based on our initial assumptions, and adapt it to fit our dataset.

- **Visualizing Data:** Through different types of visualizations we are able to gain insights, identify patterns in the data, and extract correlations.

- **Correlation-based models:** Based on these correlations we describe our simple model in terms of parameters external to the ANM, reducing our dependency on grid-specific data.

- **Data cleaning and anomaly detection:** Using visual exploration and intuitive reasoning we identify anomalies in the dataset, which are then removed or altered accordingly, increasing our model accuracy and allowing us to quantify the amount of anomalous data.

- **Neural network models:** We implement Artificial Neural Network-based models using the same parameters as the previous models, leading to slight accuracy gains, and we discuss whether this type of model is suitable for this problem.
- **Applications of dataset and insights:** We investigate the data produced by a single turbine in Orkney, and in combination with our constructed dataset, we are able to calculate an upper bound for their curtailment-based power loss. Furthermore, we propose a new method for power output estimation for the turbine.
- **Predictive models:** We combine our descriptive models with weather forecasts to create models that predict whether curtailment will happen at a given time in the future. These models are evaluated based on forecast horizon, and a method for estimating probability for curtailment is presented.

1.5 Report Structure

To assist the reader, we will here give a brief overview of the structure of this report. Sections 1-3 serve as an introduction to the problem addressed, the domain in which we are working, and describe what we believe to know about the Orkney power grid in terms borrowed from systems theory. Sections 4-11 present the main contributions of the project, from the construction of a dataset to the evaluation of a series of models of the Orkney ANM. Finally, sections 12 and 13 contain the concluding remarks and possible avenues for further work.

2 Background

In this section we present the Orkney power grid, its Active Network Management (ANM) functionality, and the problems surrounding it. We also present the concepts of Demand Response and Dynamic Demand Response, not currently implemented in the Orkney grid.

2.1 The Orkney Power Grid

To give the reader an understanding of the system we are modeling, we will present some of the relevant aspects of the Orkney power grid, namely the primary power sources and the Active Network Management system.

2.1.1 Power Sources in Orkney

Orkney has a host of impressive renewable energy sources available in its natural environment: The islands are constantly battered with large waves from the Atlantic ocean, the tidal currents passing by Orkney are among the strongest in the world, and the weather is extremely windy. Orkney is used as a test bed for novel technologies for harvesting wave and tidal energy, but the primary energy source on the islands are wind turbines.

In 2013 (the newest figures available) the wind turbines in Orkney generated an estimated 140 GWh [18]. For comparison, the entire combined (domestic and non-domestic) electricity usage for the Orkney Islands in 2014 was also 140 GWh [19]. Around 90% of this power comes from larger turbines (>50kW capacity), many of which are owned by island communities. The rest comes from over 700 domestic scale wind turbines distributed around the islands.

2.1.2 Active Network Management

One of the predominant challenges for power grid operators, is the need for constant balancing of load and generation, or simply "balancing the grid". The generation of power needs to follow the demand for power closely, to make sure that the demand can be met while also ensuring that the grid is not over capacity. The grid is over capacity if the generation exceeds the demand over a period, resulting in too much electricity present in the grid. With the global transition to renewable energy sources, this task becomes increasingly harder, as these sources are inherently unstable, and one cannot, for example, throw more coal on the fire when the load increases.

The power grid provider in Orkney, SSE, in 2009 installed an *Active Network Management* (ANM) system in Orkney, which is able to react to the state of the grid in real-time and ensure that the transmission network is not over capacity. This kind of system if often put in the category of *Smart Grids* [20].

The ANM balances the grid by reducing generation in times where the grid is over capacity. This is done by sending signals through the system issuing *curtailment* to certain generators, meaning that they are held back or completely stopped. So in situations of high renewable generation and lower demand, turbines and other generators are curtailed and forced to produce fewer emissions-free electrons than they could.

The ANM distinguishes between generators on a *firm connection* and those on a *non-firm connection*, and only those on non-firm connections can be curtailed. But the concrete logic of the ANM, when and which turbines that are issued curtailment, is unclear for a lot of the stakeholders. And an understanding of this is important for the Orkney society, as most of the wind turbines in Orkney are community-owned and curtailment hurts the economies of these small island communities. Curtailment happens without warning and so far the communities have not been able to see a clear system or pattern, making it hard for them to predict or plan for. Furthermore, energy producers in Orkney pay the highest transmission tariff in the country,[3] making the profit margin on the power generated smaller, and the impact of curtailment even greater.

This curtailment system is not optimal in regard to the greater societal goal of reducing emissions of greenhouse gases, as it is actively reducing the generation of emissions-free energy at one moment, while a couple of hours later importing energy generated from burning fossil fuels.

From this perspective the ANM might look like a strictly bad deal, but it also enabled an additional 24,2MW worth of wind turbines to be added to the grid [21], by guaranteeing that the grid will not be over capacity. This allows Orkney to run on renewable energy when wind speeds are lower or demand is higher, but results in curtailments when demand is lower or wind speeds are higher. One could even argue that a smart grid like this is only half smart, since it only scales the generation of power and is not affecting the demand for power at all. In the following section we will look at methods of affecting the load as well.

2.2 Demand Response

The concept of adapting not the generation, but the load when balancing a power grid is called *Demand Response* (DR) [22]. DR policies can be created to increase/decrease the load in certain parts of the grid at certain times in order to match the demand to the current generation. Traditionally, these policies are static, based on assumptions and experience, and their purpose is to *smooth the demand curve*, seeking to make the variations in demand as small as possible, suitable for larger fossil-fuel-based power plants, that are slow to turn on or wind down. One example could be the course of the day: People tend to use some electricity in the day time, a lot in the evening, and little at night. This is a fairly predictable cycle, and a DR policy would then try to move load-intensive tasks away from the evenings.

In the large scheme of things, people can be statistically predictable, but renewable energy, especially wind power, is highly volatile and stochastic by nature, and here the static policies come up short. In order to adapt our demand to this fluctuating source of energy, there is a need to go beyond the fixed rules of DR and instead implement *Dynamic Demand Response* (D^2R) [23], where policies are dynamic and based on the current state of the grid and of the suppliers and consumers that are connected to it. Here the purpose is not to smooth the demand curve, but rather make the demand curve follow the fluctuating generation curve.

In Denmark, the large actors in the market have implemented DR and D^2R policies, relying on consumers to scale demand. Radius, a large grid

[3] The energy producers in Orkney are charged 25£/kW, while a site near London is awarded 6£/kW for putting power into the grid [12]

provider, have chosen to make a separate tariff for the evening hours (17:00-20:00) from October to March, increasing the price by 123% in these peak hours and lowering it by 14% in the remaining hours [24]. Similarly, the largest danish energy company, Ørsted, is currently introducing per-hour dynamic pricing based on demand and generation, even allowing the prices to be negative under certain circumstances [25].

These are indirect ways of implementing of DR and D^2R, respectively, by incentivizing consumers to adapt their behavior to the (predicted or measured) state of the grid. This is based on the assumption that consumers will be able to adapt to these changing prices, which we suspect might be especially challenging when it comes to Ørsted's strategy, as few people stay continuously informed on the current energy prices.

A more automated way of going about D^2R, is proposed by orcadian start-up Solo Energy,[4]. Their solution relies on distributed energy storage in the form of home batteries, typically with a capacity of ~10 kWh, installed in the homes of regular private consumers. These batteries are then remotely controlled by Solo Energy's software platform, buying energy and charging batteries when prices are low, and releasing it back into the grid when prices rise. Their business model also includes aspects of peer-to-peer trading and blockchain technology, which are not relevant to this project.

The energy storage solution proposed by Solo Energy could be a way to reduce curtailment in Orkney, but its actions are governed by the electricity prices which reflect the energy situation in a larger area and might not reflect the curtailment situation in Orkney very well.

2.3 Summary

In this section we have described the Orkney power grid, its connections, power sources and the Active Network Management system that helps balancing load and generation on the grid by curtailing generators when the grid is reaching its maximum capacity. This system allowed for more generators to be added to the grid, but it also presents some problems for the communities of Orkney that own a great deal of these generators, as they are unable to predict or plan for curtailments, which is hurting their economies.

We also presented the concept of Dynamic Demand Response, a way of responding to balancing issues in the grid by scaling the load and not just the generation. We presented some current implementations of this, although we found none that specifically address the issue of curtailment by increasing the load at times of high generation.

In the following sections, we will develop a series of models in our effort to understand the ANM and the issue of curtailment in Orkney.

[4] See https://solo.energy/ last accessed May 14th, 2019

3 The Orkney Grid as a System

To understand the ANM system, it is relevant to understand the system of which the ANM is a subsystem, the Orkney power grid. In order to do this we will look at the grid through the lens of a few concepts from systems theory as described by Donnela Meadows in [26]. We will introduce the basic terms, and then describe the grid as a *one-stock system*.

3.1 Systems Theory Concepts

A system can be described as a collection of *stocks*, *flows* and *feedback loops*. A *stock* is a countable or measurable aspect of the system. It is a quantity or accumulation of something relevant to the system. The stocks of a system can change over time through the actions of flows.

A *flow* is an input or output of the stock, either increasing or decreasing the stock. If a stock has no incoming or outgoing flows, its level will be static, but if it has more outgoing flow than incoming, its stock will decrease, and vice versa. If the incoming flow matches the outgoing flow, the stock will be in a state of *dynamic equilibrium*, maintaining its quantity even though the elements of the stock are constantly in flux.

The third basic element is the *feedback loop*, which is a reaction to the state of a stock, which then affects one or more flows. Feedback loops can either be *reinforcing*, meaning, for example, that an increase in stock results in an increase in incoming flow, or they can be *balancing*, meaning that they will try to affect flows in a way to keep stock at a desired value.

Systems have an inherent temporal aspect, as stocks change over time according to the rate of flows. Change is not instant, and feedback loops and their reactions are not instant either. There is always a delay in reaction, and the effect of the reaction might not be fast or strong enough to counter a given change in stock. These interrelations and their properties are all important elements of systems, but they can be hard to identify as they are not always tangible or explicit.

Systems described in this way can be almost arbitrarily abstract or complex, and the level of detail in the stocks, flows, and loops depends on the desired utilization of the model. In our case, we will be describing a very high level model to highlight the interrelations between different kinds of flow and a single stock, omitting several other elements from our description.

3.2 The Grid as a One-Stock System

For many systems it is a primary task to secure dynamic equilibrium of a single stock. For example, a factory producing a single commodity will want to match its production to its sales. In the case of power grids, this translates to what we have called "balancing the grid", meaning to ensure a balance between power generation and load. The grid has a constant outgoing flow of electricity demand and a constant incoming flow of power production, and these two flows need to match within a certain margin.

To formalize this we can describe the Orkney grid as a system with a single stock: The amount of energy in the grid. The visual model of this system is shown in Figure 1. The single stock has a desired value, determined by the size and elements of the grid, and the level of the stock is determined by four flows:

Generation (in): The collective electricity production in Orkney.

Transmission (in): 20MW transmission cable from mainland Scotland

Demand (out): The collective electricity consumption in Orkney.

Transmission (out): 20MW transmission cable to mainland Scotland

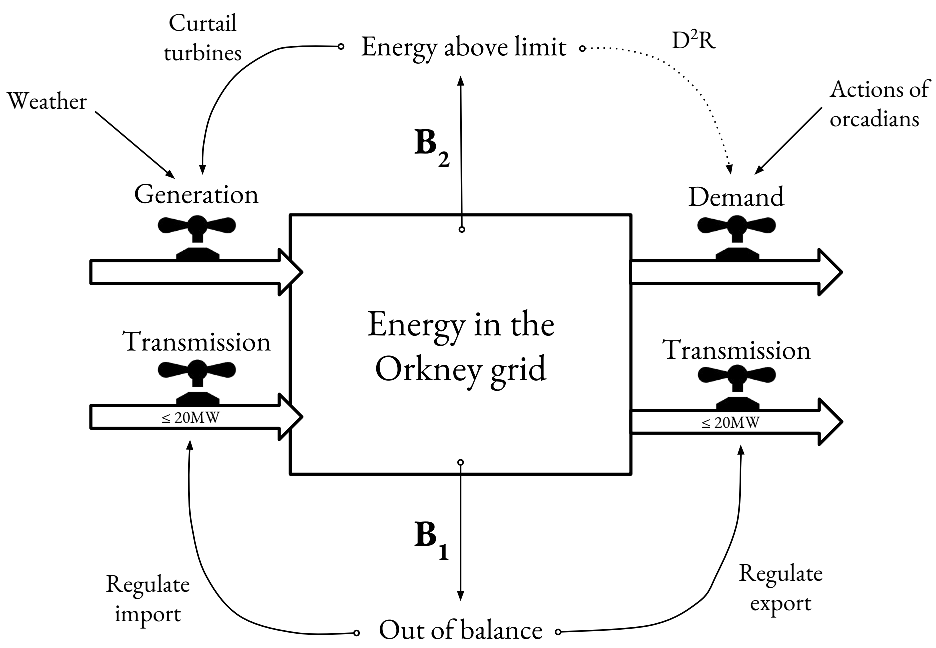


Figure 1: The Orkney power grid described as a one-stock system

If the grid operator was the only actor adjusting the flows, this would be a rather simple task: The generation would simply be set to match the demand and that would be that. The transmission flows could be turned off.

In the case of the Orkney grid however, the flows are also adjusted by external factors. The generation is determined by the state of the renewable sources (especially the winds), and the demand is determined by the actions of all orcadians on the grid. These changes in flow can alter the level of the stock and create a discrepancy between the actual value and the desired value.

To return the stock to the desired value, the system reacts by invoking two balancing feedback loops, denoted as $\mathbf{B_1}$ and $\mathbf{B_2}$, in hierarchical order. $\mathbf{B_1}$ adjusts the transmission to and from mainland Scotland, and is the first loop to be invoked to regulate the level of the stock. However this loop has a limitation to its effect: the transmission cables have a maximum capacity of 20 megawatts each way.

In case $\mathbf{B_1}$ is not capable of regulating the stock sufficiently (often because the inflow from generation is simply too high), $\mathbf{B_2}$ is invoked. It reacts to the case of too much energy in the grid by issuing generator curtailments, thereby decreasing the flow from the energy generation.

A second possible (but not currently implemented) reaction of $\mathbf{B_2}$, would be to increase the outgoing flow of demand, a case of D^2R.

3.3 Summary and Discussion

We now have an understanding of the Orkney grid as a system trying to regulate the level of a single stock, the amount of energy in the grid, by adjusting a series of incoming and outgoing flows through balancing feedback loops.

This model is of course very simplified. All of the many heterogeneous generators in Orkney are put in a single category, and curtailment is applied to the collection as a whole, but in reality the choice of which generators to curtail is more complex.

And though it can give an intuitive understanding of the grid, the model says nothing about the delays or effectiveness of the loop's reactions. This is an important aspect of the system, as slow reaction times could cause $\mathbf{B_2}$ to be invoked even though the transmission lines are not at full capacity. It might also be that the reaction time of $\mathbf{B_1}$ is lower than that of $\mathbf{B_2}$, making it desirable for the system to reserve some capacity on the transmission lines for sudden spikes in generation that $\mathbf{B_2}$ cannot react to in a timely fashion. In this case the two loops would not be as hierarchical as we have assumed.

The main take-away here is that while we can understand the basic functionality of the grid, there is still a lot of unknowns in the concrete interrelations of the elements. To investigate this we will construct a dataset and use it to evaluate our assumptions of the grid as a system.

4 Constructing a Dataset

To be able to evaluate our assumptions about the ANM, we have contructed a dataset, which we will describe in this section. We present the relevant data sources available, and how we collected and stored the data that is the basis for both developing and evaluating our models. We will also discuss the limitations and known biases of this constructed dataset.

4.1 Data Sources

In order to have something to evaluate our models against, we need data sources that can tell us something about the curtailment situation and its context.

4.1.1 SSE Website

On the SSE ANM website[5], there are two types of real time data available about the current state of the ANM:

1. The current *demand and generation* for the grid of all of Orkney, along with the peak demand and the total renewable generation capacity.
2. The status of each of the *ANM zones*, showing whether generators are currently being curtailed or stopped in each zone.

Figure 2 shows a simplified map of Orkney, and of the zones and measurement points of the ANM system. Curtailment is issued based on the state at the measurement points, which SSE calls the "constraint points",[6]. It is not clear what these constraints are, or how they interact. The total generation of Orkney is split into the generation of the firm and the non-firm (ANM-connected) generators.

The SSE does not expose any official API for this data, so in order to gather a historical dataset, we have scraped the data from their website every 10 minutes since the 30th of November 2018, storing the values in a local database.

Furthermore they make a note on the website that *"Although we make reasonable efforts to update the information on our site, we make no representations, warranties or guarantees, whether express or implied, that the content on our site is accurate, complete or up-to-date."* [27]. But since this is the only available ANM data source both we, and the orcadians, have, we choose to use it regardless.

[5]See https://www.ssen.co.uk/ANM/, last accessed May 14th, 2019

[6]See "What are ANM zones" at https://www.ssen.co.uk/ANMFAQ/ last accessed May 14th, 2019

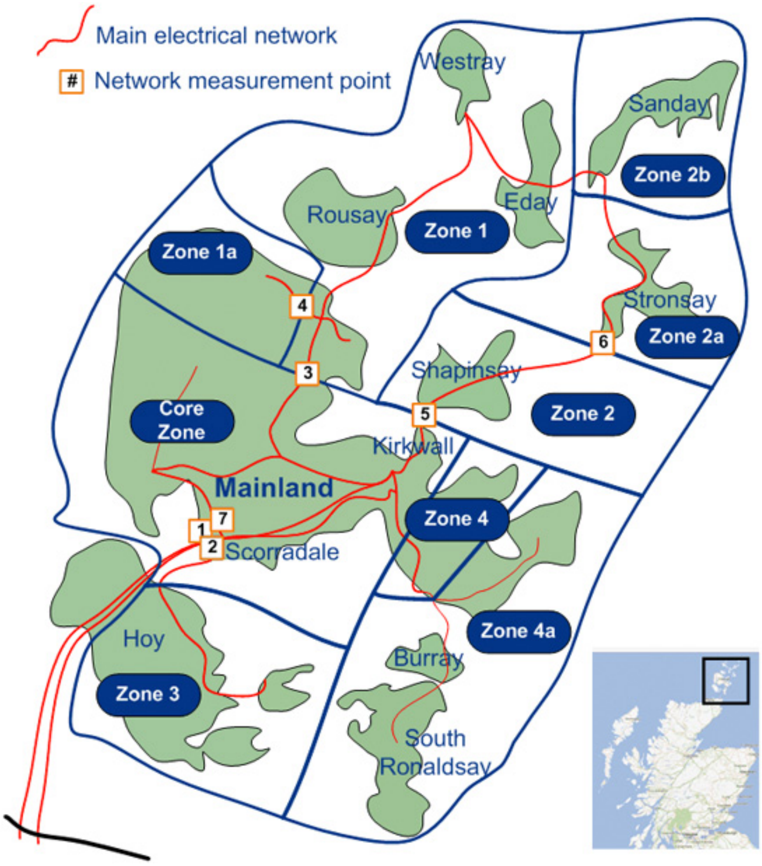


Figure 2: Map of the ANM zones

4.1.2 Weather Data

Another very relevant data source when it comes to wind energy is numerical weather data and forecasts. Here we use OpenWeatherMap[7], which exposes a free API for weather data and allows for bulk download of historical weather data.

The need for collection of weather data became apparent after the data collection had already begun. To compensate for this we have downloaded hourly readings from a station on the isle of Westray from the 1st of December 2018 until the 11th of February 2019. After this, weather data was collected every 10 minutes from the free API, though the service only guarantees to update the data once an hour, stored in the same way as the SSE data.

[7] See https://openweathermap.org/api, last accessed May 14th, 2019

4.2 Implementation

In order to collect the data continuously, we set up a small dedicated virtual server running Ubuntu 18.04. All data is scraped using Python 3 [28] and the library Requests [29] for http-requests and parsing.

For the SSE data on the current demand and generation, we inspected the HTML-page on which it is displayed and saw a request to an URL that returns the data in JSON-format. By requesting this address directly, we avoid having to parse HTML to extract the information. The ANM status page, on the other hand, makes no such requests, and we are forced to download the HTML-page to get the data we want. For this we use the HTML-parsing library Beautiful Soup [30], allowing us to query elements on the page by type or class, among other things. We parse the ANM-status for each zone and save it in a Python dictionary, which easily serializes to a JSON-object. The weather data is collected by sending a request to the OpenWeatherMap API, specifying the geographical coordinates and a free API-key, returning a JSON-object.

All data is fetched using Linux cron jobs and stored in a MongoDB database [31] hosted on the same server, with each source in a separate collection. The reason for using MongoDB is that is a *document database*, allowing for storage of JSON-like documents without the need to specify fields or keys, and allowing for fields to change over time. This is suitable for our case, since we rely on external data sources, which might change structure over time, and the highest priority for us is to keep collecting the data available. If the fields change we can simply account for that in our code.

Since our data collection is quite small, and performance is not of the highest priority for us, we choose to collect all of the data available from our sources, instead of selecting only the parts we know we want. If the system is deployed and faster querying of the database is needed, we can port the necessary data to a relational database and change our data collection accordingly.

To assemble the actual dataset, we query our different collections and use the data analysis library Pandas [32] to join them to a single dataframe, where each row constitutes a data sample. We then convert the status from each ANM-zone to a Boolean value, which is True if there is any curtailment reported in that zone at that time. We can then do a Boolean OR operation on all the zone columns, resulting in a single column denoting whether any curtailment is issued in the ANM at a given time. It is this column that we will use as the primary truth labels.

4.3 Biases and Limitations

Any real-world dataset contains biases and limitations, and this is of course also the case for the dataset we are constructing here. In this section we will try to examine some of the known biases and how they will affect the models we develop.

4.3.1 Granularity and Resampling

We only collect discrete values, and we do so every 10 minutes. This interval was somewhat arbitrarily chosen based on the fact that the monthly reports of a single known turbine (described in section 10) uses this interval. This means that we might miss curtailment situations that last less than ten minutes, and that a recorded curtailment situation will always last at least 10 minutes, even though in reality it might be shorter.

Moreover, we do not collect all data at exactly the same time. Data is gathered over the internet, each scrape sequentially, and exact times are affected by response times in the network. In order to align the data, we resample the data to 10 minute intervals, aligning the timestamps. By doing this we lose a tiny amount of precision in the data in exchange for easier handling.

4.3.2 Seasonal Bias

The data we have been collecting is only for the winter months of December, January and February, and as the seasons change, so does the weather that the wind turbines are dependent on. In fact, models in this field are often evaluated separately for the different seasons [33, 34]. But as SSE does not expose a way to get historical data and temporal continuity inhibits us from collecting real-time data from the past or future, this is a shortcoming of the dataset that we cannot fix, and therefore a bias we need to take into account when evaluating our models.

This bias is also worth noting as the winter months are usually where the problem of curtailment is least prevalent, since the general power demand is higher, meaning we might see relative low amounts of curtailment in our dataset.

4.3.3 Truth Values

When evaluating the models we develop, we will use the data on the status of the ANM zones scraped from the SSE website as the ground truth. Using this as truth has several implications for our models. First of all, we assume that the data SSE themselves expose is truthful and valid. Also, the granularity of this data is limited to the zones, so we will not be able to tell whether any single generator is curtailed. The resolution of status in each zone is also limited to three color codes:

Red: At least one generator in the zone is fully stopped

Yellow: At least one generator in the zone is partially curtailed

Green: No generators in the zone are curtailed

This has a series of shortcomings: A zone with several partially curtailed generators might lose more watts than a zone with a single fully stopped generator, but still receive a ”less severe” color code. Also, two zones with code red does not necessarily mean more curtailment than a single zone with code red. Besides, the ANM zones have differing capacities for generation, with the majority of larger turbines located in the northern and western parts. This means we have no precise way of quantifying the amount of curtailment, and by using this as truth values, we are only able to model whether curtailment is present, but not how much power is being lost.

4.4 Summary

In this section we have described our dataset, which we constructed by collecting publicly available data every 10 minutes for 3 months from December 2018 to February 2019. We have collected data on the state of the Orkney grid from the SSE website through web scraping, and weather data through the OpenWeatherMap API, supplying with bulk historical data.

We have also described noteworthy limitations and biases of our dataset. Data is only collected every 10 minutes, and the time period of our dataset is the three winter months, resulting in a significant seasonal bias. Finally, we also discussed the data we will be using as truth values, and its limitations, making it difficult for us to quantify the amount of curtailment. It is important to note these issues, as they affect what our model is actually modeling.

5 A Simple Descriptive Model

In order to understand how the ANM curtailment scheme works, we will first try to develop a purely descriptive model. This kind of model can take as input any type of contextual data, current or historical, and the output is whether to issue curtailment at that time.

Based on the maximum transmission capacity of 20MW out of Orkney, a simple curtailment model GD (for **G**eneration and **D**emand) could be the Boolean expression:

$$GD(t) \iff G_t > D_t + 20 \tag{1}$$

Where $GD(t)$ denotes whether to issue curtailment at time t, G_t is renewable generation at time t, and D_t is the power demand at time t, both of which are expressed in megawatts.

This is a purely theoretical model, only looking at how a perfectly balanced latency-free system could work. And when evaluating this model it becomes clear that this simple rule does not model the operation of the ANM very well.

5.1 Evaluating the Simple Model

We evaluate the model against our dataset, consisting of 12658 samples, at each sample recording whether the output of the model matches the truth value of the dataset, and finally calculating the accuracy, meaning the percentage of times the model was correct.

Model Name	**Accuracy**
GD	41.33%
Never Curtail	41.13%
Always Curtail	58.87%
GD_5	**74.76%**

Table 1: Accuracy of descriptive models

When evaluating the $GD(t)$ model, we see that the accuracy is only 41.33% (see Table 1 for all results), which must must be considered quite poor. For comparison a model always outputting False has an accuracy of 41.13%, and conversely a model always outputting True has an accuracy of 58.87% ("Never Curtail" and "Always Curtail" respectively in Table 1), which is considerably better than our model. Clearly, this is not the threshold the ANM operates by. However, the overall model structure might still work, if we just change the constant a bit. A more generalized version could be:

$$GD_k(t) \iff G_t > D_t + k \tag{2}$$

When we evaluate this model with different values for k, shown for values $[-20; 20]$ in Figure 3, we see that the plot peaks at $k = 5$ with an accuracy of 74.76%.

As we can see on the slope of the plot in Figure 3, our results do not indicate that the ANM uses a sharp threshold of this sort to issue curtailments. But as the accuracies are still relatively high, we assume that this type of model could still be very useful. In order to understand the ANM and our data better, we will explore the data further through visualization.

6 Visualizing the Data

To improve our descriptive model, we want to use our dataset for more than just evaluating our models — we want to extract the knowledge and patterns in the data and use it to design new models. We do this by creating

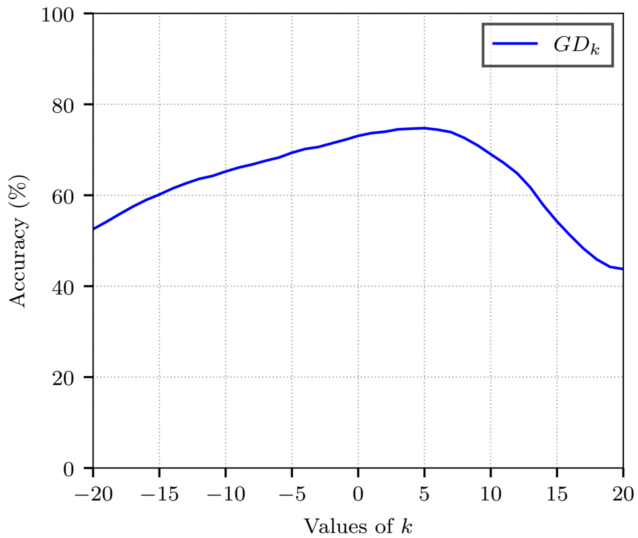


Figure 3: Accuracy of $GD_k(t)$

graphs and plots presenting an overview of the contents of our data from different angles, and from this we can identify correlations that can later be quantified numerically.

These visual representations are made using the MATLAB-inspired plotting library Matplotlib [35], which features tight integration with Pandas dataframes, making plotting possible without extensive data wrangling. Regressions are calculated through the machine learning library Scikit-Learn [36].

6.1 Timeline graphs

To ease the exploration of our data and look at concrete incidents, we have set up a web service allowing visualizing of the data as a series of graphs.[8] In Figure 4 we show the data collected for the week from the 14th to the 20th of January 2019 as an example, showing the curtailment in individual zones as horizontal tracks and the difference between generation and demand as a black line with gray fill. By looking at the graph we can extract several **clues**:

1. There seems to be a strong correlation between curtailment and periods where generation exceeds demand by more than ~10MW.
2. It looks like curtailment is often issued while generation is rising rapidly.

[8] As of May 31st available at `http://curtailment.net/`, implemented using Flask, a web development framework for Python. Source code is available at `https://github.com/NielsOerbaek/ANM-Graph-Tool`

3. It looks like most curtailment is happening in Zones 2, 2A and 2B, with 2A often being fully stopped.

4. Curtailment is not necessarily long lasting. Especially in Zone 2, where curtailment is imposed and lifted several times within a single hour.

However, Figure 4 also raises some **questions**:

1. On the evening of the 18th of January, generation rises rapidly and exceeds 7-8MW above the demand. However, no curtailment is issued. How come? What makes this situation different? How often does this happen?

2. On the 16th and 19th of January, short periods of curtailment are issued, and they are also issued in some of the zones that are not curtailed when generation is high. How come these zones are curtailed when generation is low?

Besides the data shown in Figure 4, several interesting situations occur in the data. Often, when generation is high, other zones are curtailed as well. Also, zones are sometimes curtailed for prolonged periods of times, seemingly not affected by changes in generation and demand. For example, Core Zone had a code red for over 10 days straight from the 17th to the 27th of February, while the generation fluctuated between 22MW below and 17MW above the demand.[9]

In general, we can see a pattern suggesting that most curtailments are issued when generation exceeds or is about to exceed a certain limit above the demand. But this is not all-encompassing; there are several incidents seemingly not following this rule at all.

This is an example of both finding answers and questions through data exploration. Some of the patterns we identify will not be addressed by our models, and some of the questions we have raised will not be answered. As expressed in the aphorism *"all models are wrong, but some are useful"*[37], we do not expect to construct a perfectly accurate model that describes the ANM with all its aspects. We do, however, wish to develop more general models that are useful for identifying curtailment. Still, our assumptions and omissions are important to keep in mind when both evaluating and applying our models.

6.2 Wind Speeds and Orkney Energy Production

As one would imagine, wind energy generation is highly dependent on wind speeds, and for a single wind turbine a normal way of describing the relationship between wind speed and power output is a *power curve* [38, 39].

[9] A graph of this can be seen in Appendix A, Figure 24

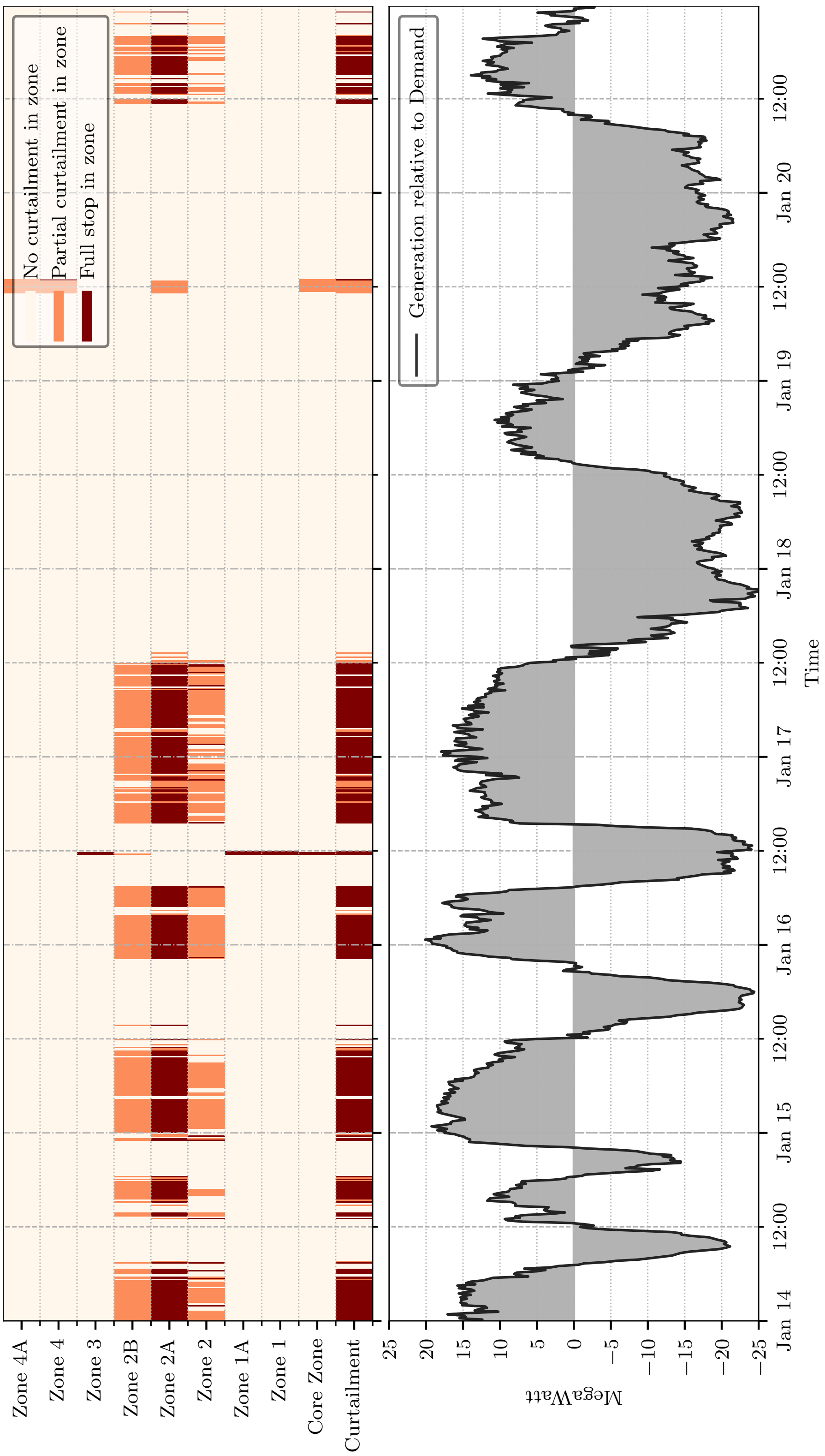


Figure 4: Correlation between (*Generation* − *Demand*) and curtailment in the individual zones. Each zone is shown in individual horizontal tracks corresponding to the labels on the upper y-axis. The black line shows the generation relative to the demand for all of Orkney.

A power curve is constructed by grouping all of the power output measurements by wind speed and taking the median of each group. Since our dataset is fairly limited in size, we round the wind speeds to the nearest integer before grouping.

Since the primary power source in Orkney is wind energy, we expect our data to show a strong correlation between wind speeds and renewable energy generation. In order to verify this, we make a scatter plot of the wind speed and generation data in our dataset. This plot is shown in Figure 5, on which we also show the calculated power curve.

It is clear from this plot that generation is often high when wind speeds are high, but our data also shows very high generation when wind speeds are near zero, which is highly unexpected. Another thing that stands out is the horizontal grouping of a lot of the data points, appearing as vertical lines on the plot.

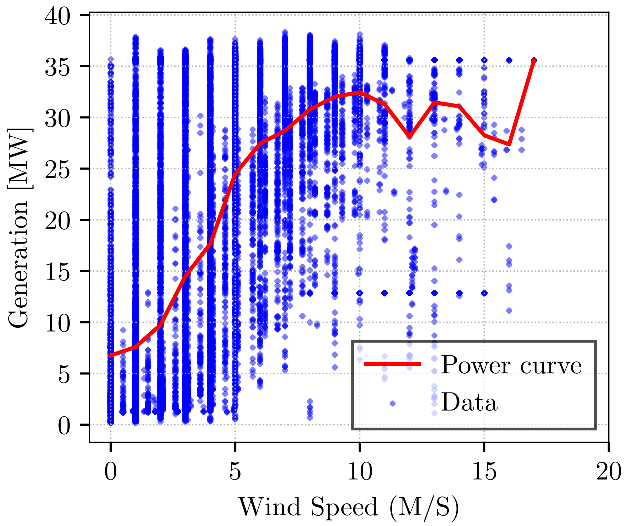


Figure 5: Scatter plot showing relationship between wind speed and power generation.

Upon nearer examination of the dataset, we find that the bulk download of historical weather data only supplies wind speeds as integers, automatically grouping the data points together. Still, this does not account for the low-wind-high-generation data. To check the validity of the historical data downloaded, we check a single data point falling into this category: December 31st 2019 00:00:00. According to the data scraped from the SSE website, generation at this time was ~35MW, but according to our weather data the wind speed was only 3 m/s. When checking against `timeanddate.com`[10],

[10]See `https://www.timeanddate.com/weather/@2640298/historic?month=12&year=2018`, last accessed March 7th, 2019

another web service that allows users to search for historical weather data (but exposes no API), the wind speeds recorded for this time and place was 10 m/s, which seems more plausible, considering the energy generation at the time.

This leads us to suspect that the bulk data we have used is not sufficiently accurate, and draw the same plot using only the data we have collected through the OpenWeatherMap API since the 12th of February 2019. This plot is shown in Figure 6, and it matches our expectations much better. Although the data is still scattered, it is clear to see that energy generation is low when wind speeds are low, and higher when wind speeds are higher.

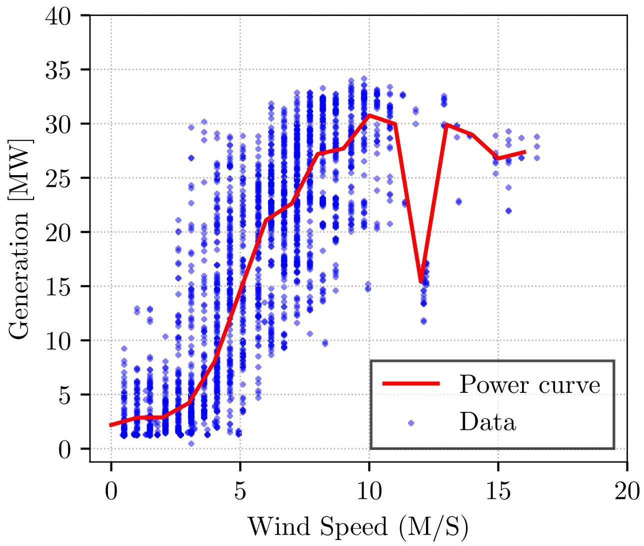


Figure 6: Scatter plot showing only weather data from API. The full line shows the collective power curve for the Orkney generation.

Here it is worth noting that although we expected a correlation, we did not expect as close a correlation as in a normal power curve. This is due to the fact that Orkney has other energy sources that are not coupled to wind speeds (solar, tidal, and in part also wave), and the wind turbines in Orkney are both diverse in type and scale and distributed across the islands, where the weather can differ slightly from our single wind speed value.

We calculated the discrete values that constitute a combined power curve for all of Orkney, also shown in Figure 6, which can then be expanded to cover all real-valued wind speeds through linear interpolation.

The few high-wind-low-generation data points can be attributed to the storm control mechanism of many wind turbines, shutting down the generation when gusts of wind go above a certain threshold. These data points affect our power curve quite a bit, especially at 12 m/s, and it might be useful to sacrifice this portion of the data in exchange for a smoother curve that works well most of the time. We therefore limit our dataset to wind speeds ≤ 11 m/s, which is still covering 96.8% of the data, and estimate an output of 30 MW when above this.

We also investigated the relationship between generation and a series of other factors, including wind direction, outside temperature, and atmospheric pressure, but no significant correlation was found.

6.3 Time, Date and the Energy Demand in Orkney

The power demand of the Orkney grid is far more constant than the fluctuating generation, with over 90% of our demand data in the interval $[15, 25]$MW. In fact, the standard deviation[11] of our demand data is $\sim$3.1MW, while it is $\sim$12.0MW for generation. Still, we assume that demand impacts curtailment, and therefore it is relevant to explore this aspect as well.

As mentioned in subsection 2.2, the course of the day is a pattern often used for scheduling fixed DR policies, as there is an observable correlation between the energy use of private consumers and the 24 hours in a day. Based on this, we examine if such a correlation is also present in our data by again producing a scatter plot, shown in Figure 7, and looking for patterns.

While not as prominent as the correlation between wind speeds and energy generation, a pattern is still visible. The energy usage somewhat follows the course of the day. However, we see a larger spread in the data for certain hours, so we suspect more correlations to be at play.

Another typical temporal correlation for energy usage is the days of the week, and to investigate this we group the data in Figure 7 by weekdays as shown in Figure 8.[12] Here we can see that most workdays (solid lines) act the same, but the weekend (dashed lines) sticks out in especially one place: Lower demand in mornings from 8:00 to 11:00. Besides this, Mondays and Tuesdays seem to indicate high demand, especially in the afternoon.

Instead of using regression to find equations that approximate these lines, we will instead create a lookup table of weekdays and hours, allowing us to get an estimation of the demand at the nearest hour of any given point. The table is then a 7×24 matrix, and again linear interpolation can be added for finer-grained values if needed, but will be omitted in this version.

This is one of the places where the seasonal bias mentioned in subsubsection 4.3.2 is extremely relevant. Besides just time and place, the energy demand is expected to correlate tightly with factors such as outside temperature[13] and hours of daylight, both factors that are affected by the time-of-year. Since we only have data from the winter months, this distinction is

[11] The standard deviation is defined as $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$

[12] Simply plotting weekdays against demand gives no visible patterns. This graph is shown in Figure 22 in Appendix A

[13] Only a slight correlation was found between temperature and energy demand, plotted in Figure 23 in Appendix A

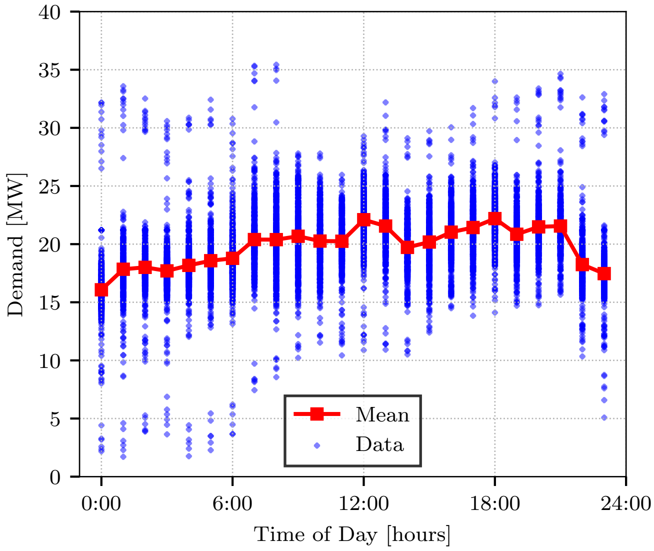


Figure 7: Scatter plot showing relationship between power demand and time of day grouped by hour

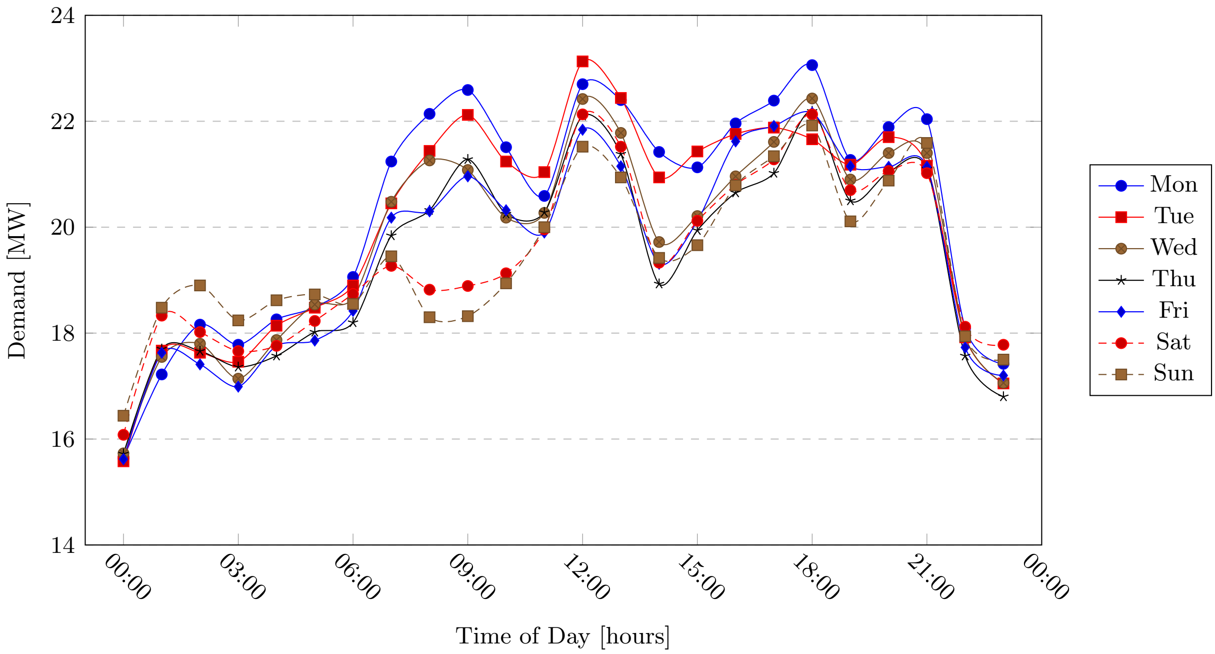


Figure 8: Median demand per hour per weekday

not relevant for us, but if we had data for a full year or more, it would make sense to add a third dimension of "season" or "month" to the lookup table.

6.4 Summary

In this section we have visualized the data in our dataset, looking for correlations, verifying our assumptions, and even finding some erroneous data in our historical weather data. We have identified some patterns and correlations affecting the behavior of demand and generation in the Orkney grid.

We have seen that most curtailment correlates to generation relative to the demand, but that certain cases seem completely unrelated. This leaves us with a choice to make. Ideally we would love to make a model that describes the whole system's behavior in a few simple rules, but that is an unobtainable goal. So we need to decide whether to embrace the complexity, or to make a simpler model that works most of the time. We choose to keep our models simple, modeling only certain aspects of the system, and adding complexity only if needed and useful.

We have also found correlations for the two variables in our initial simple descriptive model, allowing us to model the state of the grid using data external to the grid itself, which we will explore in the following section.

7 A Correlation-based Model

As we saw in the previous section, there are relationships between wind speeds and the energy generation in Orkney, and between time and energy demand. The next logical step is then to change our simple model from section 5 to be based on these correlations, and from Equation 2 derive a model WT (from **W**ind and **T**ime):

$$WT_k(t) \iff WG(W_t) > DM_{w_t,h_t} + k \tag{3}$$

Where W_t, h_t and w_t is the wind speed, hour of day, and weekday at time t, respectively, DM denotes the lookup matrix, and WG is the interpolated power curve from Figure 6.

7.1 Evaluating the Correlation-based Model

When evaluating this new model using the same method as in subsection 5.1, we see that the correlation-based model is performing significantly worse, with a peak accuracy of 66.0% at $k = -9$. These results are listed in Table 2 and shown as the full lines in Figure 9.

Recalling that large parts of the weather data from our dataset seemed erroneous, we instead evaluate our two models on only the part of the dataset where weather data was collected from the open API, which consists of 2248

Model Name	Accuracy
GD_5	74.76%
WT_{-9}	66.0%
GD_{-20}	72.96%
WT_{-20}	73.77%

Table 2: Accuracy of descriptive and correlation-based models

samples collected over 17 days from February 12th to March 1st. These results are shown as the dashed lines in Figure 9.

In this evaluation $WT_k(t)$ performs significantly better (73.77% at $k = -20$) than with the full dataset, but $GD_k(t)$ performs slightly worse. Looking at the graph, the two dashed curves behave somewhat similarly, which is what we would expect, since the new model is based on approximations of correlations. Still, the results are not satisfactory and both models peak at $k = -20$, the lowest value of k in our experiments, which leads us to look a bit closer at this part of the dataset.

Limiting the dataset to only 17 days allows a single incident of exceptional data to play a much larger role, and as mentioned in subsection 6.1 there was an exceptional situation from the 17th to the 27th of February (shown in Figure 24 in Appendix A), where the ANM status reported a continuous full stop in Core Zone.

It becomes apparent that we need to look at systematically cleaning our dataset, if we want to derive and evaluate a usable model.

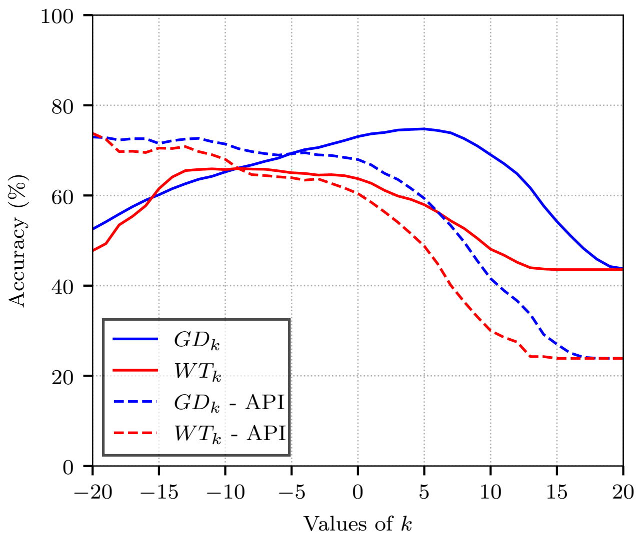


Figure 9: Accuracy of correlation-based curtailment model

8 Data Cleaning

As we saw in the timeline graphs in subsection 6.1, most of the curtailment situations follow the relationship between generation and demand in the grid as a whole. However, several incidents seem completely unrelated to this. And while it seems improbable that these curtailments were purely coincidental, they might be caused by factors that we do not expect our model to be able to encompass, like maintenance or technical issues with the grid.

For example, from around 20:00 on March 3rd, to about 8:00 the following morning, all zones in the ANM reported a full stop. This was despite generation being 5-10MW lower than the demand.[14] Our contacts in Orkney explained to us that this incident was caused by a "power cut", presumably an exceptional situation outside the normal functionality of the ANM. They speculated that a "code red" in the ANM status could also simply mean that the generator was down, not necessarily with the purpose of reducing generation.

Situations like this are outside the intended scope of our models, but in order to remove these anomalies from our dataset, we need to first identify them. In the following section we will define methods and rules for classifying anomalies and determining what to do with them.

8.1 Defining Anomalies

Cleaning anomalous incidents from our data can have a dual function: It can both help us evaluate models that are more generalizable and work better most of the time, but it can also help us identify the situations where the ANM functions counter-intuitively, against our basic idea of how it should work. We can then quantify these situations to get an idea of how often curtailments happen, even though we would not expect them to.

If we choose to exclude these situations from the model domain, we need to remove them from our dataset, and then we are faced with a new problem: How do we detect these anomalies? If we are not careful, we might very well indirectly end up with a rule of *"if it does not fit our model, then it is considered an exception"*, redact them from our dataset, and *voilà*, our model now has 100% accuracy.

Still, our intuitive understanding of the purpose of the ANM tells us that curtailments should not be happening for prolonged periods of time where demand is higher than generation. A possible model-independent rule could be *"if a curtailment situation stretches for six consecutive hours or more where demand is higher than generation, then it is considered an exception"*. In this case a *curtailment situation* defines a period of time from the status changes to yellow or red, until the status becomes green again. This is done

[14] A timeline graph of this incident is shown in Figure 25 in Appendix A

on a per-zone basis, and the 6-hour limit is chosen somewhat arbitrarily with the reasoning that we want the rule to apply only to exceptions of a significant size. From here on we will refer to this rule as the *6-hour rule*.

Another way of detecting anomalies is to define a *de minimis level*, a lower threshold for what we consider curtailment. The only way of quantifying the amount of curtailment we have at our disposal is the amount of zones that are affected, a measure that is both coarse and potentially imprecise (See the discussion in subsubsection 4.3.3). However, from looking at the data, we can see that often in the situations where we would expect curtailment, more than one zone is affected, indicating that this might be a suitable way to detect anomalies after all. Our de minimis rule is then: *"If at any point in time, two or more zones in the Orkney ANM are curtailed, we consider it curtailment, otherwise we do not"*. This method only affects the reduced view of curtailment, meaning whether there is curtailment in the Orkney ANM as a whole, and not in our data for the individual zones. This is worth noting, should we try to model the behavior of the ANM in each zone at a later time.

Finally, we look for time periods where the data seems incorrect or erroneous. This could be holes in the data or unexpected and unexplained spikes. We identify these periods by visual exploration, plotting the data on a timeline and looking for anything unusual.

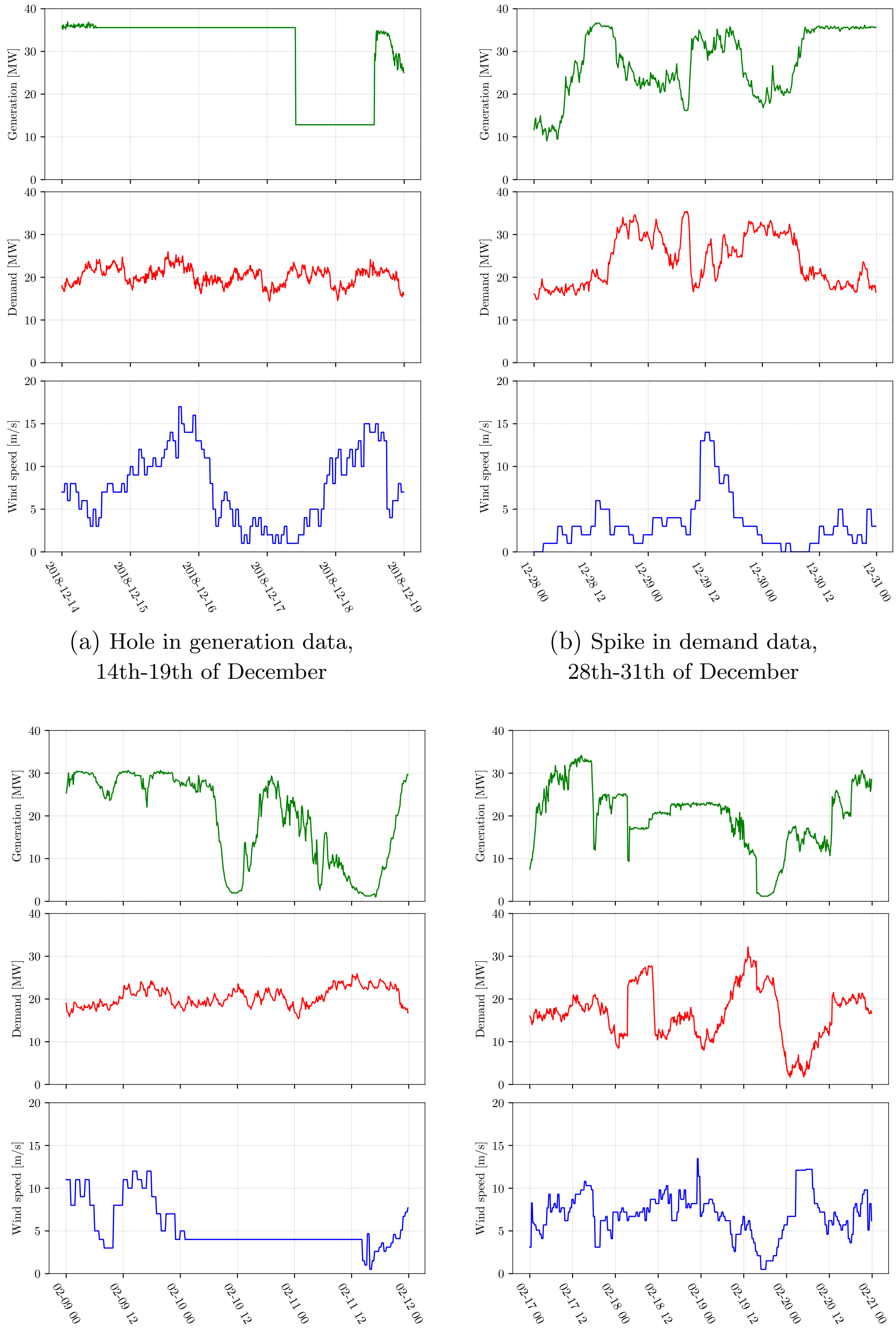
We identified five anomalous periods, of which four are shown in Figure 10: one hole in the wind data (Figure 10c), one hole in the SSE data (Figure 27 in Appendix A), one case of erroneous values from the SSE (Figure 10a), and two cases of unexpected spikes in the demand data (Figure 10b and 10d). These unexpected spikes in demand are larger than what we think is plausible, and they seem to coincide with drops in the more fluctuating generation, leading us to suspect that they are caused by faulty readings and should be classified as anomalies.

8.2 Cleaning Anomalies

After identifying the anomalies that need to be cleaned from our data, we can handle them in two different ways: We can either delete the samples that contain anomalous data, making our dataset smaller, or we can alter the data, trying to change the data to a more normalized state.

For the situations detected through the 6-hour rule, we choose to alter the data by changing the ANM-status in the given zone from red or yellow to green. And for the de minimis rule we recalculate our Orkney-wide curtailment labels to be only true if two or more zones are curtailed at a given time.

Altering the data in this way is reasonable since we are working with truth as a binary classification, meaning that there is either curtailment or not, with no in-between. So if our detection says that this classification



(a) Hole in generation data, 14th-19th of December

(b) Spike in demand data, 28th-31th of December

(c) Hole in wind data, 10th-12th of February

(d) Spike in demand data, 17th-21th of February

Figure 10: Periods with exceptional data

might be an anomaly that we do not want in our dataset, there is only one other option, making the actual alteration quite trivial.

For the visually identified erroneous time periods we are working with real-valued data, and altering these values is a lot more complex. We could try to interpolate across gaps or smooth out peaks, but our altered data would only be guesses. Instead we choose to remove these sections completely from our dataset, reducing its size.

8.3 Quantifying Anomalies

Now that we have defined rules for identifying anomalies and exceptional situations which we do not want to evaluate our model against, it is interesting to look at how much curtailment there is, and how much of this we define as anomalies.

When evaluating our methods for anomaly detection we will differentiate between the original dataset, or *full set*, and the dataset after manually removing the anomalous time periods, which we will call the *redacted set*. The amount of curtailment after applying our two anomaly detection rules to each set is shown in Table 3.

The full set covers 90 days, which is 2160 hours, and for 1219 of those hours, or 56.45% of the time, some curtailment was reported in the Orkney ANM. When applying the 6-hour rule, the curtailment time decreases to 801 hours, or 37.08%. When only applying the de minimis rule, the amount of curtailment is reduced to 842 hours, or 39%, and when applying both of these rules the amount is further reduced to 716 hours, or 33.14%. From this we can conclude that the two sets gained by applying each of our two rules have a great amount of union.

When excluding the identified anomalous periods, we choose, for simplicity, to split the data by whole days. In total we exclude 17 days, reducing the size of our dataset by 18.89%. The redacted set, now covering 73 days,

	Full Set		**Redacted Set**		
Method	**Hours**	**%**	**Hours**	**%**	**% of Full Set**
No Cleaning	1219	56.45%	878	50.16%	40.66%
6-hour Rule	801	37.08%	658	37.59%	30.48%
De Minimis	842	39.00%	636	36.33%	29.45%
Both	716	**33.14%**	587	33.53%	**27.18%**

Table 3: Amounts of curtailment in the datasets after applying different methods for anomaly detection and data cleaning.

or 1752 hours, has 878 hours of curtailment, corresponding to 40.66% of the full 90 days. As with the full set, we see a significant effect with each of the methods, and an only slightly greater effect by combining the two methods, resulting in a curtailment ratio of 27.18% when applying both methods on the redacted set.

The fact that the two anomaly rules have a significant union set, can be interpreted as supportive for the validity of our results, as the majority of the anomalies are detected by both of our intuitively defined rules.

From our results we can see that $\frac{1219-587}{1219} = 51.85\%$ of the curtailment in our dataset is what we define as anomalies, of which only a fraction is caused by our own data collection process. These results only the describe the ratio of time with curtailment, and does not necessarily reflect the amount of power lost due to curtailment. Still, we believe that our findings in this section indicate that the ANMs curtailments are not purely based on efforts to balance demand and generation.

These methods are applied in the preprocessing of the data. For a visual comparison, Figure 26 in Appendix A shows the same time period as Figure 24 after applying the two anomaly detection rules.

8.4 Evaluating with Cleaned Data

When redoing the evaluations shown in Figure 9 on the cleaned data, we see the accuracies rise significantly, as shown in Figure 11. The accuracies for the optimal values for k are shown in Table 4, peaking at ~89.5% for the GD-based model, and ~84% for the WT-based models, when only using the

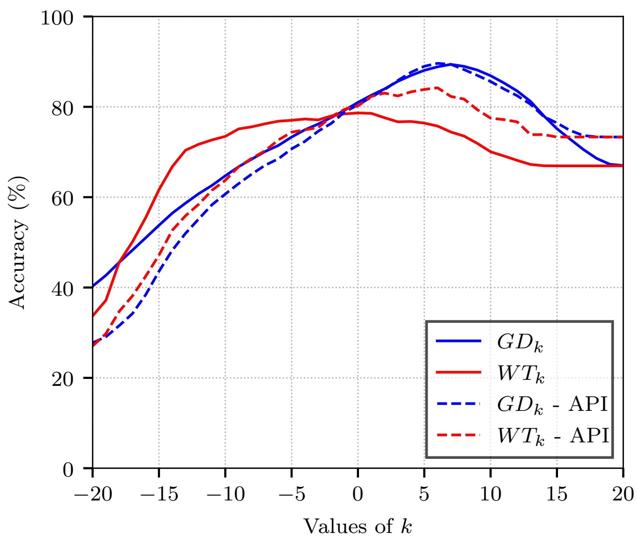


Figure 11: Accuracy of correlation-based curtailment model evaluated on cleaned data

Model Name	Accuracy
GD_7 - All data	89.64%
WT_0 - All data	78.55%
GD_6 - Only API data	89.89%
WT_6 - Only API data	84.06%

Table 4: Accuracy of descriptive models evaluated on cleaned data

weather data collected through the API.

From this point on, unless explicitly stated otherwise, we will be using the cleaned curtailment data when evaluating models.

8.5 Summary and Discussion

In this section we have made efforts to clean our data by identifying anomalies and removing them. This includes both technical anomalies, caused by errors in the data collected or in our data collection itself, identified by visual exploration, and anomalies in the curtailment patterns, which we deem outside the scope of our model. We identify anomalies by formulating an intuitive rule and by applying a de minimis level. The results of applying these two methods to our dataset have a significant union set, strengthening our confidence in the combined method's validity. We handle the detected anomalies differently based on data type, altering the ANM status when it matches one of our two rules, and removing complete days from the dataset when the real-valued data seems erroneous.

After cleaning our data, the ratio of time with some curtailment in the ANM decreases by 632 hours, or 51.85%. And when evaluating our models on this cleaned dataset, we see improvements in accuracy of over 10 percentage points for both model types. This indicates that our anomaly detection scheme has been able to identify many of the cases of curtailment that cannot be described by our model, resulting in accuracy gains, while also enabling us to quantify the amount of curtailment we do not understand.

It is worth noting that we only quantify curtailment in hours, and not in watts. We suspect that the relative amount of power-loss cleaned from the dataset is far lower, as we primarily alter data where only one zone is affected.

9 Using Neural Networks

One of the primary trends in wind power and demand forecasting research is the use of Artificial Neural Networks (ANNs), which have produced state-of-the-art results for many applications [34, 33, 40, 20, 23, 41]. The primary benefit of these machine learning and deep learning algorithms, is that they are able to "learn" their own parameters by training on datasets of samples and truth labels. The network approximates the relationship between sample and truth by adjusting matrices of weights through the backward propagation of errors. In this section we will train a series of neural networks with the dataset we have created for the Orkney power grid, discussing both the accuracy of the resulting models and the implications of this technique.

As machine learning and ANNs are not the key focus of this project, we will not spend a lot of time explaining these concepts. Instead we refer to [42] and [43] for an introduction to this subject.

9.1 Implementation

To implement the neural networks, we again use the programming language Python 3 and the library Keras [44]. Keras is a high-level frontend for several popular deep learning libraries, allowing for quick prototyping of many of the most common network architecture types. We use Keras with TensorFlow [45] as the backend that handles the tensor calculations. Python has become the de facto standard for deep learning research, and Keras and TensorFlow are the two most popular libraries in their category [46].

Our reason for choosing this setup, is that the types of networks that we will experiment with are standard architectures that are built into Keras. This allows us to define a relatively complex network with hundreds of nodes and thousands of parameters in just a couple of lines of code, and reduces development overhead during experimentation. If custom behavior is needed, we can still access the TensorFlow backend and manipulate the tensors directly.

9.2 Experiments

In these experiments we evaluate networks of two architectures: A single *perceptron*[15] (sometimes also called *neuron*), and a *feed-forward neural network* (FFNN) with 3 fully-connected hidden layers and 64 perceptron-nodes in each layer. A single perceptron is only able to learn linear regressions and classifications, while a deeper network with several hidden layers is able to model more complex behavior, hence the name *deep learning*.

Initially we create two ANNs, based on the same inputs as the GD-model from Equation 2 and the WT-model from Equation 3, respectively. When designing the architectures of neural networks, it is important to create a model with a suitable complexity, here referring to the amount of weights. If the network is not complex enough, it is not able to properly model the system, but if it is too complex, it might hit the pitfall of *overfitting*, that

[15] Technically not a neural *network*, as there is only one node, but we will not make that distinction here. See [47, Ch. 4] for more info

the network learns to give the right output on exactly the input it is trained on, but not learning anything general, making it unsuitable for classifying unseen data.

In order to detect overfitting, we have to evaluate the models on unseen data, but still we would like to both train on all data and evaluate on all data. A technique to accomplish this is *k-fold cross-validation*, in which the data is shuffled and split into k partitions, and over k iterations each partition is held as the validation set, while the rest is used for training. For larger datasets and more complex models, this can be infeasible, as training becomes too time-consuming, but for our situation it is very suitable. In fact, because our dataset is so small, we can increase the confidence in our results by doing 2 iterations of 10-fold validations, with different random seeds for the shuffling.

We train all models for 30 *epochs*, meaning 30 iterations over the training data, using the RMSProp optimizing algorithm, and with binary crossentropy as the loss function.

9.3 Results

The results from the cross-validation evaluations are shown in Table 5, and from the mean accuracy of the 20 folds, which is the traditional metric for this type of evaluation, we see that all of our ANN-based models perform worse than their manually constructed counterparts. Upon further investigation, we discover that the standard deviation of the individual accuracies are quite large for both of the GD-based models. And if we look at the per-fold results (shown in Table 10 in Appendix B), we see a few folds with accuracies at around one and two thirds of the rest.

Model	Mean	STD	Median	Min	Max
$GD_7(t)$	89.64%	-	-	-	-
GD Percep	85.84%	±13.87%	90.09%	32.44%	91.04%
GD FFNN	89.10%	±5.16%	90.42%	67.78%	91.71%
$WT_0(t)$	78.55%	-	-	-	-
WT Percep	78.19%	±0.94%	78.27%	76.55%	79.69%
WT FFNN	77.91%	±1.71%	78.27%	73.66%	80.46%

Table 5: Accuracy of ANN-based models on entire dataset. The ANN-based models are evaluated with double 10-fold cross validation

Since our dataset is relatively small, it raises the chances for "unlucky" folds, where the data in the test set does not represent the data in the training set very well. When looking at the median, the accuracies are both higher and more stable. Making a boxplot from the per-fold results, as shown in Figure 12, also highlights how outlying these outliers actually are.

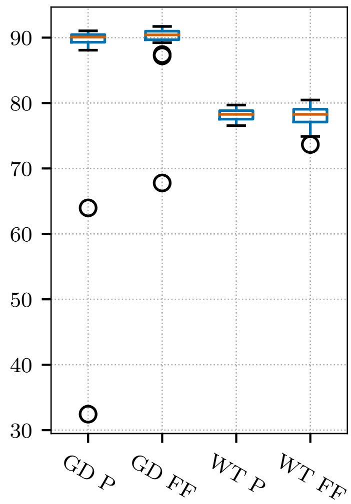


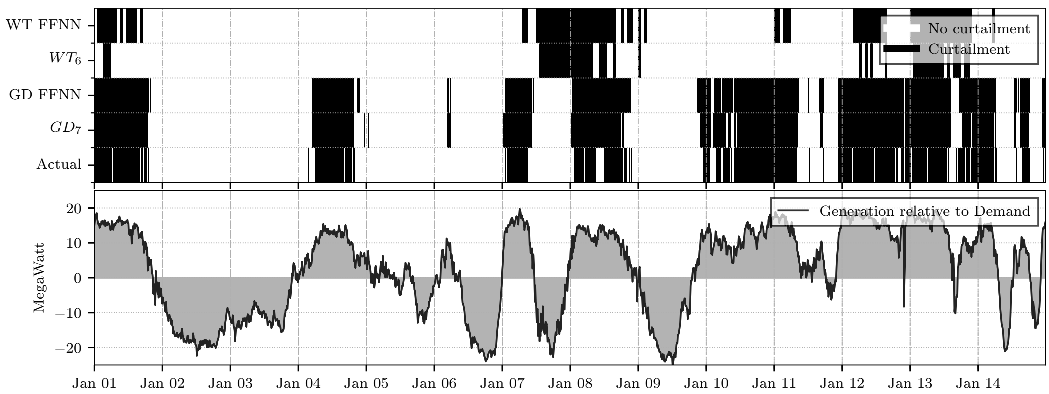
Figure 12: Boxplot of the fold accuracies

By training the models on our entire dataset, and evaluating it on the same data, we get slightly higher but similar results, shown as the second column in Table 6. This indicates that overfitting is not a significant issue for our models.

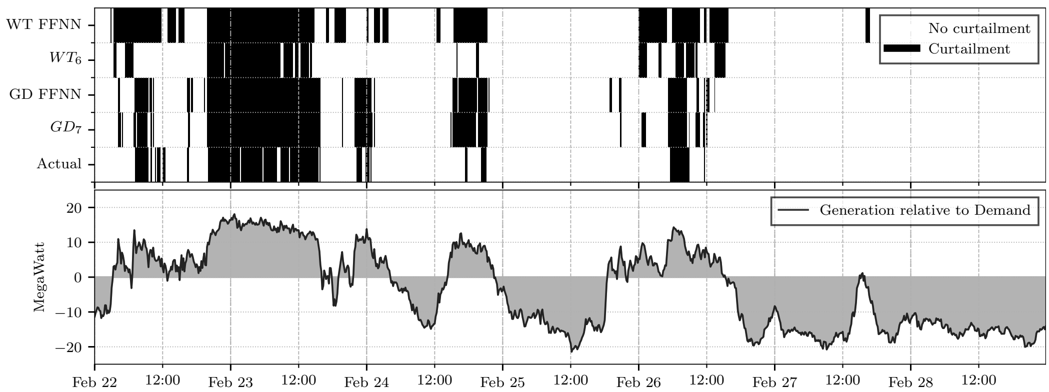
One of the lures of deep learning is that, when the dataset is large enough, the job of feature selection becomes less important, since the network will detect if some of the inputs are irrelevant for the truth value, and disregard them. Even though our dataset is quite small, we tried including more features in the WT FFNN, to see if it had an impact on the accuracy. In addition to the wind speed, hour and weekday, we also included wind direction, atmospheric pressure, temperature and month. The mean accuracy of this model was 49.82%, about as good as a coin toss, indicating that selecting features more carefully is more suitable in our case.

To see how the models behave on concrete data, we have included three timeline plots comparing the output of the different models, shown in Figure 13. We have chosen three periods that we deem interesting: One from before we started collecting weather data from the API (Figure 13a), one after (Figure 13b), and one with unseen data (Figure 13c). The numerical accuracies for these periods are also listed in Table 6.

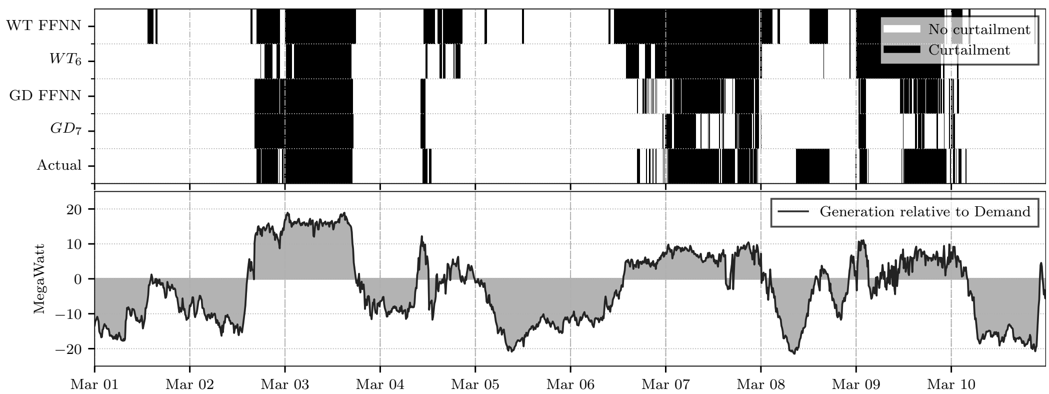
From these figures we can extract several insights, some of which confirm what we already know: We can see that both of the WT-models perform poorly in Figure 13a as it contains faulty weather data. The WT FFNN consequently underperforms WT_6, which is expected, since it has been trained primarily on this faulty weather data. And finally, we can see that the GD FFNN outperforms GD_7, also on the unseen data, again an indicator that the model is not overfitted. We can also see that the WT-FFNN is much more likely to output curtailment than WT_6, while GD_7 and GD FFNN seem more similar except for in Figure 13c, where the FFNN outperforms GD_7 significantly.



(a) 1st - 14th of January



(b) 22nd - 28th of February



(c) 1st - 10th of March

Figure 13: Timeline comparison of different models for different time periods

Model	**Full**	**Figure 13a**	**Figure 13b**	**Figure 13c**
GD_7	89.64%	90.72%	91.67%	84.38%
GD FFNN	**91.03%**	**90.82%**	**91.96%**	**89.10%**
WT_6	75.45%	62.15%	87.40%	83.40%
WT FFNN	78.36%	69.99%	79.96%	79.58%

Table 6: Numerical accuracies for the full dataset and the plots in Figure 13

9.4 Summary and Discussion

In this section we have used ANNs to model the ANM, and evaluated these models through statistical analysis of their experimental accuracies. We found that utilizing the technique of machine learning for this problem yields only slight accuracy gains, but it does have other aspects that makes it interesting for our case.

Since the models are learned from the data, it is trivial to adapt them to new incoming data, as we can simply train a new network with the extended dataset. If the dataset becomes large, this would be a computational burden, but with the amount of data we collect, it would take many years for dataset to grow to a size where this is a problem. Another advantage of these models, is that their output also describes how well the input fits the model, which can be used as certainty. A very certain output would be closer to 0 or 1, while a more uncertain output will be closer to $\frac{1}{2}$.

The use of ANNs can also potentially work as a way to detect overlooked patterns. If our neural networks had resulted in much higher accuracies, it would be an indication that there are patterns in the data that we have overlooked or misunderstood. But, since the models we train are defined by matrices of unnamed parameters, it becomes hard for us to extract any tangible knowledge about the rules governing the ANM. The ANN-models simply mimic the data they are given, and their functionality is not human-readable. And for orcadians, the understanding of how the ANM functions might be more important than model accuracy, as it affects the further development of renewable energy on the islands.

In short, simply piping our data through a neural network could not replace the EDA-process carried out through this project. It does not give the same insights into the system, even though it yields similar results. But the two methods can compliment each other, and we expect ANNs to become even more useful as the dataset grows larger and more patters (e.g. by month or temperature) emerge.

10 Case: A Single Wind Turbine

In this part, we will look at a dataset produced by a single 900kW wind turbine in Orkney. By exploring this dataset, we can gain insights about the concrete workings of generator curtailment, and we can use our collected data to shed new light on the models used by the turbine administrators.

Orkney Wind Energy (OWE) is a company based in Orkney[16]. The company was founded by a local community and the main purpose of the company is to administer wind turbines.

This wind turbine was erected after the introduction of the ANM system, and as such was given a non-firm grid connection, meaning that it can be curtailed. And curtailment has been a major issue for OWE, with an average energy-loss for the year of 2018 of over 40%.[17] This loss is not caused solely by the ANM, as it also includes technical faults in the turbine or grid, but according to OWE the majority is from ANM-based curtailment.

To help us relate our Orkney-wide dataset to the data of a single turbine, OWE has allowed us access to the data reports generated by their turbines SCADA-system. These are monthly reports consisting of key measurements recorded at 10-minute intervals, of which the most relevant to us are wind speed and power output. We have data reports corresponding to the three months in our dataset.

10.1 Plotting the Power Curves

In order to get an overview of the statistical power curve of the turbine and possible signs of curtailment, we plot the measured wind speeds against the power output, both of which are the the mean over 10 minutes of measurements. This plot is shown in Figure 14.

At first glance we notice the predominant power curve rising to the maximum output of 900 kW at around 13 m/s and then decreasing again at around 25 m/s, presumably due to the turbines built-in storm control feature. Besides this, we notice two clear horizontal lines, one at 0 kW and one around 675 kW, which we suspect is due to curtailment.

In order to verify this, we can join the OWE dataset with our Orkney-wide dataset on the time and date, and filter the measurements on whether the SSE reported any curtailment at that time. OWE is located in Zone K in the ANM, so we can limit our filtering to only that zone. However, when looking at the resulting plots, shown in Figure 15, we see that most of the measurements in Figure 15a are located where wind speeds are above 10 m/s, but the horizontal lines are not limited to this plot, and there is no apparent connection to the SSE-reported curtailment.

[16]Company name and locations are fictional, as some of the data used is commercially sensitive. The real names are known to the author

[17]These numbers are from an annual curtailment report supplied by OWE

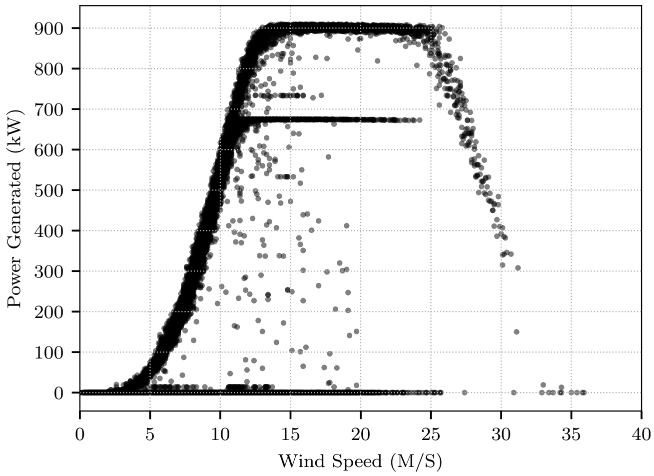
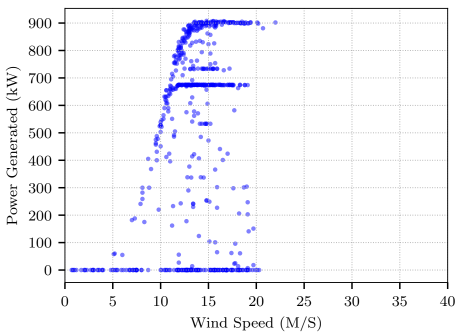
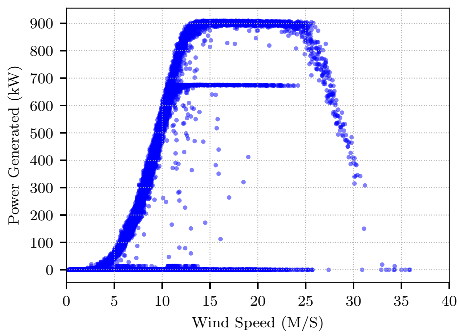


Figure 14: Scatter plot showing relationship wind speeds and power output of OWE turbine



(a) Curtailment in Zone K



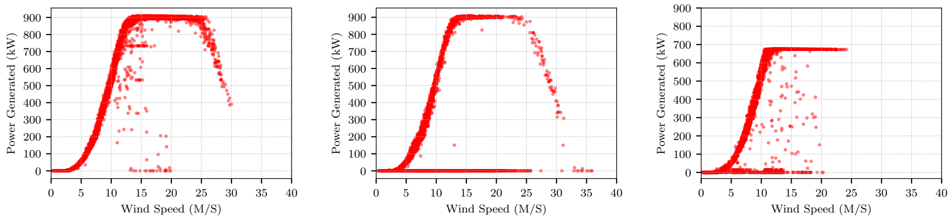
(b) No curtailment in Zone K

Figure 15: The wind-power data in Figure 14, split by curtailment in Zone K

In the OWE curtailment report, they describe December 2018 as a "model month" with low loss, and that January was affected by a technical fault that halted the turbine for for an extended period of time, meaning that it did not output any power in this period.

This information encourages us to split the data by date instead, as we might see patterns this way. We split the data into the three months and plot them again, shown in Figure 16, which illustrates the descriptions given in the curtailment report.

December (Figure 16a) is almost a perfect power curve, with few scattered outliers, and January (Figure 16b) has a very dominant line at 0kW caused by the technical fault mentioned above. Interestingly, it seems that all power output in February (Figure 16c) is capped around 675kW, regard-



(a) December (b) January (c) February

Figure 16: The wind-power data in Figure 14, split by month

less of the ANM status reported by SSE.

It is not clear what the cause of this cap is, but unless the truth values in our dataset are false, this cap is not caused by the ANM curtailment system, indicating that it must be caused by some other factors, possibly internal to the turbine.

10.2 Categorizing Loss

OWE reports a combined output for the three months of 716,131 kWh, with a loss of 31.52%. In order to calculate this loss, they estimate the expected power output by rounding the mean wind speed up to the nearest integer and multiplying it with the corresponding value in their power curve, which is defined for integer wind speeds up to 25 m/s. This is a modified version of the power curve supplied by the turbine manufacturer, adjusted to fit the experienced power output of the turbine. If the maximum wind speed in each 10-minute window is above 30 m/s, the actual output is used as the expected power output.

If we calculate the loss only on the measurements where the SSE reported curtailment in Zone K, we can estimate an upper bound on the amount of power lost due to ANM curtailment, listed in Table 7. When we use our data to calculate the actual and expected power outputs, we get marginally different values from those supplied by OWE, with a combined output of 708,951 kWh, and a loss of 31.44%. Filtering out measurements with no reported curtailment, we get a loss of 2.86%, which is only 9.09% of the

Method	**Expected**	**Total Loss**	**Curtailed Loss**
OWE Method	1,034,088 kWh	31.44%	2.86%
Interpolation	969,772 kWh	26.90%	2.71%

Table 7: Estimations of power loss in the OWE turbine. The curtailed loss only includes times where SSE reported curtailment in Zone K

total lost energy.

Further, if we use linear interpolation of the power curve, instead of rounding the wind speed up, when estimating the expected power output, both the estimate and the losses decrease. Using this method the total loss is 26.90%, and the upper bound of the curtailed loss is 2.76%.

10.3 Expanding the Power Curve

Both the power curve provided by turbine manufacturer, and the one used by OWE, are limited in that they only go up to 25 m/s, but from our data we can see a pattern of decreasing output when the wind speeds are higher. So a possible enhancement for the curtailment estimation is to extend the power curve above 25 m/s, to also reflect the apparent decrease in power output at these wind speeds.

In our dataset only 1.61% of the measurements have a mean wind speed above 25 m/s, but the difference in expected and actual output can still be significant as the output at these speeds drops quite quickly.

To compare the different power curves we have in play, we have plotted them all in Figure 17. We have calculated two new power curves from the data: One for all three months and one only for the "model month" of December. These are again calculated by rounding the wind speeds to the nearest integer, grouping the power outputs by wind speed, and taking the median of each group.

As we can see, the curve calculated from all three months is quite a lot lower than the others, due to the power losses we have described. The curve

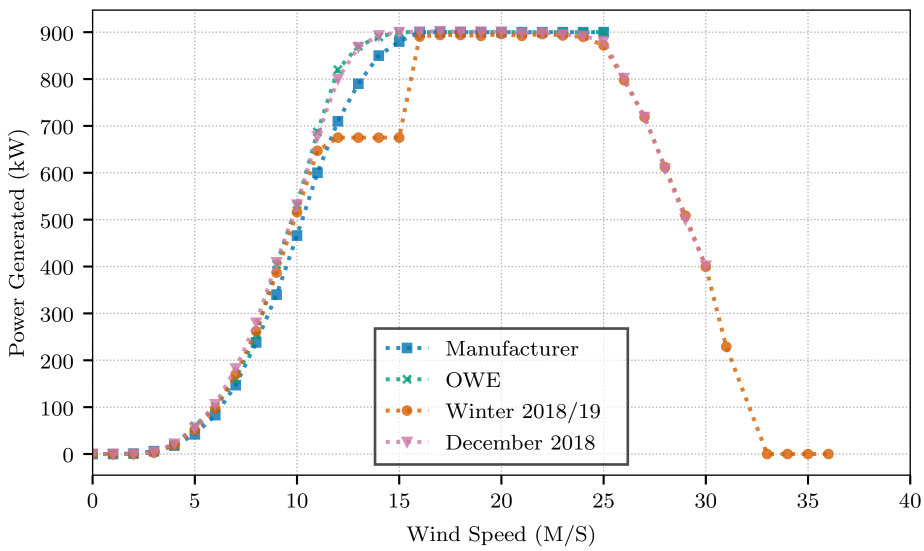


Figure 17: The different power curves for the OWE turbine.

provided by the turbine manufacturer is slightly lower than the OWE curve, while the curve calculated only from the data of December follows it quite tightly up until 25 m/s, where the OWE curve stops and the December curve drops.

From this it seems that the curve calculated from the December data could be a suitable replacement for the OWE power curve, adding some information about how the turbine behaves in high wind speeds.

10.4 Modeling with OWE Data

From the plots presented in this section, it seems that the wind measurements collected by the OWE turbine are more stable and of higher quality than those we have gathered from OpenWeatherMap. And as one of the main issues with the WT-models is the poor quality of the weather data, it becomes interesting to look at modeling with the OWE data instead.

In Figure 18, we plot the mean wind speeds measured at the OWE turbine against the total power generation in Orkney. From this we can see a clear relationship, indicating that the measurements from the turbine might be more useful to our ANM models than the data collected from OpenWeatherMap.

When training the WT-FFNN with wind data from the OWE turbine, the mean accuracy increases by over 4 percentage points, as shown in Table 8. From this, it is clear that there is potential for accuracy increases by

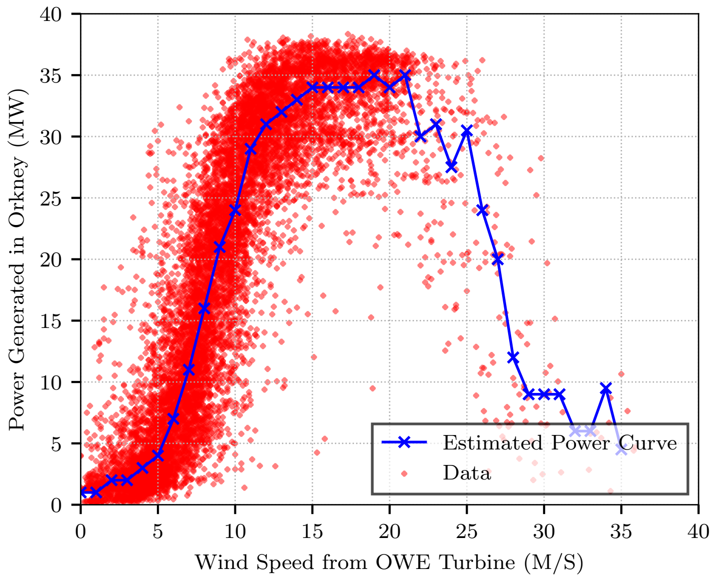


Figure 18: Scatter plot showing relationship between wind speeds from OWE turbine and total generation in Orkney from SSE

Wind Data Source	Mean	STD	Median	Min	Max
OpenWeatherMap	77.91	±1.71	78.27	73.66	80.46
OWE Reports	83.28	±2.36	83.37	76.91	87.51

Table 8: Accuracies of WT-FFNNs with different sources for wind data

using higher quality weather data, but it is uncertain whether simply collecting more data from the OpenWeatherMap API and retraining the model on a newer dataset could yield similar results.

10.5 Summary

In this section we have looked at the data generated by a single 900kW wind turbine in Orkney. We have analyzed the data to identify the amount of curtailment issued to the turbine. We found no clear connection between the power loss reported by OWE and the ANM curtailment status reported by SSE, but we did find a clear cause for power loss in a technical fault that halted the turbine for 18 days, and an apparent cap on the power output effective for the full month of February. By using the SSE data as truth values, we found that a maximum of 9.09% of the reported loss could be due to ANM-based curtailment.

Here it is worth noting again that the three months we are looking at, are the months least prone to curtailment, as the energy demand is higher in the winter. And while the calculations in this section can seem like an undermining of the problem of ANM-based curtailment, it should rather be seen as the development of tools and methods to help classify power loss. These classifications can potentially be used to identify other types of problems, or to identify problems in the ANM itself, insofar the information reported by SSE fails to reflect the reality of the people dependent on these turbines.

We have also extracted a power curve for the turbine, and proposed an interpolation-based method of estimating the expected power output, which might help OWE get a more precise idea of how much power they are losing.

Finally, we found a strong correlation between the wind data measured by the OWE turbine, and the power generation for all of Orkney, indicating that this data could be usable for modeling. We then used the wind speed data acquired from OWE to retrain our WT-FFNN model, resulting in significant accuracy gains.

11 A Predictive Model

One of the primary reasons for creating a curtailment model based on wind speeds was the fact that weather forecasts, which are the results of advanced predictive meteorological models, are publicly available, and modeling in terms of wind allows us to combine our models with these to obtain a predictive curtailment model.

According to our orcadian contacts, the most precise weather forecasts for the region are those generated by the UK Met Office, who also deliver forecasts to local newspaper The Orcadian. The UK Met Office offers a free API service, returning 5-day forecasts at fixed 3-hour intervals as a single JSON-object. These forecasts are recalculated once every hour, and we have been collecting these weather predictions since the 28th of March 2019. These forecasts, along with the time and weekday that they are forecasts for, can be used as input to our WT-based models to predict whether curtailment will occur.

11.1 Evaluating the Predictive Model

When constructing a dataset from the weather forecasts, we record both the time at which the forecast was computed, and the target time of each forecast. From this we can calculate the range of the forecast, which is the amount of hours-ahead a forecast is. In each JSON-object we receive, there is between 24 to 30 forecasts ranging from 96 to 120 hours into the future, depending on the time of day it is retrieved. Each of these forecasts is a single row in this new dataset.

Using this data, we construct a dataset for the month of April, the first month where we have full data both from SSE and from the Met Office, joining the forecasts with our ANM-status data on the forecast time. Each row then contains forecast time, predicted wind speed, hours-ahead, and curtailment status, and the data is cleaned using the 6-hour rule and the de minimis level. In total we evaluate on 19,584 unique forecasts ranging from 1 to 96 hours-ahead. We group our results by hours-ahead, with around 200 forecasts in each group, to get an idea of how well our predictive model works for different forecasting ranges.

The results of our evaluations are listed in Table 9, where the accuracies are further grouped into intervals of 12 hours. We evaluate WT_6 and our two ANN-based models, the perceptron and the FFNN, and see that our manually constructed model WT_6 has accuracies of ∼74% for same-day forecasts, over 10 percentage points lower than in the descriptive version. After this, the accuracy drops to about 71% at 36-48 hours, and then, curiously, rises again to meet its peak value of 74.57% at 84-96 hours. The ANN-based models are performing poorly throughout with total accuracies of 63-65%, the perceptron slightly better than the FFNN.

Both WT_6 and the neural networks have been designed and evaluated using the weather data from OpenWeatherMap, which we found had quality issues. To see the effect of using another data source, we include the OWE data in our dataset, adapt WT_6 to use the power curve shown in Figure 18, and re-train the ANNs with the OWE wind data. The results for these models are listed with suffix "(OWE)" in Table 9. The OWE-based models show significantly higher accuracies, especially the ANN-based ones, but follow the same pattern as WT_6, with accuracies that fall and then rise as the forecast range increases.

To gain a better visual understanding of this accuracy development, we plot the accuracies of the four best-performing models in Figure 19, applying a moving average (window size = 6 hours) to smooth the curves and increase readability. This figure illustrates the peculiar rise in accuracies when the forecast range increases. It does not seem plausible that the models used by the UK Met Office display this characteristic on a general scale, so instead we suspect it is a feature specific to this concrete dataset. This of course limits our confidence in our results, and as data collection continues, we plan a larger evaluation in order to declare our models accuracy with more certainty.

Beside the difference in quality of our wind data sources, another difference is the physical location. The OWE data is sourced from the top of a 40m tall turbine, while OpenWeatherMap and the Met Office presumable

Range [hours]	WT_6	Percep	FFNN	WT_6 (OWE)	Percep (OWE)	FFNN (OWE)
(0, 12]	74.27%	65.15%	62.11%	77.01%	79.30%	78.81%
(12, 24]	74.29%	65.72%	64.05%	76.54%	79.46%	78.25%
(24, 36]	72.40%	66.68%	65.61%	74.04%	77.86%	75.47%
(36, 48]	70.98%	66.32%	65.14%	70.53%	76.08%	73.03%
(48, 60]	72.00%	65.77%	63.49%	72.75%	75.14%	74.34%
(60, 72]	73.23%	65.88%	62.95%	74.04%	76.54%	74.65%
(72, 84]	74.08%	64.41%	62.33%	77.53%	78.94%	77.95%
(84, 96]	74.57%	62.05%	60.20%	77.35%	80.67%	78.13%
All	73.24%	65.26%	63.25%	74.98%	**78.00%**	76.34%

Table 9: Accuracy of predictive models on different forecasting ranges in hours, divided into 12-hour intervals.

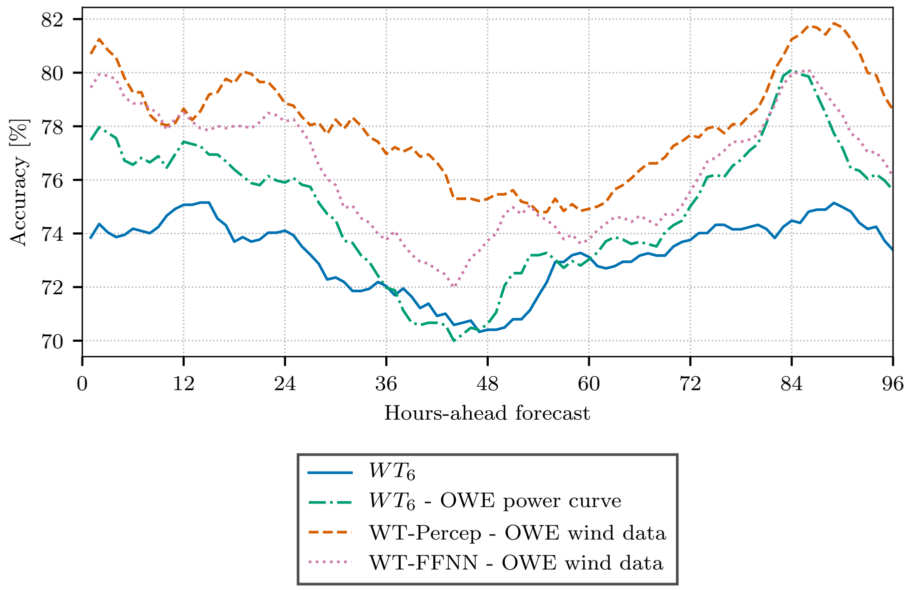


Figure 19: Accuracy of the four best-performing predictive models for different forecasting horizons. Each plot is smoothed with a moving average with a window size of 6 hours. See Figure 28 in Appendix A for full results.

are sourced nearer to ground level. The OWE data is located on the west coast of one of the isles, while the other two are located on the east coast of the (rather flat) island of Westray. There might also be differences in sensor quality and calibration. These aspects all create discrepancies between our data sources, which can add to the model inaccuracy.

A final aspect to consider is the seasonal bias of our dataset, which might affect our results as well. The models are designed and trained solely on winter data, while the data we are evaluating on is from April, the middle of spring. We suspect that models based on a dataset covering all seasons would have higher accuracies.

11.2 Probabilistic Prediction

As mentioned in section 9, an interesting feature of the ANN-based models, is that their output is not just a Boolean value, but a floating point value between 0 and 1, where anything above 0.5 is converted to True and anything equal to or below is converted to False. How close the prediction is to these extremes can be interpreted as the *certainty* of the model, and to improve the richness of our output, we can use this value to estimate the probability of curtailment for a given sample by looking at the relationship between output value and accuracy.

In order to do this, we split our predictions into 20 groups, each an interval of 0.05 of the prediction value. We can then calculate the statistical probability of curtailment for each of these groups. The interval size is a compromise between resolution and confidence, as increasing the number of groups improves the resolution, but reduces our confidence in the calculated accuracies, as there are fewer samples in each group. In Figure 20 we show the results of applying this to the two versions of the WT perceptron. We also show the amount of samples in each group.

When looking at the results for the perceptron trained on OWE data Figure 20a, the best performing of our models, we see that the probability of curtailment starts rising already around an output of 0.3 and then peaks in the interval $(0.6 - 0.65]$. After 0.8 it declines rapidly, although this could be associated with the low number of samples in this group. The vast majority of predictions are in the interval $(0.15 - 0.75]$, meaning that the model rarely outputs values in the extremes of its range.

When applying the same method to the perceptron trained on data from OpenWeatherMap, shown in in Figure 20b, we see a very different progression. This model is much more likely to output a value closer to 1, predicting curtailment for 74.37% of our samples, while the perceptron trained on OWE data only predicts curtailment for 47.22% of samples.

Applying this method to the two corresponding FFNNs,[18] yields similar progressions, indicating that this difference in model output is a result of the discrepancies in wind data sources. We expect that a model using the same data source for both training and predicting would have a better utilization of the full output range.

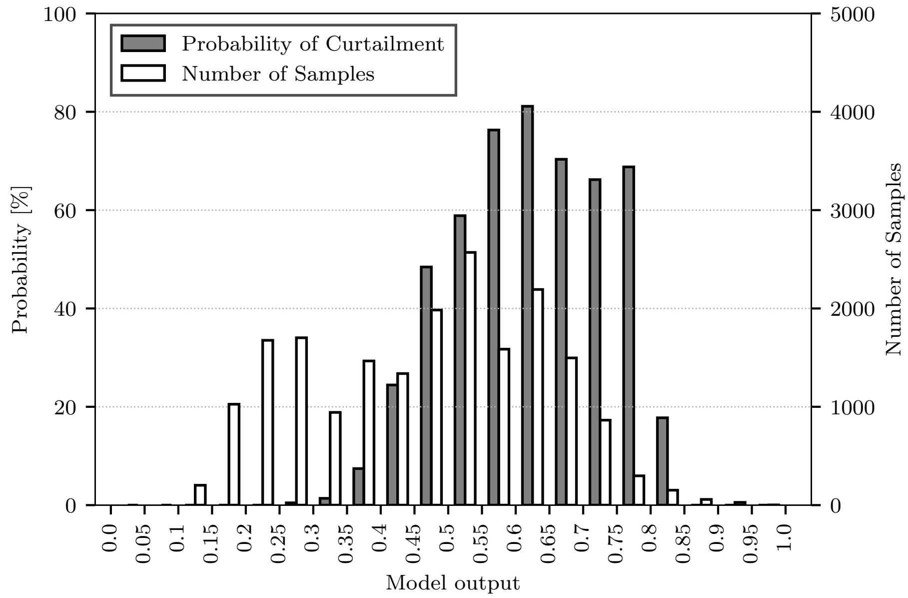
By using these grouped accuracies and linear interpolation, we can estimate the probability p associated with model output x as:

$$\begin{aligned} i &= (x - 0.025) * 20 \\ p &= A_{\lfloor i \rfloor}(\lceil i \rceil - i) + A_{\lceil i \rceil}(i - \lfloor i \rfloor) \end{aligned} \tag{4}$$

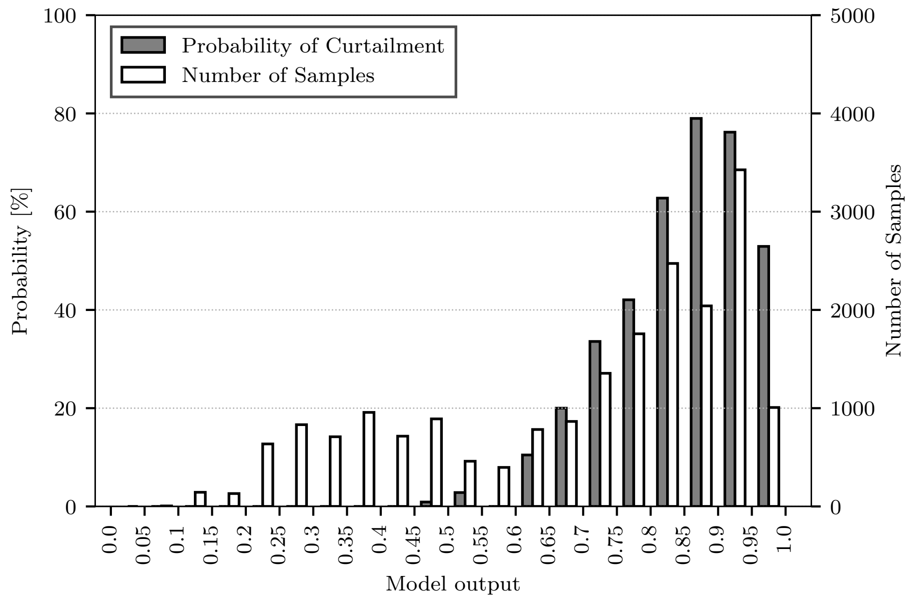
Where A_k denotes the accuracy of group k. The offset value 0.025 is subtracted from the output to compensate for the fact that the groups are indexed by the start of the interval.

It should be noted that these are fairly crude statistical probabilities that do not necessarily translate directly to the real world, and should taken as such. For example, by using this method, our models will output zero probability for curtailment for many of the lower outputs, and from what we have learned from our investigation of the ANM, this is far too bold a statement.

[18]Shown in Figure 29 in Appendix A



(a) WT perceptron trained on OWE wind data.



(b) WT perceptron trained on OpenWeatherMap wind data.

Figure 20: Curtailment probability and number of samples, grouped by model output in intervals of 0.05

11.3 Summary

In this section we have described a predictive curtailment model. This was obtained by combining weather forecasts, themselves results of meteorological models, with the WT-based models developed in this project. We evaluated the models on around 20,000 unique forecasts for the month of April 2019 collected through the UK Met Office, ranging from 1 to 96 hours in forecast horizon. From our evaluation it is clear that WT_6 performs better than the ANN-based models, but when exchanging our wind data for the data acquired from OWE, we see accuracies rise significantly.

These predictive models have lower accuracies than their descriptive counterparts, caused in part by the addition of the inaccuracies of the weather forecasts, the discrepancies between data sources, and the seasonal bias present in our dataset. Interestingly, the accuracies measured do not decline steadily as the forecast range increases. Instead they fall and then rise again, which we interpret as a special characteristic for this dataset, and not a general feature of the combined model.

The accuracies of our predictive models leave room for improvement, but we believe that the current models are sufficiently accurate to be useful, and, furthermore, we believe that improvements will be obtained as our dataset grows larger and we can train and predict on the same data source.

The best performing model in our evaluation is the WT perceptron trained on the OWE wind data. When applying this model in practice, it is an additional advantage that a degree of certainty is built into the output, making it possible to estimate the probability of curtailment with more nuance.[19]

12 Conclusion

In this project we have been working to gain an understanding of how the Orkney ANM curtailment system works. Both the nature and severity of the problem, and to determine how it can best be modeled. The purpose of this investigation has been, in part, creating knowledge and insights for the orcadians, for whom curtailment is a pressing issue, but also developing models that can be used in the proposed Orkney Cloud data center pods for charge-controlling. Concretely, we have been addressing the research question: *How can we identify curtailment of turbines in the Orkney power grid by inferring the context from publicly available data sources?*

[19] As an example, a prototype curtailment forecast service is available at `http://forecast.curtailment.net/` as of May 31st 2019. This JavaScript-based implementation runs on the client-side, asynchronously fetching the newest weather forecast from the UK Met Office API and piping it through a model converted and run using TensorFlow.js, meaning that the server only has to serve static files. Source code available at `https://github.com/NielsOerbaek/CurtailmentForecast`

To answer this question, we have created a dataset through continuous web scraping, collecting relevant available data at 10-minute intervals spanning three months. This dataset allowed for us to evaluate our assumptions about the ANM system, and, through a series of visualizations, identify and subsequently extract correlations and patterns in the data. These correlations enabled us to model curtailment both in terms of factors internal to the grid (generation and demand), but also through external factors (wind speeds, time-of-day and day-of-week).

When evaluating and analyzing our models, we have found that data quality is paramount to the development of usable models, especially due to the limited size of our dataset. We identified anomalous parts of our data that represent situations or aspects that we did not want our models to encompass, leading us to define rules for systematic data cleaning and anomaly detection. We have found that a small majority (51.85%) of the times where some curtailment was reported, it did not fit our understanding of how the ANM should work. This does not, however, imply that the majority of curtailed power loss (in Wh) is caused by these anomalies, as we have limited ways of quantifying power loss.

We have trained simple ANN-based models, using the features identified in our data exploration, and found slight accuracy increases when using this technique. These models have the additional advantages that they are trivial to adapt to new data, and that they output a degree of certainty. They do not, on the other hand, provide tangible and human-readable knowledge about how the ANM works, making them suitable only in combination with the manually constructed rule-based models.

When redesigning and retraining our models with the wind data obtained from a single turbine in Orkney, we saw significant accuracy increases, further underlining that data quality is extremely important. And while this wind data is not publicly available, we believe that it is possible to get publicly available weather data of the same quality.

Finally, we have combined our developed wind-and-time-based models with publicly available weather forecasts to assemble a predictive model, forecasting whether curtailment will occur in the future. We presented a method of estimating probability for curtailment, which could be suitable both for a public curtailment forecast service, and to implement power-controlling mechanisms in a grid-aware decentralized data center.

13 Future Work

As described in section 1, the initial aim of this project was to investigate the vision of an Orkney Cloud data center, and while we have uncovered much useful knowledge in this project, there are still a lot of interesting avenues for further exploration that could be considered. In this section we will highlight some of these.

13.1 Other Perspectives on the Dataset

While we have spent a great amount of energy in this report exploring our dataset, there are still perspectives that we have not delved into.

For example, the data scraped from the SSE website splits the generation into firm connections and ANM-connections. When comparing these numbers over time, as for example in Figure 21, it seems that the two numbers behave differently, and there might be some interesting information to extract from this upon further analysis. Possibly it could even give us clues as to how we can quantify the amount of power loss.

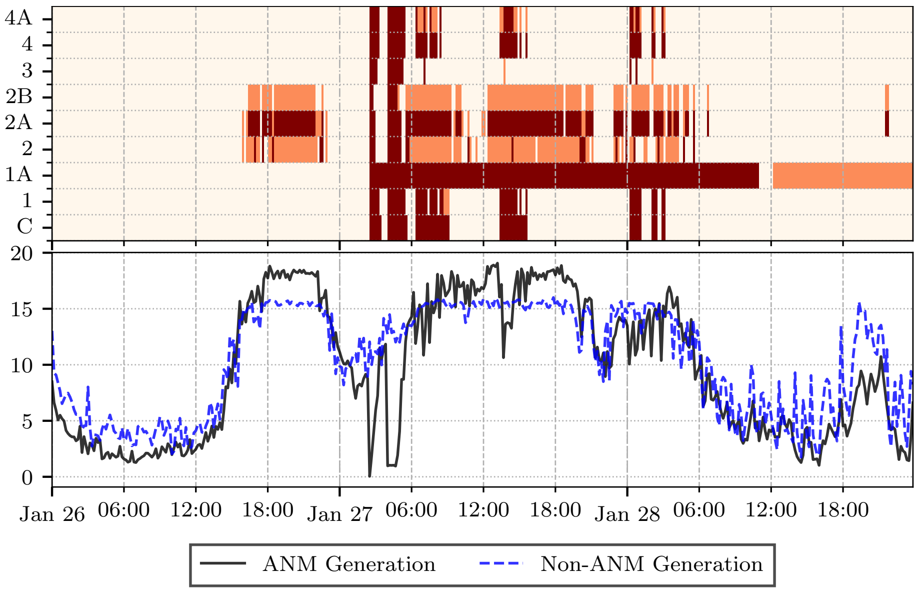


Figure 21: Timeline showing ANM and Non-ANM generation over time and curtailment in individual zones. Some labels and legends are ommitted for brevity. See Figure 4 for full labels.

Another option is to look into developing more nuanced models describing curtailment in each zone. Our current models only look at the ANM as a whole, but there might be interesting insights from looking at the zones individually, possibly developing a way to automatically create and compare the models for each individual zone.

Finally we could expand the dataset with features such as sunlight, tidal current speed and wave height, and look into the effect of these renewable energy sources on the total Orkney generation.

All of these perspectives become only more interesting as our data collection continues and we are able to create larger datasets.

13.2 Other ANN Architectures

In this project we have only used very basic neural network architectures, but it could be interesting to look into other architectures as well. Our dataset is actually a time series, with each sample in sequence describing a progression in the data, but currently we are only feeding single samples into our ANN models. One of the larger fields within deep learning research is that of *sequence prediction*, usually implemented using *recurrent neural networks* (RNNs), specifically designed to tackle this kind of problem. Several types of RNNs are also built into Keras, the deep learning library we are using, and experimenting with models of this type is an obvious next step.

References

[1] Cisco. *Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper*. en. 2017. URL: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html (visited on 04/24/2019) (cit. on pp. 5, 6).

[2] M. Dayarathna, Y. Wen, and R. Fan. "Data Center Energy Consumption Modeling: A Survey". In: *IEEE Communications Surveys Tutorials* 18.1 (2016), pp. 732–794. ISSN: 1553-877X. DOI: 10.1109/COMST.2015.2481183 (cit. on p. 5).

[3] Tom Bawden. "Global warming: Data centres to consume three times as much energy in next decade, experts warn". In: *The Independent* (Jan. 2016) (cit. on p. 5).

[4] Niels Møller Kjemtrup. "Skyen ligger på en dansk mark – og den æder strøm". da. In: *Ingeniøren* December, 2018 (Dec. 2018), pp. 44–49. URL: https://ing.dk/artikel/skyen-ligger-paa-dansk-mark-aeder-stroem-222587 (visited on 04/04/2019) (cit. on p. 5).

[5] Arman Shehabi, Ben Walker, and Eric Masanet. "The energy and greenhouse-gas implications of internet video streaming in the United States". en. In: *Environmental Research Letters* 9.5 (May 2014), p. 054007. ISSN: 1748-9326. DOI: 10.1088/1748-9326/9/5/054007. URL: https://doi.org/10.1088%2F1748-9326%2F9%2F5%2F054007 (visited on 04/04/2019) (cit. on p. 6).

[6] Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli. "Fog Computing and Its Role in the Internet of Things". In: *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*. MCC '12. event-place: Helsinki, Finland. New York, NY, USA: ACM, 2012, pp. 13–16. ISBN: 978-1-4503-1519-7. DOI: 10.1145/2342509.2342513. URL: http://doi.acm.org/10.1145/2342509.2342513 (visited on 05/02/2019) (cit. on p. 6).

[7] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu. "Edge Computing: Vision and Challenges". In: *IEEE Internet of Things Journal* 3.5 (Oct. 2016), pp. 637–646. ISSN: 2327-4662. DOI: 10.1109/JIOT.2016.2579198 (cit. on p. 6).

[8] The Orkney Cloud Project. *Orkney Cloud Website*. 2018. URL: http://orkneycloud.org/ (cit. on p. 6).

[9] Laura Watts, ed. *Orkney Cloud - Community-led data services*. The Orkney Cloud Project, 2018. URL: http://orkneycloud.org/publish/orkneycloud2018.pdf (visited on 03/20/2019) (cit. on p. 7).

[10] Philippe Bonnet. "Orkney Cloud Data Center". In: *Orkney Cloud - Community-led data services* (2018), pp. 36–38. URL: http://orkneycloud.org/publish/orkneycloud2018.pdf#page=38 (visited on 03/20/2019) (cit. on p. 7).

[11] Niels Ørbæk Christensen. *Preparing to Prototype the Orkney Cloud*. Thesis Preparation. IT University of Copenhagen, Dec. 2018 (cit. on p. 7).

[12] Laura Watts. "Energy at the End of the World: An Orkney Islands Saga". English. In: (2018). URL: https://www.research.ed.ac.uk/portal/en/publications/energy-at-the-end-of-the-world-an-orkney-islands-saga(a3dca776-b9ae-43c5-8288-2184095374cf).html (visited on 02/05/2019) (cit. on pp. 7, 12).

[13] Xiao-chao Fan, Wei-qing Wang, Rui-jing Shi, and Feng-ting Li. "Analysis and countermeasures of wind power curtailment in China". In: *Renewable and Sustainable Energy Reviews* 52 (Dec. 2015), pp. 1429–1436. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.08.025. URL: http://www.sciencedirect.com/science/article/pii/S1364032115008710 (visited on 01/31/2019) (cit. on p. 7).

[14] Debra Lew, Lori Bird, Michael Milligan, Bethany Speer, Enrico Maria Carlini, Ana Estanqueiro, Damian Flynn, Emilio Gómez-Lázaro, Hannele Holttinen, Nickie Menemenlis, Antje Orths, J. Charles Smith, Lennart Söder, Poul Ejnar Sørensen, and Yoh Yasuda. "Wind and Solar Curtailment: International Experience and Practices". In: *International Workshop on Large-Scale Integration of Wind Power Into Power Systems as Well as on Transmission Networks for Offshore Wind Power Plants* (2013) (cit. on p. 7).

[15] Gareth Davies. "Developing sustainable futures — an Orkney Way". In: *Orkney Cloud - Community-led data services* (2018), pp. 10–17 (cit. on p. 8).

[16] John Wilder Tukey. *Exploratory Data Analysis*. en. Addison-Wesley Publishing Company, Jan. 1977. ISBN: 978-0-201-07616-5 (cit. on p. 9).

[17] Natalia Andrienko and Gennady Andrienko. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. en. Springer Science & Business Media, Mar. 2006. ISBN: 978-3-540-31190-4 (cit. on p. 9).

[18] Orkney Renewable Energy Forum and Community Energy Scotland. *Orkney-Wide Energy Audit 2014: Energy Sources and Uses*. 2014. URL: http://www.oref.co.uk/wp-content/uploads/2015/05/Orkney-wide-energy-audit-2014-Energy-Sources-and-Uses.pdf (visited on 03/07/2019) (cit. on p. 11).

[19] Department for Business, Energy & Industrial Strategy. *Sub-national electricity sales and numbers of customers 2005-2017*. 2018. URL: https://www.gov.uk/government/statistical-data-sets/regional-and-local-authority-electricity-consumption-statistics (visited on 03/07/2019) (cit. on p. 11).

[20] Luis Hernandez, Carlos Baladron, Javier M. Aguiar, Belen Carro, Antonio J. Sanchez-Esguevillas, Jaime Lloret, and Joaquim Massana. "A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings". en. In: *IEEE Communications Surveys & Tutorials* 16.3 (2014), pp. 1460–1495. ISSN: 1553-877X. DOI: 10.1109/SURV.2014.032014.00094. URL: http://ieeexplore.ieee.org/document/6803101/ (visited on 02/06/2019) (cit. on pp. 11, 37).

[21] Orkney Renewable Energy Forum. *Innovations*. en-GB. 2019. URL: http://www.oref.co.uk/orkneys-energy/innovations-2/ (visited on 02/11/2019) (cit. on p. 12).

[22] Pierluigi Siano. "Demand response and smart grids—A survey". In: *Renewable and Sustainable Energy Reviews* 30 (Feb. 2014), pp. 461–478. ISSN: 1364-0321. DOI: 10.1016/j.rser.2013.10.022. URL: http://www.sciencedirect.com/science/article/pii/S1364032113007211 (visited on 05/02/2019) (cit. on p. 12).

[23] S. Aman, M. Frincu, C. Chelmis, M. Noor, Y. Simmhan, and V. K. Prasanna. "Prediction models for dynamic demand response: Requirements, challenges, and insights". In: *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. Nov. 2015, pp. 338–343. DOI: 10.1109/SmartGridComm.2015.7436323 (cit. on pp. 12, 37).

[24] Radius. *Radius indfører nye timetariffer*. da. 2017. URL: https://radiuselnet.dk/Nyheder/Nyhedsoversigt/Nyhedsarkiv/2018/01/News-Article---Radius-indfoerte-nye-timetariffer (visited on 05/02/2019) (cit. on p. 13).

[25] Ørsted. *Hvad er timepris?* da. 2017. URL: https://orsted.dk/Privat/Kundeservice/Aftaler-og-priser/Hvad-er-Timepris (visited on 05/02/2019) (cit. on p. 13).

[26] Donella H. Meadows. *Thinking in Systems: A Primer*. English. Ed. by Diana Wright. White River Junction, Vt: Chelsea Green Publishing, Dec. 2008. ISBN: 978-1-60358-055-7 (cit. on p. 14).

[27] Scottish and Southern Electricity Networks. *ANM About*. URL: https://www.ssen.co.uk/ANMAbout/ (visited on 05/01/2019) (cit. on p. 17).

[28] Python Software Foundation. *Python Language Reference*. 2015. URL: https://www.python.org/ (visited on 05/10/2019) (cit. on p. 19).

[29] Kenneth Reitz. *Requests: HTTP for Humans*. 2017. URL: https://2.python-requests.org/en/master/ (visited on 05/10/2019) (cit. on p. 19).

[30] Leonard Richardson. *Beautiful Soup*. 2018. URL: https://www.crummy.com/software/BeautifulSoup/ (visited on 05/10/2019) (cit. on p. 19).

[31] Kristina Chodorow and Michael Dirolf. *MongoDB: The Definitive Guide*. 1st. O'Reilly Media, Inc., 2010. ISBN: 978-1-4493-8156-1 (cit. on p. 19).

[32] Wes Mckinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference* (2010) (cit. on p. 19).

[33] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong. "Probabilistic Forecasting of Wind Power Generation Using Extreme Learning Machine". In: *IEEE Transactions on Power Systems* 29.3 (May 2014), pp. 1033–1044. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2013.2287871 (cit. on pp. 20, 37).

[34] Gregor Giebel, Richard Brownsword, George Kariniotakis, Michael Denhard, and Caroline Draxl. *The State-Of-The-Art in Short-Term Prediction of Wind Power*. Report. ANEMOS.plus, 2011. DOI: 10.11581/DTU:00000017 (cit. on pp. 20, 37).

[35] J. D. Hunter. "Matplotlib: A 2D graphics environment". In: *Computing In Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55 (cit. on p. 23).

[36] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. "Scikit-learn: Machine Learning in Python". In: *J. Mach. Learn. Res.* 12 (Nov. 2011), pp. 2825–2830. ISSN: 1532-4435. URL: http://dl.acm.org/citation.cfm?id=1953048.2078195 (visited on 05/10/2019) (cit. on p. 23).

[37] George E. P. Box. "Science and Statistics". In: *Journal of the American Statistical Association* 71.356 (Dec. 1976), pp. 791–799. ISSN: 0162-1459. DOI: 10.1080/01621459.1976.10480949. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1976.10480949 (visited on 05/20/2019) (cit. on p. 24).

[38] Patrick Milan, Matthias Wächter, Stephan Barth, and Joachim Peinke. "Power curves for wind turbines". en. In: *WIT Transactions on State of the Art in Science and Engineering*. 1st ed. Vol. 1. WIT Press, June 2010, pp. 595–612. ISBN: 978-1-84564-205-1. DOI: 10.2495/978-1-84564-205-1/18. URL: http://library.witpress.com/viewpaper.

asp?pcode=9781845642051-018-1 (visited on 03/07/2019) (cit. on p. 24).

[39] Danish Wind Industry Association. *Power Curve of a Wind Turbine.* 2003. URL: http://xn--drmstrre-64ad.dk/wp-content/wind/miller/windpower%20web/en/tour/wres/pwr.htm (visited on 03/07/2019) (cit. on p. 24).

[40] Jaesung Jung and Robert P. Broadwater. "Current status and future advances for wind speed and power forecasting". In: *Renewable and Sustainable Energy Reviews* 31 (Mar. 2014), pp. 762–777. ISSN: 1364-0321. DOI: 10.1016/j.rser.2013.12.054. URL: http://www.sciencedirect.com/science/article/pii/S1364032114000094 (visited on 02/04/2019) (cit. on p. 37).

[41] Zhi Li, Lin Ye, Yongning Zhao, Xuri Song, Jingzhu Teng, and Jingxin Jin. "Short-term wind power prediction based on extreme learning machine with error correction". en. In: *Protection and Control of Modern Power Systems* 1.1 (June 2016), p. 1. ISSN: 2367-0983. DOI: 10.1186/s41601-016-0016-y. URL: https://doi.org/10.1186/s41601-016-0016-y (visited on 02/01/2019) (cit. on p. 37).

[42] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning.* MIT Press, 2016 (cit. on p. 38).

[43] Yoav Goldberg. "A Primer on Neural Network Models for Natural Language Processing". In: *arXiv:1510.00726 [cs]* (Oct. 2015). arXiv: 1510.00726. URL: http://arxiv.org/abs/1510.00726 (visited on 04/29/2019) (cit. on p. 38).

[44] François Chollet. *keras.* GitHub, 2015. URL: https://github.com/fchollet/keras (cit. on p. 38).

[45] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* 2015. URL: http://tensorflow.org/ (cit. on p. 38).

[46] Francois Chollet. *Deep Learning with Python.* English. 1 edition. Shelter Island, New York: Manning Publications, Dec. 2017. ISBN: 978-1-61729-443-3 (cit. on p. 38).

[47] Hal Daumé III. *A Course in Machine Learning*. English. 2nd Edition. Self-published, 2017. URL: http://ciml.info (cit. on p. 38).

Appendix

A Additional Graphs

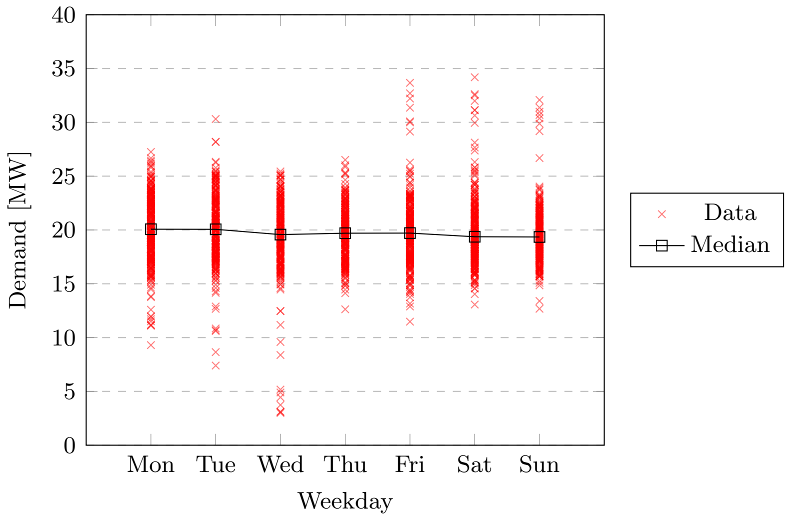


Figure 22: Scatter plot showing relationship between weekday and power demand

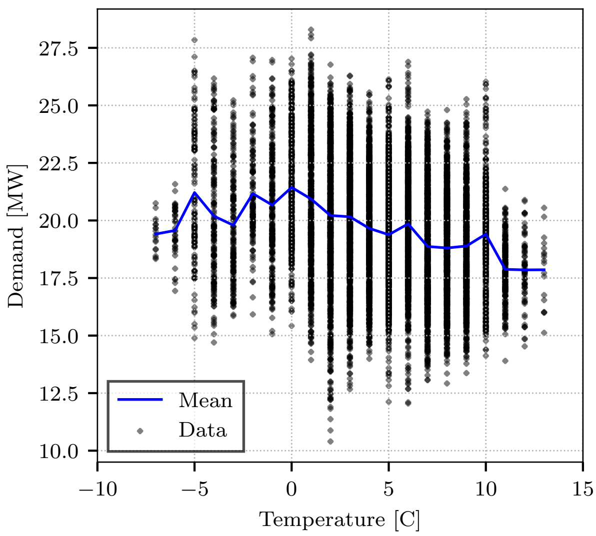


Figure 23: Relationship between energy demand in Orkney and outside temperature

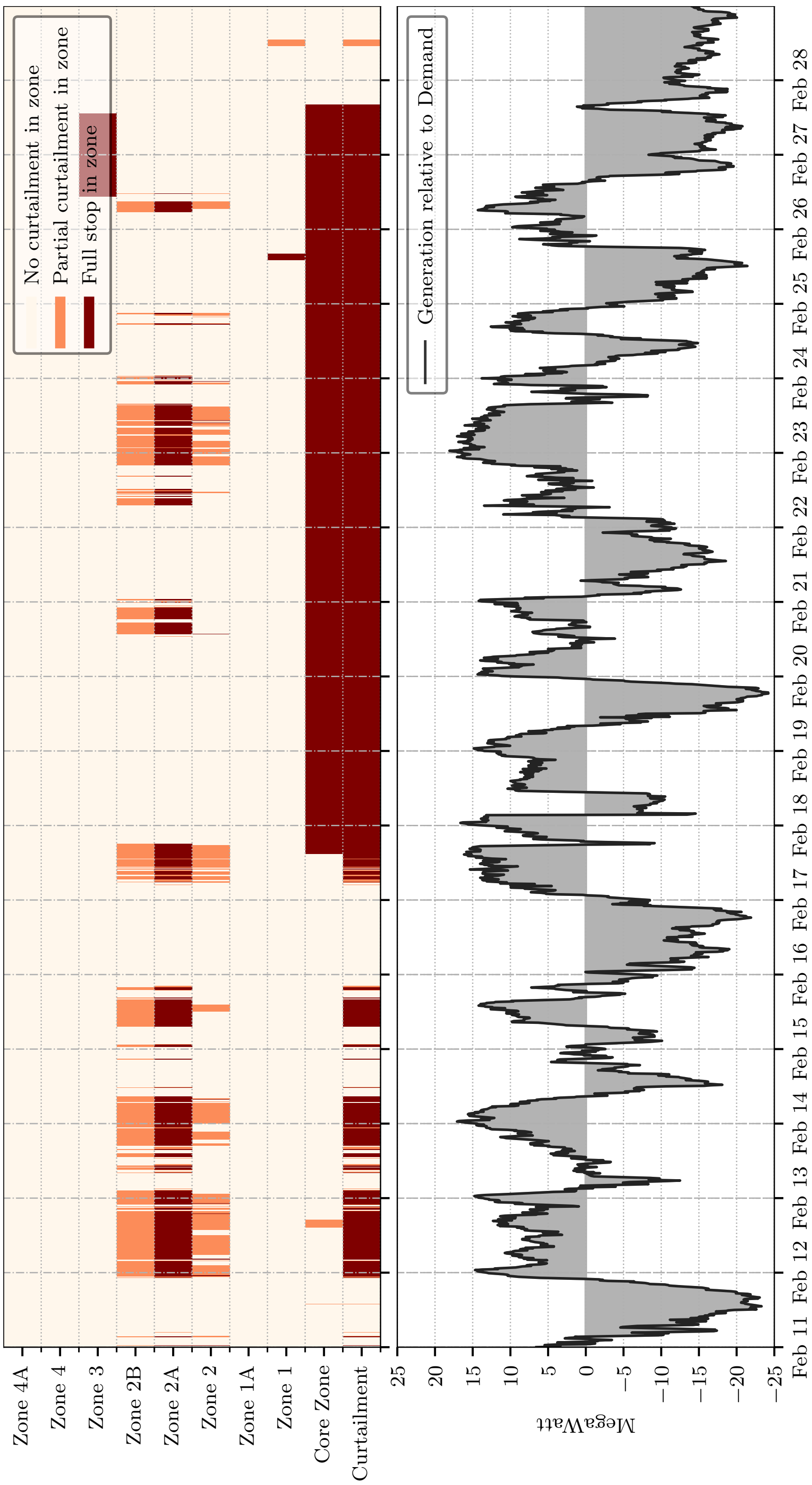


Figure 24: Prolonged curtailment situation in Core Zone

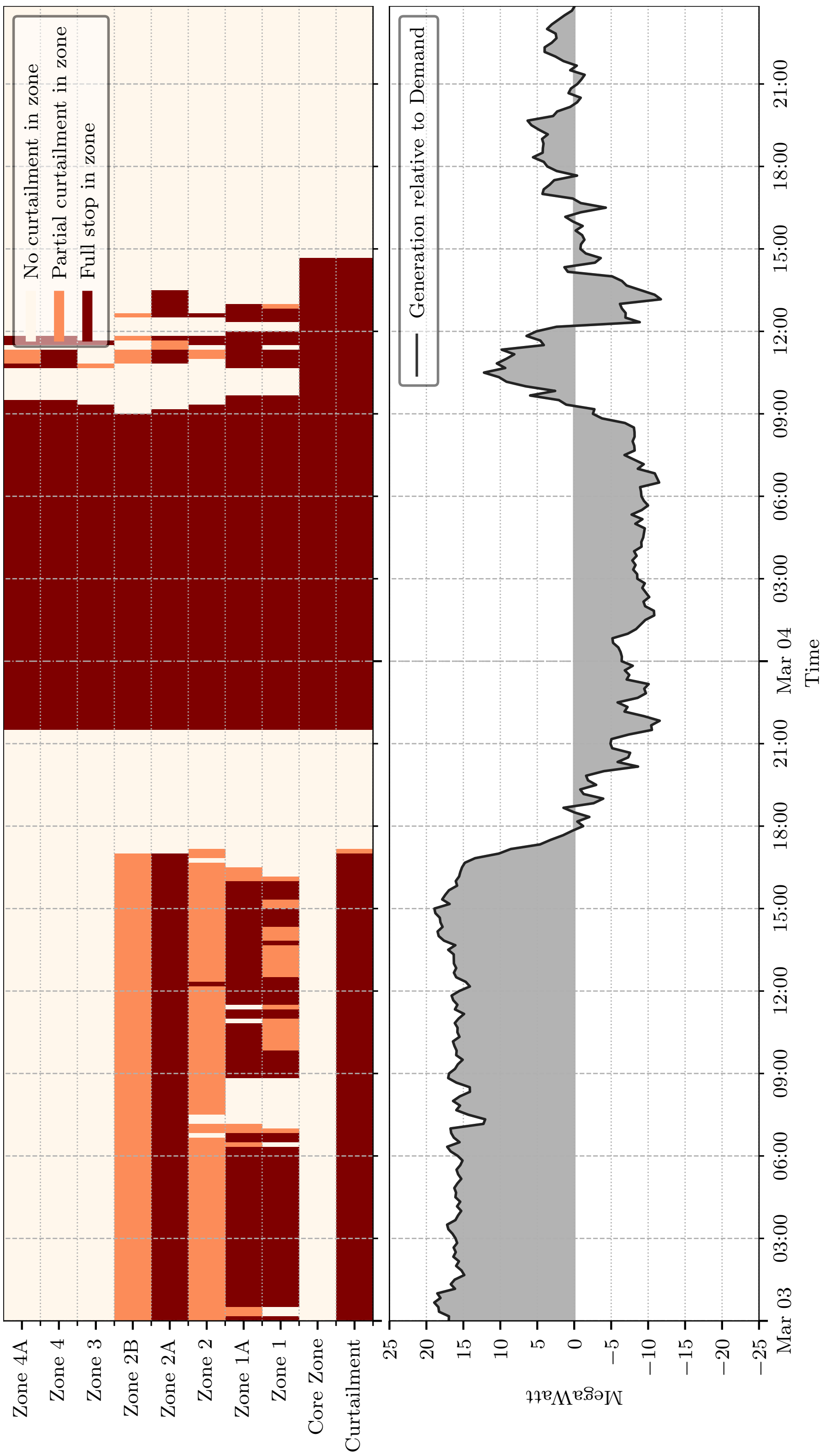


Figure 25: Curtailment situation with full stop in all zones

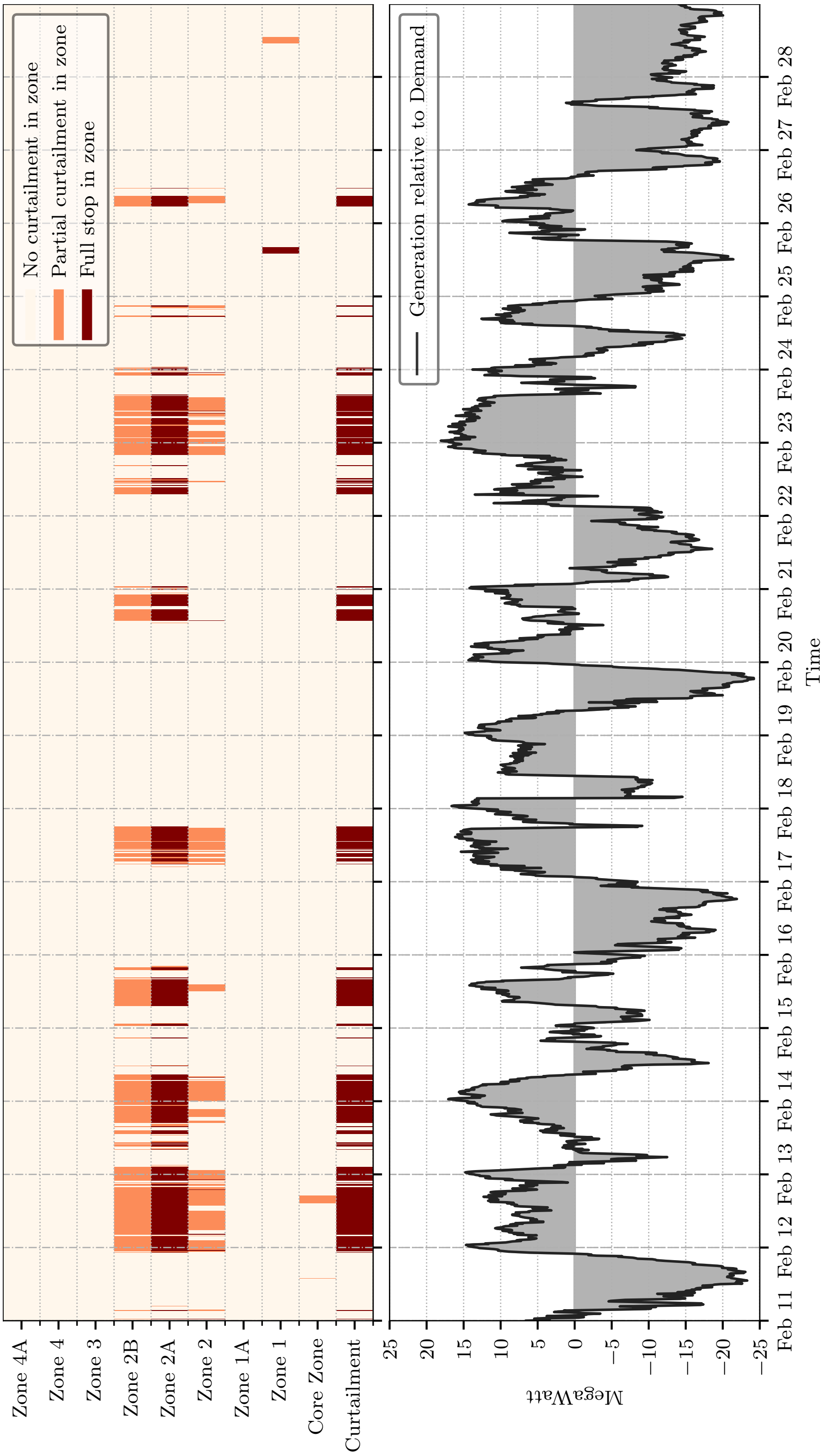


Figure 26: Figure 24 after data cleaning

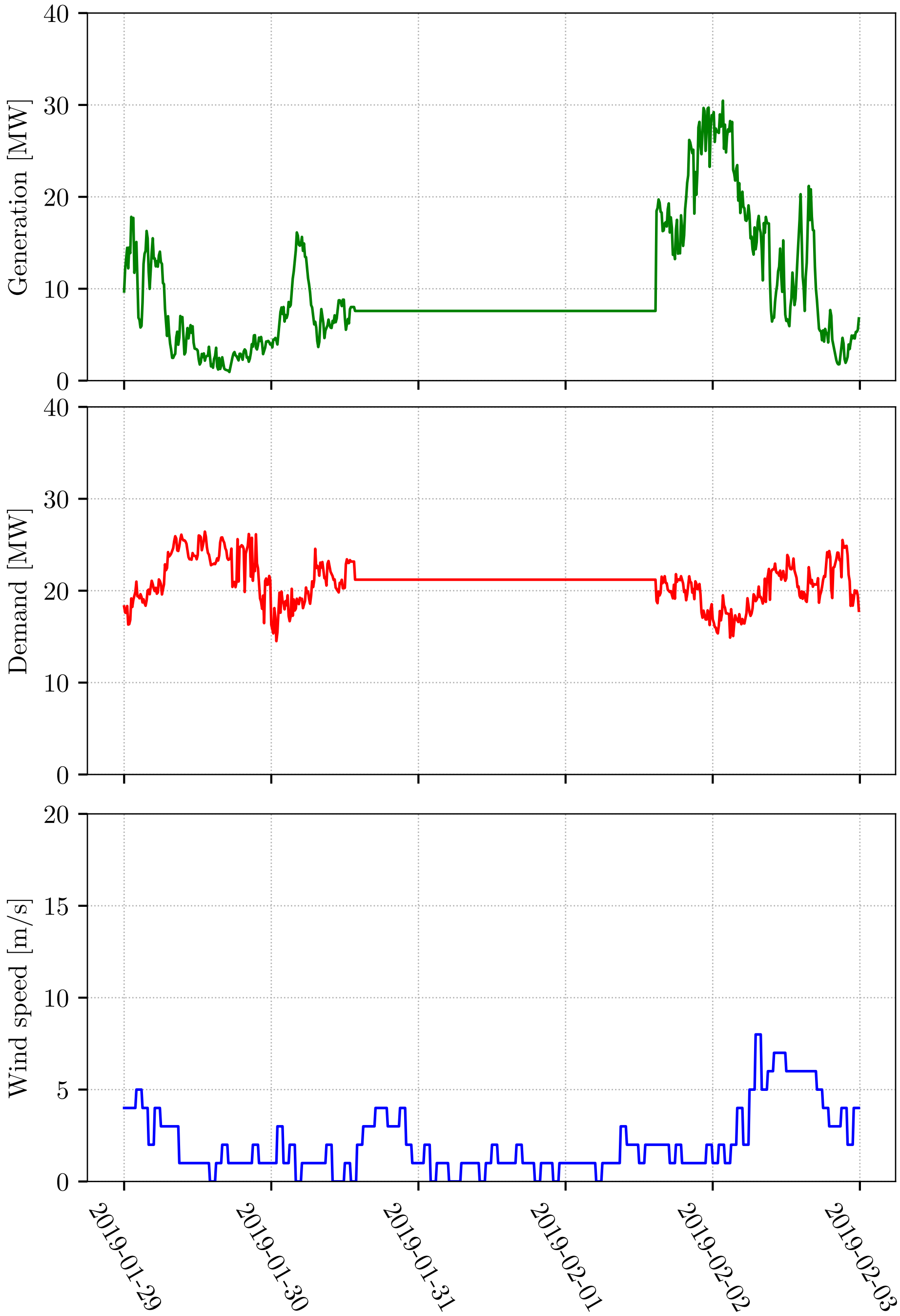


Figure 27: Hole in SSE-Data. January 30th to February 2nd

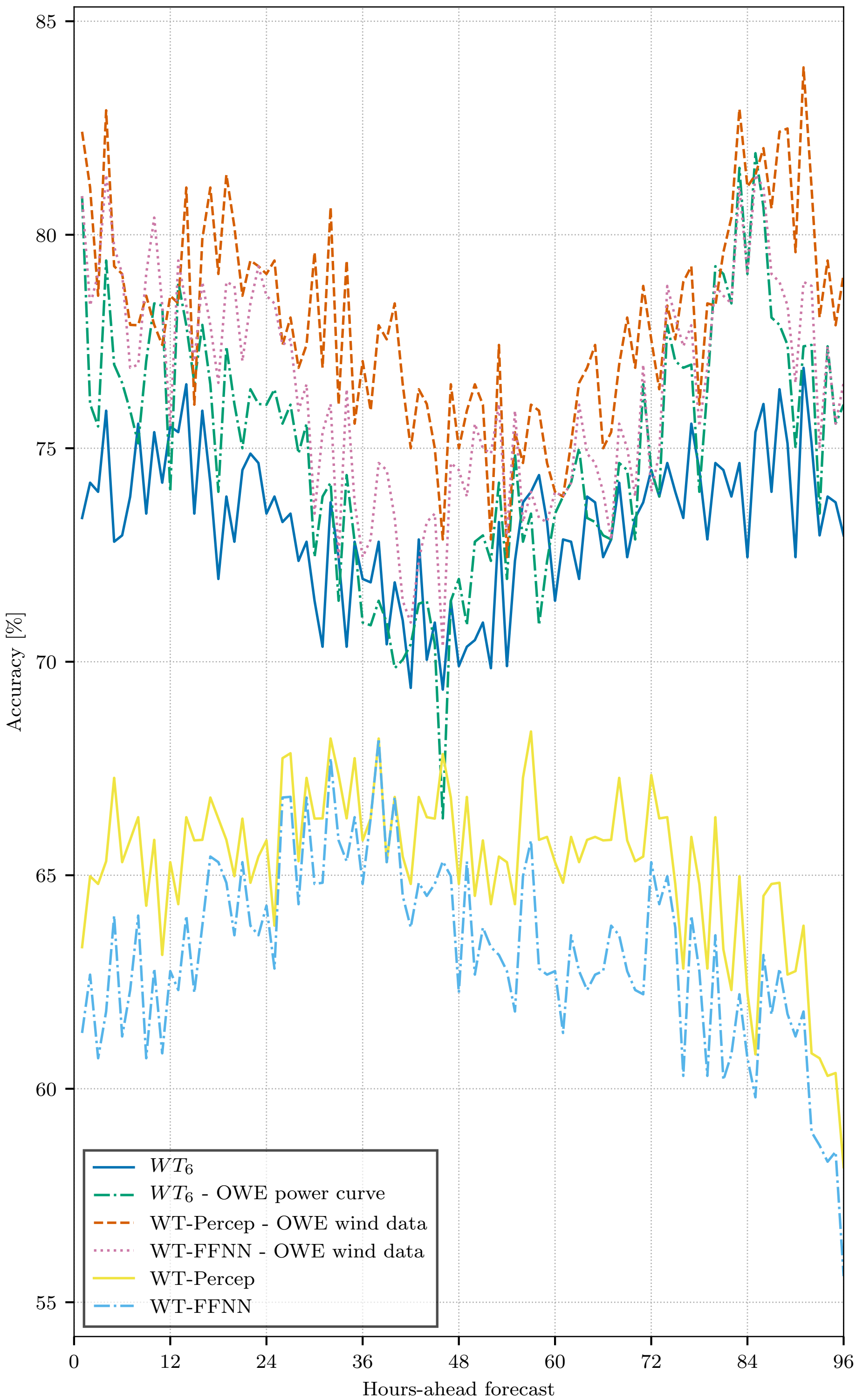
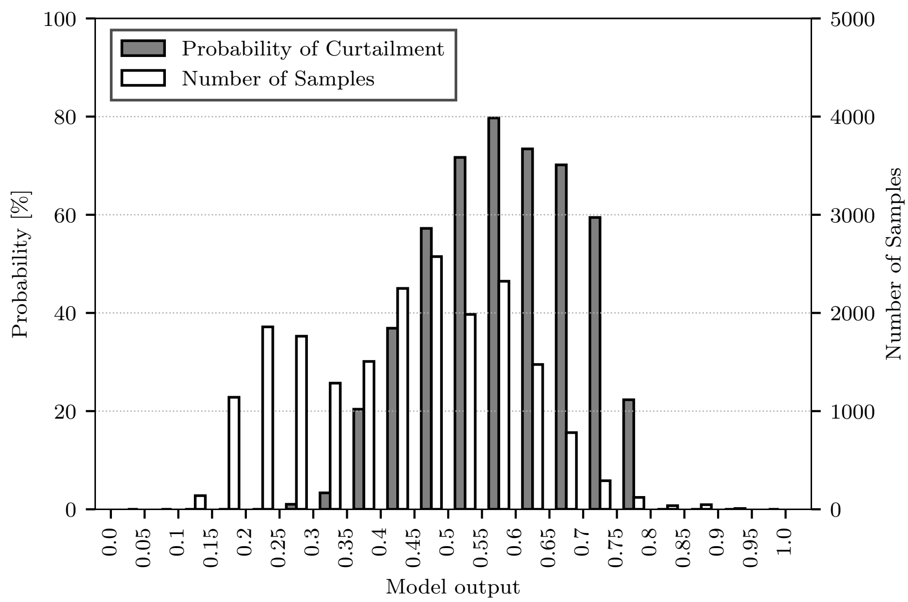
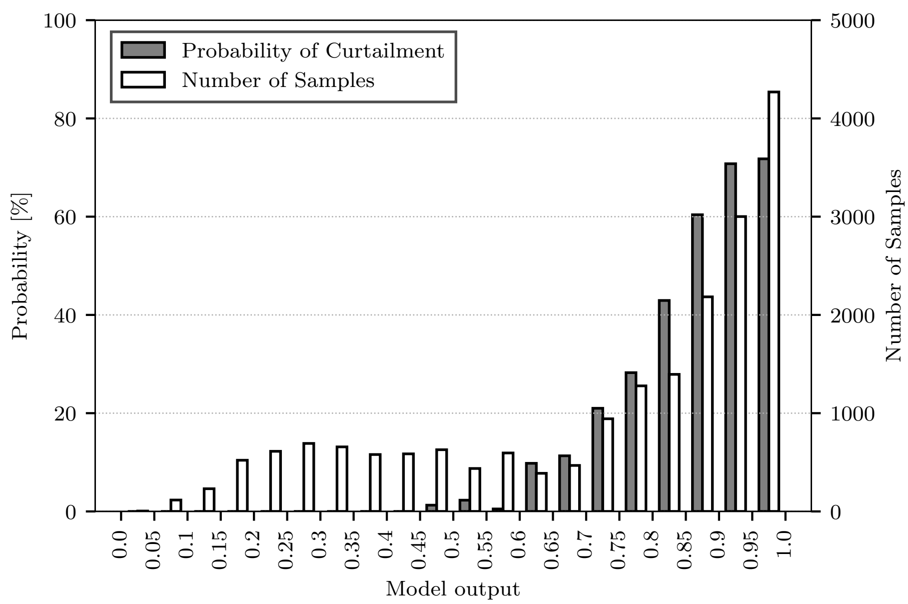


Figure 28: Accuracy of predictive models for different forecasting horizons



(a) WT FFNN trained on OWE wind data.



(b) WT FFNN trained on OpenWeatherMap wind data.

Figure 29: Curtailment probability and number of samples grouped by FFNN model output in intervals of 0.05.

B Additional Tables

I	F	GDP	GDFF	WTP	WTFF
0	0	89.13%	90.66%	77.41%	77.12%
0	1	90.75%	91.71%	78.74%	78.65%
0	2	90.18%	90.85%	78.74%	79.60%
0	3	90.85%	91.33%	79.22%	80.46%
0	4	88.08%	87.42%	78.27%	78.27%
0	5	90.37%	67.78%	79.12%	79.50%
0	6	90.09%	89.32%	77.41%	77.03%
0	7	90.94%	89.80%	78.27%	78.74%
0	8	90.09%	90.94%	77.69%	75.41%
0	9	89.22%	87.21%	76.72%	74.90%
1	0	89.32%	90.47%	77.88%	77.98%
1	1	90.09%	90.18%	79.69%	79.12%
1	2	63.97%	90.37%	79.69%	79.60%
1	3	91.04%	91.13%	76.93%	77.79%
1	4	89.70%	90.37%	78.27%	77.22%
1	5	90.37%	91.23%	78.46%	78.27%
1	6	90.75%	89.23%	76.55%	77.03%
1	7	89.61%	91.42%	77.88%	79.03%
1	8	89.80%	90.47%	79.31%	78.74%
1	9	32.44%	90.08%	77.58%	73.66%

Table 10: Full results of ANN cross-validation. **I:** Iteration, **F:** Fold,**GDP:** Generation/Demand Perceptron,**GDFF:** Generation/Demand FFNN,**WTP:** Wind/Time Perceptron,**WTFF:** Wind/Time FFNN