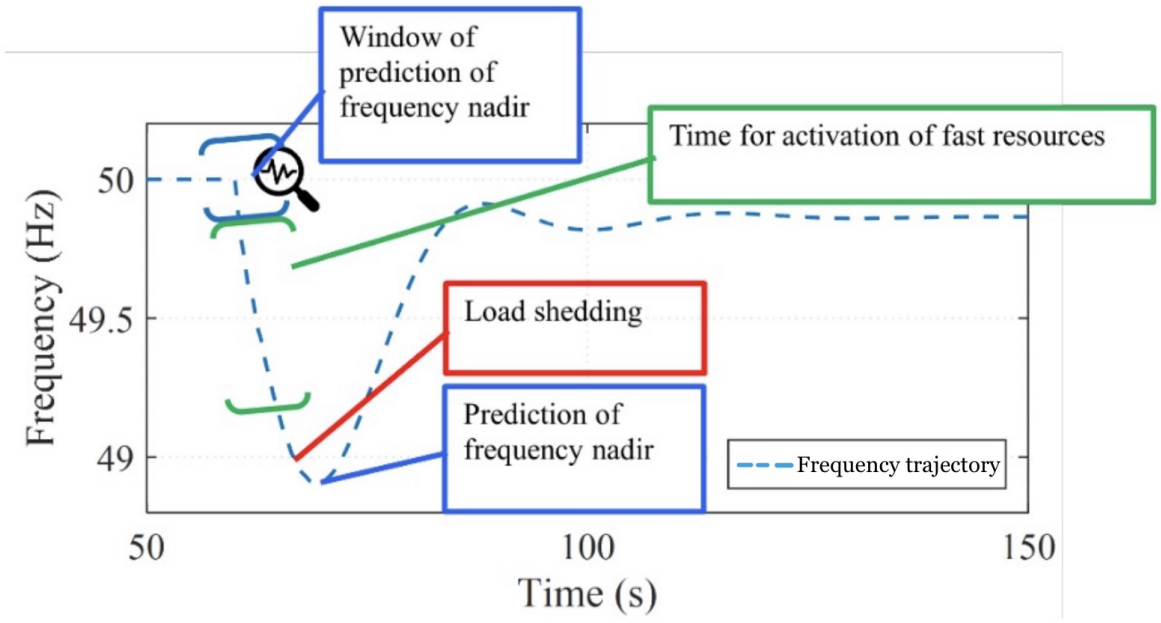
CHALMERS
UNIVERSITY OF TECHNOLOGY



Forecasting Frequency Nadir Using a Deep Recurrent Neural Network

Master of Science Thesis in Electric Power Engineering

JOHN BOJESTIG
KRISHNA ROHITH

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021

Master thesis EENX30

Forecasting Frequency Nadir Using a Deep Recurrent Neural Network

John Bojestig
Krishna Rohith

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2021

Forecasting Frequency Nadir Using a Deep Recurrent Neural Network
John Bojestig, Krishna Rohith

Department of Energy and Environment
Division of Electric Power Engineering
Chalmers University of Technology, Gothenburg, Sweden

Supervisor
Hannes Hagmar, PhD Student
Department of Energy and Environment
Division of Electric Power Engineering
Chalmers University of Technology, Gothenburg, Sweden
E-mail hannes.hagmar@chalmers.se

Supervisor
Claes Sandels, Researcher
Electric Power System Unit
RISE Research Institutes of Sweden
E-mail claes.sandels@ri.se

Examiner
Ola Carlson, Professor
Department of Energy and Environment
Division of Electric Power Engineering
Chalmers University of Technology, Gothenburg, Sweden
E-mail ola.carlson@chalmers.se

Master's Thesis 2021
Department of Electrical Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2021

Abstract

With an everyday growing share of intermittent renewable power generation, the inertia in the system is declining at a similar pace. Inertia refers to the energy stored in the rotating masses, which improves the stability of the power system. This presents a challenge in the area of frequency stability, primarily during disturbances. Traditional measures to improve frequency stability are likely to be insufficient in the future; hence there is a growing interest in additional solutions. A power system is designed to operate at a specified frequency, where bigger variations can harm the power system equipment and the loads, and in the worst-case, cause a blackout.

To ensure operation close to the nominal frequency, system operators utilise a reliable control system. In the event of disturbances, the capability to detect the onset of instability and initiate suitable control actions is of high importance. Typically, control actions such as activating power system reserves may take time before the effects are at full power. Thus, significant effort has been put into improving this. One method is through forecasts, provide information about the disturbance severity to the operator. An indicator of the severeness is the lowest frequency point during the disturbance, called frequency nadir. Receiving this information at an early stage enables the operator to act faster and more precisely, thereby enhancing the frequency stability in the system.

Frequency nadir forecasting have been conducted before, but to the authors' knowledge, only using data from a single time point. This project aims to develop a data-driven model that uses sequence data to forecast the frequency nadir. Since the model is data-driven and historical data from disturbances is scarce, data were simulated using a power simulation software called PSS®E 35.0. The model of choice is a recurrent neural network with long short-term memory, which is then compared to a baseline neural network model with input from one time point. Later, studies were performed to visualise how different data impacts the predictions by the model.

The results show that the model can forecast the frequency nadir with high accuracy with the generated data, and the use of sequence data reduces the prediction error up to 50%. Additionally, the model can make good predictions as soon as 0.1 seconds after the disturbance. The results also show that the model needs data from merely 10% of the system, easing the challenge of online implementation of the model. Given the data is available, the model can forecast the frequency nadir, thus improving the power system's frequency stability.

Keywords: *Nordic power system, frequency stability, frequency nadir, forecasting, machine learning, neural networks.*

Sammanfattning

Med en ständigt växande andel av icke-reglerbar förnybar kraftproduktion minskar trögheten i systemet i samma takt. Trögheten refererar till energin som finns i roterande massa, vilken förbättrar stabiliteten i elnätet. Detta leder till en utmaning för frekvensstabiliteten, framförallt vid störningar. Traditionella åtgärder för att förbättra frekvensstabiliteten kommer sannolikt att vara otillräckliga i framtiden. Därför finns det ett växande intresse för innovativa lösningar. Ett elnät är designat för en specifik frekvens, där större variationer kan skada elnätets utrustning och lasterna i nätet, och i värsta fall, orsaka en mörkläggning.

För att säkerställa den nominella frekvensen, systemansvariga använder ett tillförlitligt kontrollsystem. I händelse av en störning, är förmågan att upptäcka instabilitet och initiera lämpliga kontrollåtgärder mycket viktig. Normalt sätt tar det några sekunder innan kontrollåtgärder likt frekvensreserver når full effekt. Därav har betydande insatser gjorts för att försöka förbättra detta. En metod är att genom estimeringar, ge information om allvarlighetsgraden på störningen till den systemansvarige. In indikator på detta är den lägsta frekvensen under en störning, kallad frekvensnadir. Genom att mottaga denna information i ett tidigt skede efter störningen, möjliggör för en snabbare och mer precis motaktion, vilket förbättrar frekvensstabiliteten i elnätet.

Modeller som förutser den lägsta frekvensen under en störning har genomförts tidigare, men enligt författarnas vetskap har de gjorts med hjälp av data från en enda tidpunkt. Detta projekt syftar till att utveckla en datadriven modell som använder sekvensdata för att förutsäga den lägsta frekvensen. Eftersom modellen är datadriven och historisk data från störningar är knappa simulerades data med en programvara för kraftsimuleringar som kallas PSS®E 35.0. Den valda modellen är ett Recurrent Neural Network (RNN) med Long Short-Term Memory (LSTM) som jämförs med ett neuralt nätverk med data från en tidpunkt. Senare utfördes studier för att visualisera hur datan påverkar estimeringar av modellen.

Resultaten visar att modellen kan estimera den lägsta frekvensen med hög noggrannhet med simulerad data och användningen av sekvensdata minskar estimeringsfelet med upp till 50%. Dessutom kan modellen göra bra förutsägelser så snart som 0.1 sekunder efter störningen. Resultaten visar också att modellen behöver bara data från 10% av systemet, vilket underlättar utmaningen att implementera modellen online. Om datan är tillgänglig, kan modellen i ett tidigt stadie estimera den lägsta frekvensen och därmed förbättra elnätets frekvensstabilitet.

Keywords: *Nordiska elsystemet, frekvensstabilitet, frekvensminimum, prognostisering, maskininlärning, neurala nätverk.*

Acknowledgements

We would like to start off by thanking our supervisor Hannes Hagmar for being our guiding light through the last six months, and also for being our inspiration for an extended interest in power systems. We also thank Claes Sandels from RISE Research Institutes of Sweden, for giving us the opportunity to work on this thesis. A kind thanks to our examiner Professor Ola Carlson for his guidance, support, and for having interesting discussions with us throughout this period. A special thanks goes to Mattias Persson at RISE for everything inertia and for providing us with resources that have given us a broader perspective on frequency stability. A final thanks to Erik Weihs for putting things in a different perspective. A big thanks to our supervisors, advisors, and examiner again, for making the last few months super fun and thank you for the weekly meetings. We are truly grateful to everyone.

A few words from Krishna:
I would like to thank my family, for being my pillars of strength during this pandemic times. A huge thanks to my thesis partner and bror, John Bojestig for being the best to work with, for all the mugs of kaffe and all the awesome times at Chalmers. Thank you! I also thank my friends Daniel Åkerberg, Srivatsava Tangirala, Billa Jai Bhargav and Kelechi Jude Egbe for being there and for their never ending support and encouragement. I am glad to have you all.

A few words from John:
On a day like this, I am grateful that the joy of learning and making an impact brought me here. All due credits to my family and friends that never gave any directive hands but were always full of support. Thank you! And finally, to the true warrior and thesis partner Krishna, you inspire me. I am thankful for the time we shared.

Krishna Rohith & John Bojestig, Gothenburg, June 2021.

List of Acronyms

aFRR	Automatic Frequency Restoration Reserve
ANN	Artifical Neural Network
FCR-D	Frequency Containment Reserve for Disturbed operation
FCR-N	Frequency Containment Reserve for Normal operation
FFR	Fast Frequency Reserve
FN	Frequency Nadir
FS	Frequency Stability
LSTM	Long Short-Term Memory
LTFS	Long-Term Frequency Stability
MAE	Mean Absolute Error
mFRR	Manual Frequency Restoration Reserve
ML	Machine Learning
MSE	Mean Squared Error
NN	Neural Network
NPS	Nordic Power System
PMU	Phasor Measurement Unit
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
RoCoF	Rate of Change of Frequency
RQ	Research Question
SQ	Sub-Question
STFS	Short-Term Frequency Stability
UFLS	Under Frequency Load Shedding

Contents

1

Introduction

This chapter presents a background and a problem overview of the thesis, along with a review of previous work in the research field. Furthermore, the research methodology is briefly presented, and the scope, limitations and sustainable aspects are introduced.

1.1 Background

To tackle the challenge of climate change, the power systems need to transition into being more sustainable. With a growing share of renewable power generation, the inertia in the system, which is the energy stored in the rotating masses, is declining at a similar pace. Conventional generation from large synchronous generators has traditionally provided the power system with the inertia that can support the system frequency in the event of larger disturbances. Typically, power generation from renewable energy sources such as wind and solar provides significantly lesser inertia to the system. Frequency stability refers to the capability of a power system to maintain its system frequency close to the nominal value, even during the occurrence of a large disturbance. If the frequency deviates too much from the nominal value, it can lead to system isolation, customer outages, or in the worst-case scenario, a blackout. With less inertia, maintaining the frequency close to its nominal value is more challenging, leading to a growing interest in other solutions to improve frequency stability.

To aid and allow the continuous development of renewable energy sources, system operators need better tools to handle frequency instability events. The common practice today is to activate control measures when the frequency goes below a predefined value. A problem with this approach is that the control actions cannot act instantaneously. Instead, predefined values need to be set higher to make the control measures act on time, resulting in the possibility of the actions being less powerful and precise. An alternative approach to handling this problem and improving the frequency stability during disturbances is to predict the lowest frequency point right after their occurrence. By doing so, control measures can be activated faster and with greater precision, and the power system can be safely operated with less inertia.

1.2 Problem overview

When a disturbance occurs, the power system's frequency drops, and around 8.7 seconds later [1], it starts to recover back again. The lowest point of frequency after the disturbance occurs is an indicator for frequency stability and is called the Frequency Nadir (FN). The system operator is responsible for maintaining the frequency at the nominal level. If they have information about the FN at an early stage after the disturbance, this can enable them to act quicker and more precisely, thus improving the system's frequency response. This can be achieved by forecasting the FN during the initial seconds after the disturbance. Mapping the data from the initial stage during the disturbance to the FN is a complex task. Therefore, Machine Learning (ML) models are of interest due to their capability of approximating non-linear functions when trained on large sets of data. Unfortunately, or luckily, historical data of disturbances is scarce, which limits the available data the ML models can be trained on. Hence, training data for the ML models need to be simulated using power simulations software tools and dynamic models of the actual power system. Similar models using data from Phasor Measurement Units (PMU) are not widely used today, to the authors' knowledge. PMUs are an emerging technology in power systems that can deliver fast and time synchronised data with high frequency. This will hopefully enable the implementation of FN forecasting models online in the near future.

Based on the problem overview, this thesis will examine the following Research Question (RQ) with the corresponding Sub-Question (SQ):

- **RQ:** How is the capability of a chosen ML-model to forecast the frequency nadir after a larger (N-1) disturbance?
- **SQ:** What data does the model require from the PMU network?

1.3 Research methodology

A literature review will be conducted to choose a suitable ML model and provide insights into what areas require further study. To be able to evaluate the chosen model, a baseline model will also be developed. Data generation will be performed through simulations to get a sufficient amount of training and testing data. When the model is developed, a case study will be conducted to evaluate the data required from the PMUs. The evaluation of the model and the case study will ultimately shape the answers to the RQ and SQ.

1.4 Scope and limitations

The scope of this thesis is to develop an ML model that can forecast the FN and thereby support the system operator to improve their frequency response. The data for training and testing the model will be simulated on a test grid, and testing it on a real power system is out of scope. Moreover, the data the model requires from the PMUs to operate online will be evaluated.

1.5 Sustainable aspects

Precisely knowing what the FN will be when a significant disturbance occurs can reduce the interrupted access to electricity, translating into a positive social impact on the end consumers. Furthermore, the FN forecasting model might improve renewable energy integration into the power system, which in ecological terms is a win-win. When validating the model, there is a risk that data sets which yield a better result for the thesis might be hand-picked. This selective bias is a crucial ethical aspect that needs to be consciously avoided. Therefore, emphasis has to be placed hard upon validating the results. It is also vital to not over-promise and under-deliver.

1.6 Literature review

Several studies have tried to provide system operators with information that can improve the power system stability. This has been done by either forecasting the frequency trajectory or the FN after a disturbance. The latter has gained the most attention in the literature. It is a good indicator of the frequency stability during disturbances and can provide valuable extra time to handle them. Therefore, in this section, the main focus will be to evaluate how the past literature has forecast the FN.

Methods [2] and [3] assume there are linear relationships between input parameters and the FN. In [2], a simplified linear regression model is developed for online use in the Energy Management System (EMS), with a focus on simplicity. The motivation for choosing a simplified model is an easier implementation online. It may reduce the need for data, preferable since all the data might not be easily available or fully accurate. If data with uncertainty needs to be used, it can worsen the forecast. The emphasis is on the maximum instantaneous frequency deviations for dimensioning incidents, both over and under. The model is designed to continuously forecast the worst-case scenario to monitor if the system operates in a safe manner. Accordingly, the regression model is divided into two subsequent parts. Over and under frequencies are forecasted separately as the transfer functions of the primary reserves differ during these events. Due to the simplicity of the model, it has been able to be implemented online in a real power system.

However, power systems are non-linear systems, and the previous approach is a simplification of the problem. The drawback of using non-linear modelling is time-consuming forecasts. Therefore, ML methods have been proposed to deal with this task. The benefits of using ML methods lie in their ability to map the input to the output data and perform forecasts rapidly. In [4], a Neural Network (NN) was used to forecast the FN in the system during disturbances. The study mainly focused on the input to the model. Four different inputs are considered, namely the inertia constant, size of disturbance, voltage dynamics, and frequency dynamics. Different combinations of these inputs are tried, and the prediction times depend significantly on the choice of inputs. The assumption is that in the near future, the disturbance size will be available for the system operator in 2 seconds. The rest of the parameters are assumed to be available right after the disturbance. A scenario when the inertia constant is not available and needs to be estimated is examined. The approach used to estimate the inertia constant resulted in a prediction time of around 2 seconds. Even if the inertia estimation is bad, the information can improve the forecast by the neural network. Interestingly, the result shows that the FN can be predicted without knowing the inertia constant or the size of the disturbance. Meaning, two parameters that are expected to have a great impact on the FN, but might be slow to be available, can be excluded.

Moreover, the result in [4] showed that the NN benefits from using data from different time points in a sequence. Even though they show promising results, no other studies were found using it to forecast the FN during the literature study. When a sequence of time steps is given to each input feature, the NN can gain valuable information from the parameters' changes during the disturbance. For instance, inertia is vital for the frequency response but is seldom readily available. Instead, initial changes in frequency can indicate the inertia in the system and help the NN model forecast the FN. It should also be noted that a NN is not the state-of-the-art ML model to use for sequence data.

It is good to note that most papers have not considered Under Frequency Load Shedding (UFLS) when forecasting. This can be attributed to the power systems they use, where UFLS is used as the last resort, which must be avoided. If a UFLS scheme is used more regularly, the load shedding significantly impacts the FN. To predict the dynamics with UFLS, the model either has to be more complex [5] or, as in [6], divide the task into sub-models that will create an easier forecasting task for each sub-model. Note that frequency dips down to 48 Hz have been considered, which can be compared to [4], where the lowest dip is around 49.4 Hz. These differences can affect the dynamics that are to be predicted, especially in the case where a single generator runs out of kinetic energy before FN is reached.

Several models in the literature are data-driven. The historical data from disturbances is, for most models, not enough due to the lack of numerous large disturbances. Therefore, simulated data generation becomes a crucial part. Ideally, the simulated data should mimic real data with all possible disturbance scenarios. As power systems are complex, simplifications are done to reduce the data size and time spent on the task. The simulated power systems used in research are reduced test

systems, varying in size. Validation is often done on simulations from the same test system. There is also a significant difference in the number of data samples among the models. In general, simplifications have an impact on the prediction results. For example, in [7], contingencies on two generators in the system are simulated, while in [8], over 100 different contingencies are applied, including various combinations of disturbances. The first one is probably an easier forecasting task since the testing data is also from the simulated data, consisting of just these two types of disturbances. The latter likely performs better on a test on historical data since it has learned to forecast more scenarios. The lack of historical data of disturbances makes these tests hard to implement.

1.7 Thesis structure

The rest of the thesis is organized as follows:

- **Chapter 2** provides a theoretical background on frequency stability with a closer a look the the factors affecting it and also gives an insight into the system behaviour during a contingency.
- **Chapter 3** gives a theoretical background on machine learning with a focus on neural networks and specifically recurrent neural networks with long short-term memory.
- **Chapter 4** presents a method for forecasting the frequency nadir in the event of a disturbance, and introduces the data generation process.
- **Chapter 5** provides and discusses the results.
- **Chapter 6** gives the main conclusions of the thesis and ideas about future work are presented.

2

Frequency stability

This chapter provides a theoretical background to frequency stability. The general concept of stability in the power system is presented together with a more detailed presentation of frequency stability, especially during disturbances. This concludes with the challenges in frequency stability and possible solutions.

2.1 Power system stability

Power system stability is an important factor to consider when talking about the day-to-day operation of a power system. According to the IEEE-CIGRE Joint Task Force on Stability Terms and Definitions, stability is defined as *"the ability of an electric power system, for a given initial operating condition, to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact"* [9].

Power system stability can be broadly classified into five main categories based on the considerations in [10], namely Rotor angle stability, Frequency Stability (FS), Voltage stability, Converter driven stability and Resonance stability. These classifications can be seen in Figure 2.1. Frequency stability, which is the stability factor of concern, is defined as *the ability of a power system to maintain steady frequency following a severe system disturbance resulting in a significant imbalance between generation and load* [11].

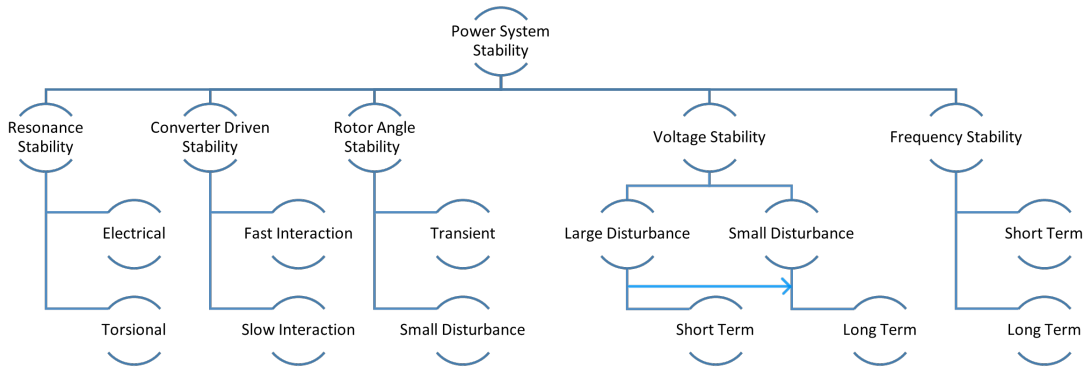


Figure 2.1: Classification of Power System Stability [11].

FS is further divided into two categories, short-term frequency stability (STFS) and long-term frequency Stability (LTFS). The focus of this thesis is on the STFS. Short-term in the case of FS can be fit into a time frame that ranges from $0 - 10$ seconds after the occurrence of a disturbance [10], which includes both the sub-transient and transient periods. STFS deals with bringing the system frequency back to the operating range when a sudden severe disturbance, such as a generator outage or a load disconnection, occurs in the system, which brings about a drastic drop in the frequency. LTFS is concerned with maintaining the system frequency in the operating range.

In numerical terms, a small disturbance such as a single transmission line trip can drop the system frequency below 49.9 Hz, or a disturbance in the form of a single line to ground fault drives the bus voltages to zero, affecting the power flow, thereby the frequency. On the other hand, a larger disturbance such as a generator trip causes the frequency to go down even further, sometimes to dangerously low levels ($\leq$ 49.4 Hz in the NPS). This can be attributed to the power mismatch between generation and consumption. Until the power balance is restored, the frequency keeps dropping. The generators are not designed to operate at lower frequencies, and there is a risk that they break down at frequencies below 49 Hz. If a generator trips or disconnects to protect itself, the frequency drops further, resulting in a cascading effect of disconnected generators, leading to an outage. Therefore, in the NPS, loads are disconnected around 49 Hz with a load shedding scheme to avoid a blackout.

In the day-to-day operation of a power system, the frequency operates within stable bounds, yet it is not constant. It floats continuously around the 50 Hz mark, with a stable operating range of 49.9 Hz $\geq f \leq$ 50.1 Hz. A frequency less than 49.9 Hz is called under-frequency, and greater than 50.1 Hz is called over-frequency. Power consumption changes continuously, and the generation tries to match it to maintain power balance in the system and keep the frequency at 50 Hz.

A few indicators in the power system give an idea about the state of the frequency stability. Two of the most interesting indicators are the FN and the RoCoF, illustrated in Figure 2.2. Here, f_n is the nominal frequency and f_{ss} is the steady-state frequency after the disturbance at a lower frequency level than the nominal one. Frequency nadir, shown as f_{nadir} in the figure, is defined as *the minimum value of frequency reached during the transient period* [12]. Having information about the FN can be valuable to determine the state of the frequency stability by checking the margin to unsafe frequency levels. RoCoF is the instantaneous rate of change of frequency and can give a picture of the system's robustness, where smaller changes are preferable.

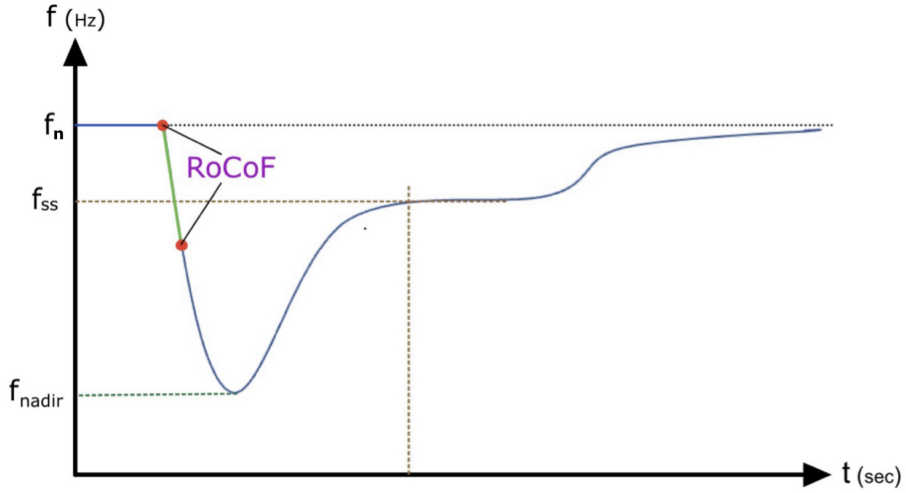


Figure 2.2: Frequency trajectory during a contingency.

2.2 Swing equation

The fundamental of maintaining frequency stability in the power system is balancing power production and consumption. The system operators monitor this balance by observing the frequency, since its derivative indicates a power mismatch. When there is a contingency in the system, i.e. a disconnection of generation or load, the frequency changes due to an active power mismatch in generation and consumption [3]. The initial reaction of the frequency can be expressed by the swing equation, which describes the motion of a rotating mass:

$$H_i \frac{df_i}{dt} = \frac{f_n^2}{2S_{ni}f_i}(P_{mi} - P_{ei}) \tag{2.1}$$

where:

- H_i : per unit inertia constant of generator i, also known as H constant.
- f_i : frequency of the generator i
- f_n : nominal frequency
- S_{ni} : apparent power of the generator i
- P_{mi} : mechanical power of generator i
- P_{ei} : electrical power of generator i

The H constant is an important quantity and can be formulated as:

$$H = \frac{\text{kinetic energy in MJ at rated speed}}{\text{machine rating in MVA}} = \frac{E_K}{S_B} \tag{2.2}$$

with the unit of H in MWs/MVA, often only expressed in seconds.

The swing equation (2.1) shows that a frequency derivative indicates an imbalance in the generator between mechanical and electrical power. Apart from the power imbalance, the derivative is also dependent on the inertia constants of the generators. Where higher inertia corresponds to a lower frequency change, it is only the inertia that instantaneously can inject active power into the system and reduce the imbalance. Therefore, the initial decline rate of the frequency is strongly dependent on inertia [13].

2.3 Inertia

In the initial seconds after a disturbance, the inertia plays a key role in frequency stability by instantaneously providing kinetic energy to reduce the power mismatch. Traditionally, inertia has mainly been provided by rotational masses from synchronous generators, which are explained further in this section. Other methods to provide a similar service are described in Section 2.6.

During a disturbance, the energy stored in the rotational masses is released and damps the dynamics. The kinetic energy of the rotational mass is given as:

$$E_k = \frac{1}{2} J w_m^2 \tag{2.3}$$

where J is the moment of inertia expressed in $kg \cdot m^2$ and w_m is the rated rotational speed in rad/s. The inertia constant H tells how many seconds the generator's kinetic energy can provide rated power determined by:

$$H = \frac{E_k}{S_B} \tag{2.4}$$

where S_B is the rated power and H expressed in seconds. Typical values for the inertia constant H varies between 2-10 seconds. In a power system, the system inertia constant is often of interest and is calculated by:

$$H_{sys} = \frac{\sum_{i=1}^{N} H_i S_{B_i}}{\sum_{i=1}^{N} S_{B_i}} \tag{2.5}$$

with N generators. Then the total kinetic energy in the system can be given by:

$$E_{k,sys} = H_{sys} \sum_{i=1}^{N} S_{B_i} \tag{2.6}$$

2.4 System response to a contingency

During a contingency, the frequency starts to drop due to a mismatch in active power. The power system responds by first reducing the mismatch and later restoring the frequency to its nominal value. A successful frequency trajectory after a contingency is shown in Figure 2.2 above.

The mismatch in power can be reduced by either increasing the generation or reducing the loads, where the priority is always to provide electricity to the customers. UFLS though, can be used as the last measure in emergencies to avoid a frequency collapse, where parts of the load get disconnected to save the rest of the system [5]. A common initial response by the power system to a disturbance can be seen in Figure 2.3. During the initial seconds after a contingency, the inertia provides most of the missing power. Frequency Containment Reserve for Disturbed operation (FCR-D) gradually increases and later covers most of the mismatch. The FCR-D cannot act instantaneously for two reasons: firstly, the reserves get the activation signal when a frequency threshold is reached, usually at 49.9 Hz, and secondly, due to mechanical delays before the reserves get fully activated after getting the signal. When the frequency nadir is reached and the frequency starts to rise, the inertial response becomes negative. There is also self-regulation of loads in the system, where the load is reduced when the frequency is dropping. This phenomenon partly improves the frequency stability by reducing the power mismatch [3].

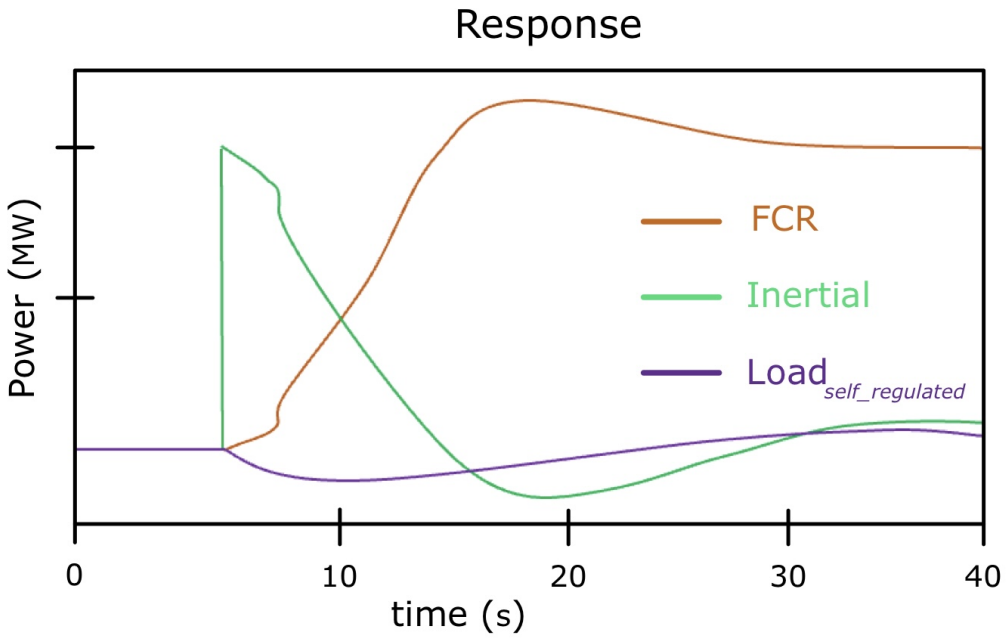


Figure 2.3: Initial system response to a contingency [14].

2.5 Frequency control

The electricity market matches generation on an hourly basis to a forecasted load on a day-ahead market. Even if the forecasted load is correct, there will be load fluctuations within the operational hour, and a control system is needed to keep the continuous balance. Often various frequency control services are used together to maintain frequency at a stable value and restore it to its nominal value if it deviates [15]. Figure 2.4 highlights the frequency control services in the Nordic market and their division into primary, secondary and tertiary controls.

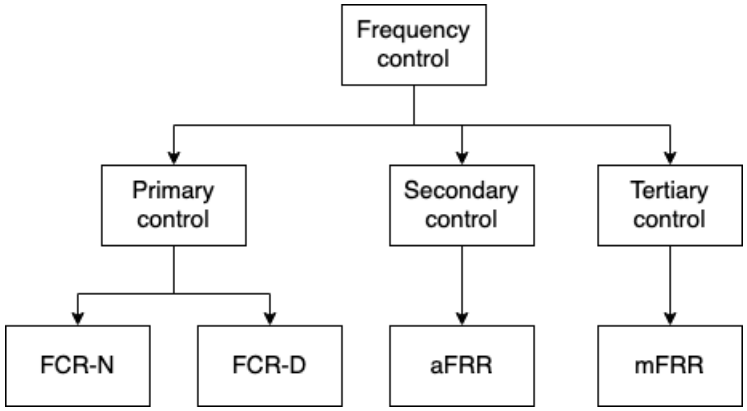


Figure 2.4: Frequency control services in the Nordic market [15].

Primary frequency control services automatically act to stabilise the frequency. These services are provided in seconds during a disturbance and up to a minute during a steady period. Frequency Containment Reserve for Normal operation (FCR-N) is used to continuously compensate for load changes within the operational hour and keep the frequency between 49.9 Hz and 50.1 Hz. Therefore, FCR-N is responsible for the small-signal stability in the normal operational bandwidth. During disturbances, meaning the frequency exceeds the limits for FCR-N, FCR-D gets activated to minimize the transient coming from the disturbance. FCR-D can bring the frequency both up and down, but in this theory section, the focus is on bringing the frequency up. Furthermore, FCR-D is responsible for the small-signal stability below 49.9 Hz. To ensure this, the FCR-D is not bringing the frequency back to the nominal value. Instead, it only stabilises the frequency at a lower level. Hence, frequency control is dependent on secondary and tertiary controls to restore the frequency using balancing reserves [16]. The whole process of the control services during a disturbance is illustrated in Figure 2.5.

To ensure that the frequency is stable, the NPS needs enough reserves to be available every hour. The countries connected to the Nordic grid share the responsibility to provide the reserves, where each country covers reserves that correspond to their part of the Nordic energy consumption. How the reserves are provided can vary among these countries. In Sweden, the reserves are decided on separate markets for support services. The technical requirement for FCR-D is to have a linear activation between 49.9 to 49.5 Hz, where the reserves have to be 50% activated in 5 seconds, 100% activated in 30 seconds and operated for at least 20 minutes. The system is

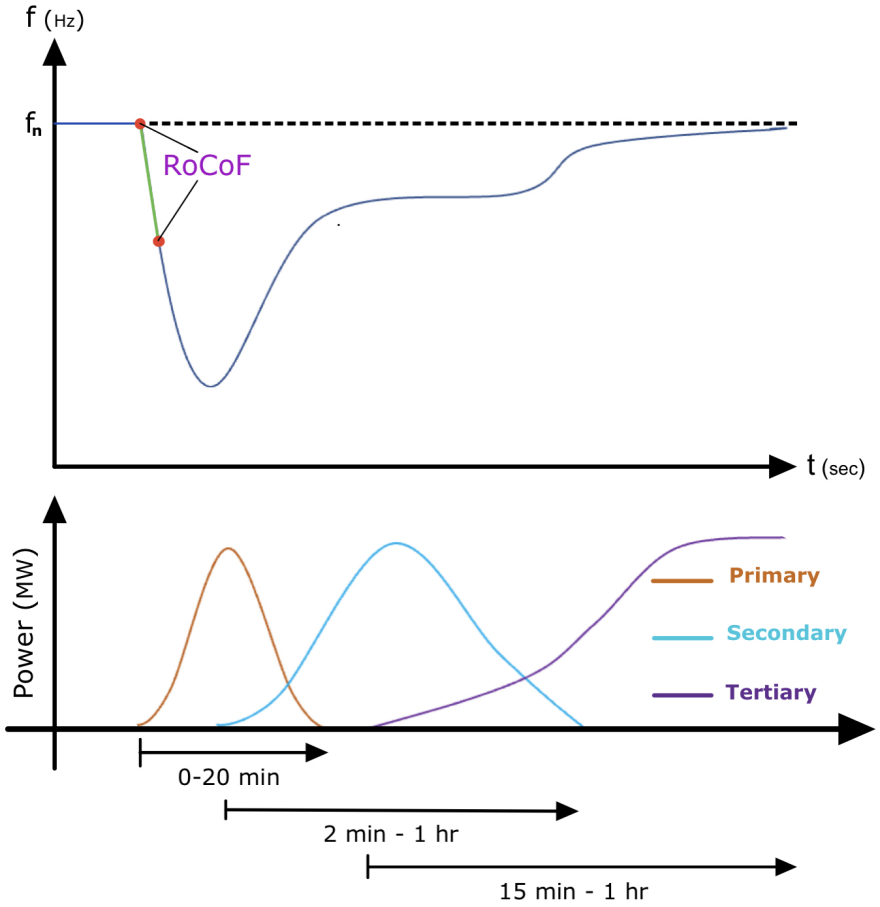


Figure 2.5: Response by the control services during a disturbance [17].

designed based on the N-1 criteria, which states that the system must withstand a single component outage and still maintain stability. In this case, the system has to be able to handle the disconnection of the biggest generator. Thereby, the amount of FCR-D reserves required depends on the biggest generator online at each moment [17].

When the primary control has stabilised the frequency, it is up to the balancing reserves to restore the frequency to the nominal value. There is an automatic service called automatic Frequency Restoration Reserve (aFRR) in secondary control, which generally acts in a few minutes. The last part is tertiary control, which has a similar service to the one before, except that it is non-automatic and is called manual Frequency Restoration Reserve (mFRR). In Sweden, this service takes between 12-15 minutes to activate. These services need to provide reserves during disturbances and when there is a mismatch between the actual and forecasted loads. The mismatch can either be more short-term, i.e. during mornings or evenings when the load is changing rapidly, or more long-term due to incorrect weather predictions. The short-term mismatches are handled by aFRR. The mFRR's purpose is to support the system and restore the frequency during long-term mispredictions and disturbances, i.e. a disconnected generator [18].

2.6 Future of frequency stability

Traditionally, reserves have been the way to solve frequency stability issues by simply adding more or adjusting how they work. Since the inertia is dropping and appears to continue in this trajectory in the future, reserves alone cannot solve this. The reason is because of certain mechanical delays in the generators. There is a risk of overcompensation with bigger reserves that leads to reserves needing to push down the frequency again. The delays, together with overcompensation, can result in unwanted frequency oscillations [19].

The main factors to keep the frequency within the operating range are rotating mass, the active power and the dimensioning incident [20]. As already mentioned, the rotating mass is reducing, and the traditional increase of reserves can be counterproductive. The dimensioning fault is the biggest fault the system is designed to handle, usually a loss of the largest generator. By reducing the biggest fault, the stability improves, and fewer reserves need to be available. This should only be done if it is necessary and can be motivated socio-economically [19]. Since the traditional measures to improve frequency stability might be insufficient, alternative methods have gained more interest. Possible solutions to improve the main factors which keep the frequency stable are shown in Figure 2.6. Some of the promising methods are fast frequency reserves, synthetic inertia and battery energy storage systems.

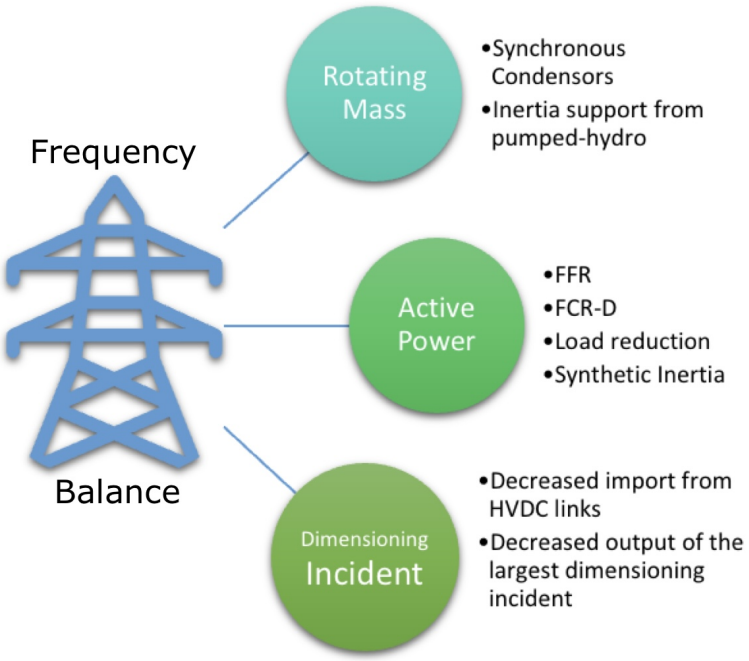


Figure 2.6: Key measures to improve frequency stability [3].

Fast Frequency Reserve (FFR) is a system frequency control service for disturbances that deliver a fast power injection that can mitigate the effect of reduced inertial response and maintain frequency stability. The FFR units offer a similar response to FCR-D but can respond faster. FFR resources are usually activated by a few triggers such as a frequency deviation of 50 Hz ± 0.1 Hz (under and over frequencies), a RoCoF value of $df/dt \leq -0.035$ Hz/s or ≥ 0.035 Hz/s, them both or, even by a circuit breaker or relay signal [3]. The FFR activation speed depends upon the source that delivers the reserve. It can then either be a power boost scheme such as load shedding or controlled to provide reserves, proportional to the frequency deviation as:

$$\Delta P_e = -K_p \Delta f \tag{2.7}$$

where ΔP_e is the power provided by the FFR, and K_p is the proportionality constant.

So far, we have seen that the traditional inertia with rotational masses provides the required energy when the frequency goes down. The same service can also be provided without any rotational mass, called **Synthetic Inertia**. Increasing renewable production by units connected in a non-synchronous manner and decreasing nuclear plants result in less rotational inertia in the system. An example of such a renewable power unit is a wind turbine generator connected to the system via power electronic converters and thus does not contribute to rotational inertia. Therefore, an option that can mimic the traditional inertial response without any rotational masses is of interest. This is done using a controller that emulates the active power response similar to a traditional synchronous generator. For example, the usage of a controller can capture the kinetic energy from both the rotor of a wind turbine and its blades by measuring the RoCoF and then applying a proportional amount of electrical torque on the rotor, which slows the rotor. This process leads to the rotor releasing the extracted kinetic energy in a similar manner to a conventional generator [3, 14].

Battery Energy Storage Systems (BESS) can be used to provide power almost instantaneously by rapidly discharging in a fraction of a second, much faster than a conventional thermal plant. This enables them to improve the frequency stability of the system by providing a speedy response to unpredictable variations in generation and demand, among other services such as regulation and load ramping [21].

3

Machine learning

In this chapter, a theoretical background of the machine learning methods used in the study is presented. The general structure and working principle of neural networks, recurrent neural networks and long short-term memory are described.

3.1 Neural network

Artificial neural networks (ANNs), also called neural networks (NNs), is inspired by how the neurons in the brain can learn from experiences to, for example, classify or predict future events. A neural network is a supervised learning algorithm that learns to map the input to the output by training on input-output data examples. The availability of more data and faster computation in recent years have raised the attention of NNs. This, combined with the NNs' capability of using large data sets and learning non-linear relations, creates new possibilities in many areas [22].

A neural network consists of many connected nodes called artificial neurons, which send signals to each other, similar to the human brain's synapses. An example of a feedforward NN with one hidden layer can be seen in Figure 3.1. A network comprises different layers of neurons (illustrated by circles), where a neuron can only connect to other neurons in the layer right before and after. Data fed into the input layer is passed through each layer until it produces the estimation in the output layer. Between the input and output, there are hidden layers. Each neuron has a bias weight $\boldsymbol{b}$ and an activation function σ, which together with the weights of the connections $\boldsymbol{W}$, control the output. The network's estimation is compared to a target value during training, where the errors are used to update the weights. The network is updated in numerous minor steps till it cannot improve further. To do this training, a large set of data is required [23].

3.1.1 The forward pass

The process of going from the input to output in a NN is called the forward pass, which is used both in training and with a trained network to perform predictions. The process of a forward pass is presented below, following [24]. As can be seen in Figure 3.1, a row vector $\boldsymbol{x} \in \mathbb{R}^n$ of input data is feed into the input layer. Thereafter, the vector is passed through the weight matrix $\boldsymbol{W_1} \in \mathbb{R}^{n \times k}$ via the connection lines

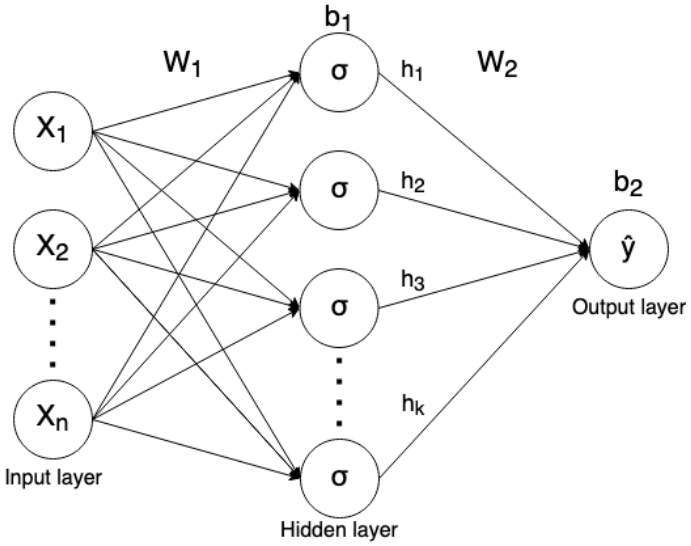


Figure 3.1: Neural network with a single hidden layer and one output.

and gets added to the bias weight $\boldsymbol{b_1} \in \mathbb{R}^k$ in the hidden neurons. The superscripts n and k represents the length of the input data vector and number of neurons in the first layer, respectively. The sum of these proceeds into the activation function in the hidden layer, and the output from the layer will be:

$$\boldsymbol{h}(\boldsymbol{x}) = \sigma(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1) \tag{3.1}$$

where σ is, in most cases, a non-linear activation function. Frequently used activation functions are Sigmoid, Tanh, or the Rectified Linear Unit (ReLU) seen in Figure 3.2 [22]. The outputs from the hidden layer go through the second weight

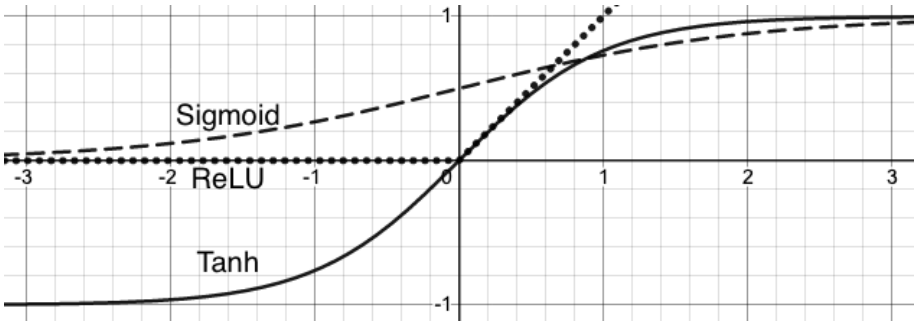


Figure 3.2: Activation functions. Dashed line Sigmoid, dotted line Relu and standard line Tanh.

matrix $\boldsymbol{W}_2$ and together with the bias weight $\boldsymbol{b}_2 \in \mathbb{R}$ compute the final output $\boldsymbol{z}$ as:

$$\boldsymbol{z}(\boldsymbol{h}) = \boldsymbol{W}_2\boldsymbol{h} + \boldsymbol{b}_2 \tag{3.2}$$

The last step is to apply the output to an activation function to generate the prediction:

$$\hat{\boldsymbol{y}} = \sigma(\boldsymbol{z}) \tag{3.3}$$

In a regression task, if the output range is from $-\infty$ to ∞, then the last activation function must be a linear one.

3.1.2 Training

In training, the model learns to fit the patterns in the training data, but the goal is to perform on unseen data. The first step to assure this is to split the available data into training, validation and testing data. As the names suggest, training data is used for training, and the model is repeatedly tested on validation data during training to check if the model improves on unseen data. If the model improves in regards to training data but gets worse with validation data, it means that the generalization is poor, referred to as overfitting. There are many approaches to avoid overfitting, where a simple one is to detect when it happens and stop the training since more training makes the model worse on unseen data, referred to as early stopping. When training is done, the model is finally evaluated with the test data [25].

The NN learns to map the input to the output target by updating the network's weights. The update starts with the loss from a suitable loss function $J(\hat{\boldsymbol{y}}, \boldsymbol{y})$. For a regression task, the Mean Squared Error (MSE) is commonly used. In multi-classification tasks, multi-class cross-entropy loss is often used. The weights are not updated after each error; instead, the average of a collection of errors are used for the update, expressed in the cost function J:

$$J = \frac{1}{S} \sum_{i=1}^{N} L(\hat{\boldsymbol{y}}, \boldsymbol{y}) \tag{3.4}$$

where S is the number of samples used before updating. Usage of the errors from the whole training set before updating the weights is called batch gradient descent. As the name hints, the weights are updated by a gradient-based optimization, whose aim is to reduce the cost J as much as possible. The gradient of J in (3.4) with respect to each weight in the network shows the impact on the cost function if the weights are changed. Based on this, a method called backpropagation tunes the parameters iteratively in small steps to minimize the cost function J [24]. With batch gradient descent, the computer can encounter problems with handling all the training sets in the memory, resulting in slower training. Instead, a common solution is to use mini-batch gradient descent, where the data is divided into smaller batches, and the model gets updated after each batch. The use of the batches from the entire training set is called one epoch. Usually, several epochs are used during training to achieve high accuracy [22].

Overfitting is a problem in deep machine learning models. The ideal, but near impossible, goal of generalization is to have low bias and variance. In reality, it is common to trade-off between these two through a method called regularization. Consequently, the model will fit the training data worse but can result in better performance on unseen data. It can be achieved by adding a regularization term to the cost function J that penalizes larger weights. The penalty decides by a parameter that requires tuning. Too low can lead to overfitting and high to underfitting, meaning the model will fit neither the training nor test data. Another way to regularize is to randomly eliminate a given percentage of the connections each time, making the model less reliant on a small part of the connections, called dropout [23].

3.2 Recurrent Neural Network

In the previous section, the model has used input parameters from one time step. In some applications, the prediction accuracy can improve by instead using several time steps. A Recurrent Neural Network (RNN) is a class of ANNs, where sequences of input data are used to take advantage of current and past data in predictions. RNNs have been applied in a wide range of applications in power systems, from instability predictions [26] and load forecasting [27] to power maximization control for wind power plants [28].

As already mentioned, feedforward NNs pass the data from input to output in one direction without containing any information about events in a sequence. The RNN works similarly, with the RNN block as the main difference, where information about earlier time steps is saved and passed to the next block. The structure of a sequence of RNN blocks is shown in Figure 3.3, where t is the number of time steps in the sequence and h_t is the output from the previous time step. The structure of the input and output of an RNN can range all the way from one input - one output to many inputs - many outputs. The RNN sequence in the figure consists of t smaller blocks, which all contain weights, connections, and activation functions. The blocks' structures can vary from a simple neuron layer to a more advanced Gated Recurrent Unit (GRU) or a Long Short-Term Memory (LSTM) [29]. In Section 3.1.2, the concept of

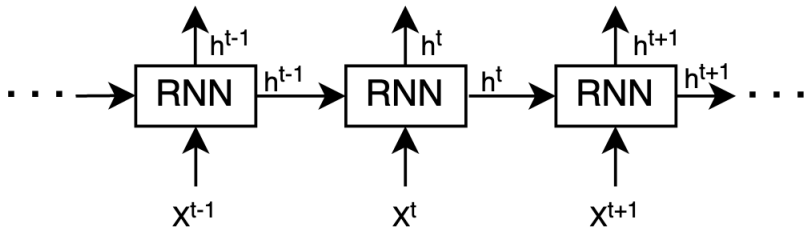


Figure 3.3: A sequence of RNN blocks.

backpropagation is explained. RNNs utilize a similar approach which also considers time dependencies, called backpropagation-through-time, that propagates through all the time steps in an RNN-sequence. In a simple RNN, the backpropagation-through-time can lead to vanishing or exploding gradients when dealing with long sequences. This implies that the gradients can either go to zero or towards infinity, resulting in poor estimations [30].

3.2.1 Long short-term memory

In 1997, Hochreiter & Schmidhuber introduced LSTM to tackle the vanishing gradient problem of RNNs. An LSTM can handle vanishing and exploding gradients better due to its cell state, which acts as long-term memory. Making it possible for efficient training with backpropagation-through-time due to uninterrupted gradient flow. This results in a reduction of the likelihood of vanishing or exploding gradients [31].

Although many varieties of LSTM exist, the most common one was proposed in 2000 with the introduction of the forget gate [32], illustrated in Figure 3.4. It is important to stress that an LSTM block consists of h number of LSTM cells. Thereby, vector notation is used, where $\boldsymbol{f}^t$ refers to h number of LSTM cells. The operation of an LSTM block can be expressed with the following formulas [33]:

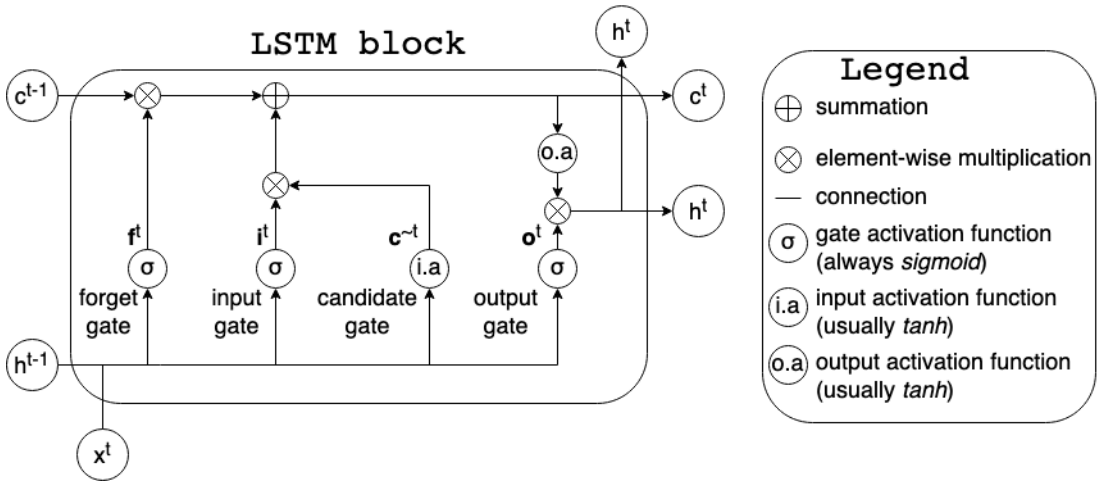


Figure 3.4: Illustration of an LSTM block.

$$\boldsymbol{f}^t = \sigma(\boldsymbol{W}_f \boldsymbol{x}^t + \boldsymbol{U}_f \boldsymbol{h}^{t-1} + \boldsymbol{b}_f) \tag{3.5}$$

$$\boldsymbol{i}^t = \sigma(\boldsymbol{W}_i \boldsymbol{x}^t + \boldsymbol{U}_i \boldsymbol{h}^{t-1} + \boldsymbol{b}_i) \tag{3.6}$$

$$\tilde{\boldsymbol{c}}^t = tanh(\boldsymbol{W}_c \boldsymbol{x}^t + \boldsymbol{U}_c \boldsymbol{h}^{t-1} + \boldsymbol{b}_c) \tag{3.7}$$

$$\boldsymbol{c}^t = \boldsymbol{f}^t \odot \boldsymbol{c}^{t-1} + \boldsymbol{i}^t \odot \tilde{\boldsymbol{c}}^t \tag{3.8}$$

$$\boldsymbol{o}^t = \sigma(\boldsymbol{W}_o \boldsymbol{x}^t + \boldsymbol{U}_o \boldsymbol{h}^{t-1} + \boldsymbol{b}_o) \tag{3.9}$$

$$\boldsymbol{h}^t = \boldsymbol{o}^t \odot tanh(\boldsymbol{c}^t) \tag{3.10}$$

where $\boldsymbol{c}^0 = 0$ and $\boldsymbol{h}^0 = 0$ and $\odot$ are referring to the element-wise product.

- $\boldsymbol{x}^t \in \mathbb{R}^n$: input vector to the LSTM block
- $\boldsymbol{h}^t \in \mathbb{R}^k$: hidden state vector
- $\boldsymbol{f}^t \in \mathbb{R}^k$: forget gate vector

- $\boldsymbol{i}^t \in \mathbb{R}^k$: input gate vector
- $\tilde{\boldsymbol{c}}^t \in \mathbb{R}^k$: candidate gate vector
- $\boldsymbol{c}^t \in \mathbb{R}^k$: cell state memory vector
- $\boldsymbol{o}^t \in \mathbb{R}^k$: output gate vector

where the superscripts n and k refer to the number of inputs and hidden units, respectively. The hidden state vector can also be called the output vector of the LSTM block.

The LSTM can add or remove information through structures called gates, illustrated in Fig. 3.4. In the first forget gate, LSTM removes irrelevant information from the previous state by the operation of (4.2). In the next step following (3.6)-(3.7), with the input and candidate gates, the LSTM stores the new relevant information into the cell state c^t. In (3.8), the cell state values get updated with the information from the previous two steps. The last one is the output gate, following (3.9), the gate controls the information that is to be passed to the next time step, together with information from the cell state by the process of (3.10). The LSTM blocks can create a chain-like structure with several blocks to finally create an LSTM network. An LSTM network constructed using a single layer is illustrated in Figure 3.5.

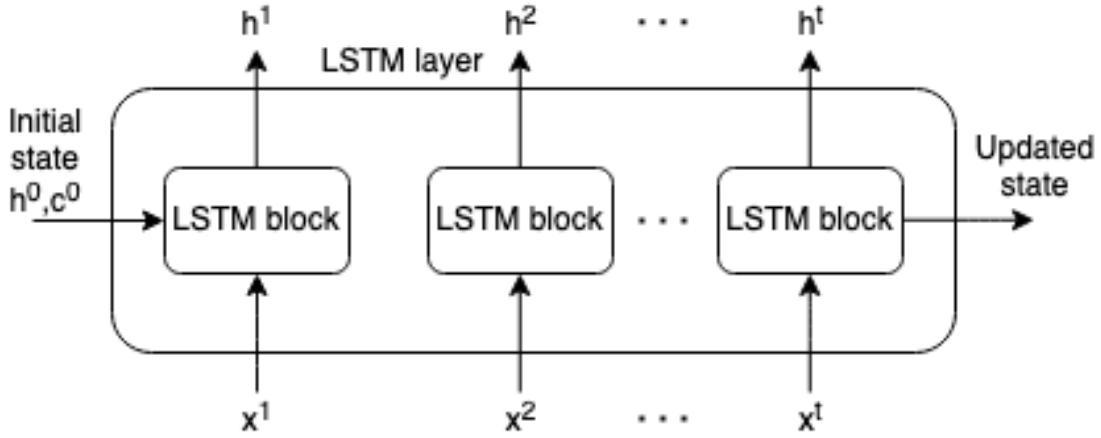


Figure 3.5: A layer of LSTM blocks.

4

Methodology

This chapter presents a method for developing an LSTM model for forecasting the FN during a disturbance. First, the generation of data and the test grid used for simulation will be explained. Second, the prediction model will be presented together with corresponding tests of the model.

4.1 Test grid

A modified version of the classical Nordic32 grid is chosen as the test grid, presented in [34]. The modified grid is similar to the the NPS, but fictitious. A single line diagram of the test system can be seen below in Figure 4.1.

The system comprises four regions, namely:

- "**North**" – Region consisting of a majority of hydro generation and small loads
- "**South**" – Region with thermal power generation, connected loosely to other areas in the system
- "**Central**" – Region with high loads and mostly thermal generation
- "**Equivalent**" – Region which is the equivalent of an external system that is connected to the "North."

The system has three different types of transmission lines operating at three different nominal voltages: 400 kV, 220 kV and 130 kV, respectively. The 400 kV ones are long transmission lines with series compensation, while the 130 and 220 kV lines represent the regional systems. Table 4.1 gives an overview of the active power generation and load in the test system. The system frequency is 50Hz, and frequency control is done by the hydro generators' speed governors in the North and Equivalent regions. An equivalent generator G20 represents generation from the "Equivalent" area. It contributes a significant amount towards primary frequency control, and the thermal generation units in the Southern and Central regions do not contribute at all.

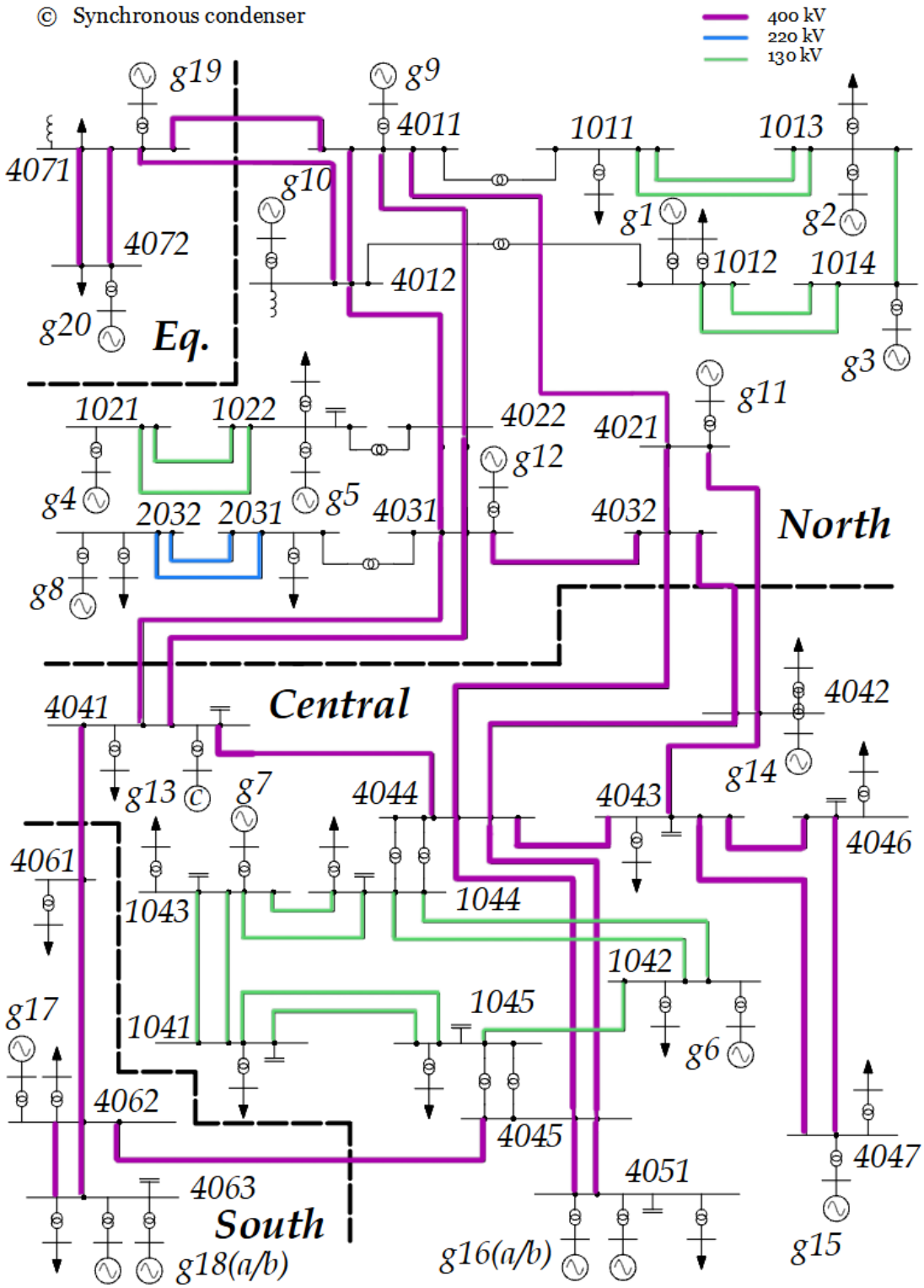


Figure 4.1: Nordic32 test grid used for developing the ML model [34].

Table 4.1: Active power generation and load for the Nordic32 test grid.

Area	Generated Power (MW)	Power Consumed by Load
North	4628.5	1180.0
South	1590.0	1390.0
Central	2850.0	6190.0
Equivalent	2437.4	2300.0
Total	**11505.9**	**11060.0**

4.1.1 Operating points

There are significant power transfers from the North to the Central regions, and also, the system is loaded heavily. This results in the system being more sensitive in these areas, where a disturbance in any transmission line has a more significant impact on the system stability than if it were to occur in other power transfer sections. The reactive power capability of the generators in the Central region and that of a few generators in the Northern region influence the maximum power that is delivered to the loads in the Central areas. Over-excitation limiters keep the reactive power levels in check, and the load tap changers restore voltages at the distribution levels, thereby the powers of the loads. In the test grid, there are a total of 74 buses, out of which 20 are generator terminal buses, 22 are distribution level buses, and 32 are transmission level buses. The grid also includes 102 branches, out of which 20 are step-up transformers, and 22 are for distribution purposes.

In the Nordic32 test grid, two operating points, namely, operating point A and B, are considered. The former operating point is said to be insecure due to the inability of the system to withstand the N-1 disturbance criteria. According to the guidelines given by the IEEE PES Task force on Test Systems for Voltage stability Analysis and Security Assessment, a few simple modifications are then made to the grid with operating point A by adding a few buses, generators, and transformers [34]. This is to make the system secure, and the new operating point is called operating point B, where the system is now able to withstand an N-1 contingency. The operating points give information about the active and reactive powers generated and consumed at the buses and the initial bus voltages. This data is attained by performing a power flow calculation in PSSE. Here, bus G20 is considered as the slack bus.

The following updates are performed to get the more stable operating point B: A generator G16B, which is identical to G16 and producing the same active power of 600 MW, which the slack bus compensates, is added in parallel to it with a similar step-up transformer. This addition decreases the power flow in the North-Central corridor by the same amount, increasing the system's robustness, but it cannot withstand a generator loss such as G15 or G18. One way to mitigate this is to reduce the contingency level by replacing G15 and G18 with two identical generators each, namely G15 & G15B and G18 & G18B, with half the original capacity. The connected step-up transformers are also replaced with similar ones that have half the initial rating.

4.2 Data Generation

The generation of data from thousands of simulations on large complex grids with numerous power system components needs to be automated. A common way to do this is to run PSS®E from the programming language Python. This can be done by utilising the FNSL Application Program Interface (API) provided by PSS®E. All the steps that show how the data generation is performed are presented in Figure 4.2 and explained in more details below:

Initialize generators and loads: The power generation is varied between 75% - 110% around the stable operation point B, explained in 4.1.1. The boundaries arises from stability issues when applying bigger variations. The change in power generation is distributed among the loads, where the change in each load is proportional to its size. The load flow of the initial operating condition is then solved using Full Newton-Raphson method. If the load flow does not converge, a new initial operating condition is generated.

Apply contingency: The dynamic simulations start with 2 seconds of simulation time, without any contingency, to get steady-state data prior to the disturbance. The applied contingency is always a generator trip in the test grid, except the equivalent generators G19 and G20. The severity of the fault is proportional to the amount of active power generated by the generator where the contingency is applied. It is vital to get close to uniform distribution of the fault magnitudes in the data set. If not, fault magnitudes that seldom occur will perform worse and vice-versa, resulting in a likelihood that the model will perform worse during the biggest fault magnitudes, when the stakes are at the highest. Since the maximum capacities of the generators in the test grid vary, bigger generators have to generate more samples at bigger fault magnitudes, to get a uniform distribution. The fault magnitudes are varied between 180 MW - 750 MW and the interval is split into two parts. Below 570 MW, a step size of 45 MW, and above this, 15 MW is applied. This results in the distribution being more uniform in between the smallest and biggest faults. It is good to note that the fault

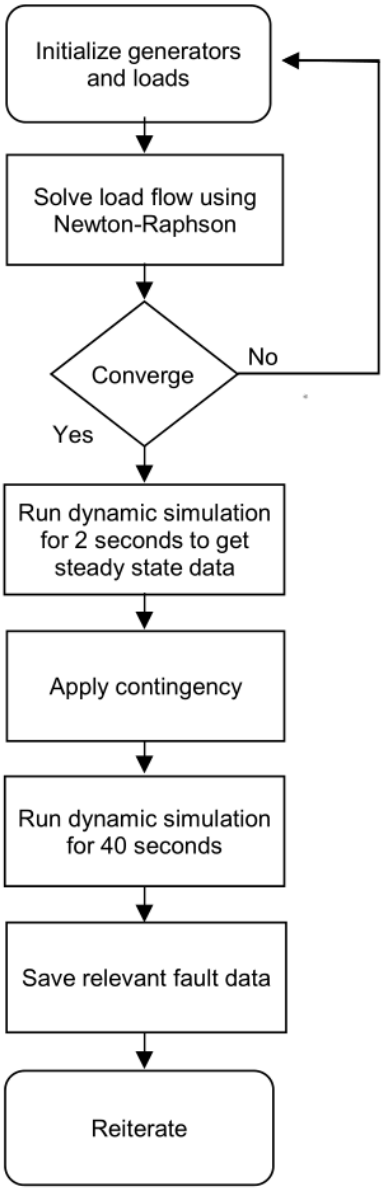


Figure 4.2: Flowchart of the data generation.

magnitudes of the generators are set before performing the load flow. Finally, after the application of the contingency, the dynamic simulation is run to observe the lowest point the frequency will reach. Based on the data collected from the PMUs in the Nordic region, the average time to nadir was found to be 8.7 seconds [1]. Thereby, the simulation is run for 40 seconds to ensure that this time-frame includes the FN.

Save relevant data: The data saved from the simulation should be used by the ML models, which require both input and output data. The input data saves a sequence of measurements starting from right before and up to a certain point after the disturbance, with parameters measured at multiple locations in the power system. More details about this are provided in the following section. The ML model should learn to map this input data to the output data, consisting of the measured FNs. Even though frequency is called a global parameter, it differs depending on the location in the grid during a disturbance. To get some representation of an average frequency, the center of inertia frequency is used:

$$f_{coi} = \frac{\sum (f_i \cdot H_i)}{\sum H_i} \tag{4.1}$$

where f_{coi} is the centre of inertia frequency, f_i is the frequency of an individual generator, and H_i is the inertia constant of a specific generator.

4.3 Frequency nadir prediction model

This section presents the proposed LSTM model's architecture and training procedure. Similarly, the NN baseline model is introduced. It ends up with an explanation about how the input data studies have been conducted.

4.3.1 LSTM model

The literature review showed that sequence data improved the forecasts of FN, even though the most common ML methods for sequence data have not been used. Consequently, an LSTM model was chosen, often used in sequence data tasks such as language translation. Moreover, the parameters that are available shortly after the disturbance can be sufficient for the forecasts. Since an earlier prediction can improve the frequency response, only these kinds of parameters were used.

4.3.1.1 LSTM architecture

The architecture of the LSTM network is presented in Figure 4.3. The model uses sequence data consisting of 55 time steps of measurements in the input vectors x^t to forecast the output y^t, in this case the FN. The architecture is based on one layer consisting of a sequence of 55 LSTM blocks, where each individual block consists of 25 LSTM cells. Each LSTM block takes one input vector x^t, follows the operation presented in Section 3.2.1 and the output forwards to the next block in the same layer. The output from the last block goes into a linear activation function to make the forecast of FN.

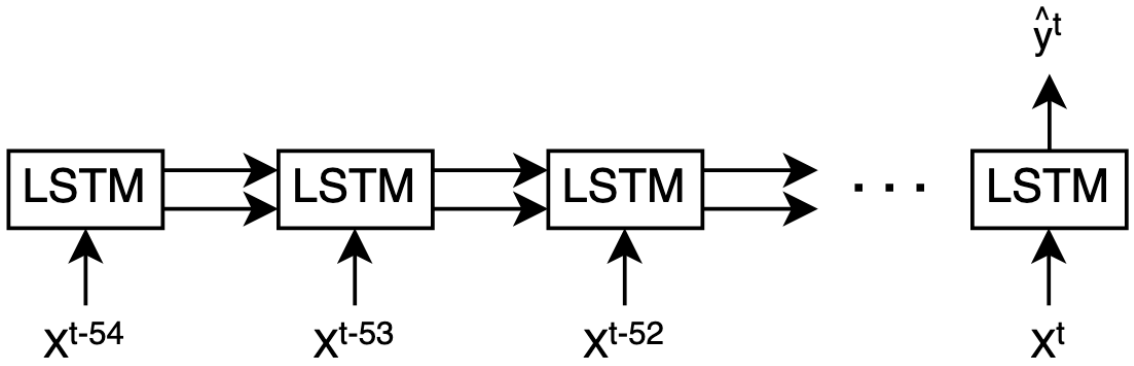


Figure 4.3: Architecture of the LSTM network.

4.3.1.2 LSTM training

The data from the generation explained in Section 4.2 were split into training, validation and testing. The training of an LSTM is based on supervised learning; hence, the model is trained on both input and output data. The input training data has the dimension ($3161 \times 55 \times 364$), representing the number of training samples, time points in the sequence, and input parameters, respectively. The time sequence consists of data from 0.1s before the disturbance until 1s after. Thus, the measurement frequency is 50 Hz. The input parameters consist of frequency, voltage and active power flow. The frequency and voltage are measured at each load, bus and generator. Meanwhile, the active power flow is measured at lines and from generators. The output training data is the target value of FN, given in center of inertia frequency. To get a successful training of the model, some pre-processing of the data is required. Firstly, the FN of the simulations is expected to be around 49.9-49.4 Hz. Due to instability issues in some simulations, a smaller part of the simulations ended up with an FN far below the given range. Therefore, these samples were removed since they are not realistic and will have a negative impact on the model. Secondly, the data is normalised between [-1,1] to improve the training as follows:

$$x_{scaled} = \frac{x - (\frac{max(x)+min(x)}{2})}{\frac{max(x)-min(x)}{2}} \tag{4.2}$$

where each parameter is scaled separately. The best performing LSTM model was found using data from the first 0.7 seconds of the time sequence, with a measurement frequency of 17 Hz and 25% of the parameters in the data set. The network was trained with an adaptable learning rate optimisation algorithm, known as "Adam" [35]. Through a method called mini-batch gradient descent, batches of 85 samples were used separately to train the model. A maximum of 500 epochs was used during training, where the training was not likely to hit the maximum epochs due to a method called early stopping. If the validation data did not improve after ten epochs, meaning the validation loss improved with less than 10^{-12}, the training stopped. The training was further improved by reducing the learning rate by a factor of 0.85 when the learning stagnated and the validation loss were less than 0.0001 over ten epochs. All parameters related to the network is summarised in Table 4.2.

Table 4.2: Architecture and parameters of the long short-term memory network.

Data	
Input parameters	364
Time steps in sequence	55
Input data type	f, V, P
Training data	3161
Validation data	400
Test data	400
Architecture	
LSTM layers	1
LSTM hidden cells	25
LSTM layer activation	Tanh
Final layer activation	Linear
Training	
Optimizer	Adam
Learning rate	0.01
Max epochs	500
Batch size	85
Loss metrics	MSE

The LSTM network presented above was implemented in Python, using the Keras library. The parameters and architecture of the model have been iteratively tuned to improve the prediction accuracy. It should be noted that the model's design is the best for the Nordic32 test grid and does not have to be for other grids. Moreover, it is clear that this process could be extended further and result in even better forecasts.

4.3.2 Baseline and testing

To examine if it is beneficial to use sequence data, the LSTM network will be compared with a NN using data from a single time point. The choice of NN as a baseline model comes from promising results in previous studies [4]. The architecture of the NN is depicted above in Figure 4.4. From iterative testing, the NN was found to perform best using data from the time point 0.6 seconds after the disturbance and 10% of the parameters in the data set. This creates an input vector that is fed into the NN. For a full explanation about how a NN operates, see Section 3.1. The input values and weights are summed up in the hidden neurons and go into the rectified linear activation function (ReLU). In the last output layer, the output from the hidden layer together with the weights are summed up and applied to a linear activation function.

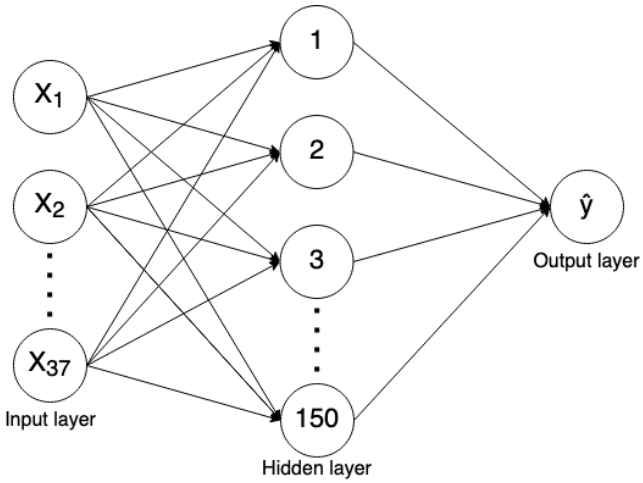


Figure 4.4: Architecture of the neural network used as a baseline.

The training was done similarly to the LSTM network, with other parameters presented in Table 4.3. These were also tuned in an iterative manner, with room for improvement of the prediction accuracy. The NN was implemented in Python using the Keras library.

Table 4.3: Architecture and parameters of the Neural Network.

Data	
Input parameters	37
Input data type	f, V, P
Training data	3161
Validation data	400
Test data	400
Architecture	
Hidden layers	1
Hidden cells	150
Hidden layer activation	ReLU
Final layer activation	Linear
Training	
Optimizer	Adam
Learning rate	0.001
Batch size	95
Max epochs	500
Loss metrics	MSE

4.3.3 Study of input data

Though the accuracy of the forecast model is of considerable interest, the time point of the prediction and the ease of implementation can also have a similar priority. This can be attributed to the data required for the model. To investigate this further, four sub-studies are conducted, where the input data is varied to answer the following questions:

1. **Frequency of measurements:** How does the reporting frequency of the PMUs impact the predictions?

2. **PMU coverage:** How many PMUs are needed in the system to achieve satisfactory performance?

3. **Input parameters:** Are all of the parameters in the data set required to use?

4. **Time point of prediction:** How fast can the model make accurate predictions?

In each study, only one parameter is varied at a time. For example, in the first sub-study considering frequency of measurements, only the measurement frequency is varied. The rest of the data structure follows the original data set explained in Section 4.3.1.2. For each measurement frequency level, the model is trained and tested with this set of data. Every time an LSTM model is trained, the network will be unique. Hence, when training the same model multiple times, there will be variations. To work around this, all of the results are an average of five times of training and testing.

The result metrics used in the report is Mean Squared Error (MSE) and Mean Absolute Error (MAE). MSE is important to highlight since that is the metric the training and testing of the model is based on, calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{4.3}$$

where $\hat{y}$ is the predicted value and y is the target value from the simulations. A more illustrative metrics to get the head around the accuracy of the model is MAE:

$$MAE = \frac{1}{n}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{4.4}$$

5

Results and discussions

In this chapter, the accuracy of the LSTM network and a comparison to the baseline NN model is presented and discussed. Finally, the required data for the model is introduced and analysed.

5.1 LSTM accuracy and baseline comparison

The prediction accuracy of the LSTM and baseline NN model is listed in Table 5.1, where the accuracy is expressed in both MAE and MSE, and the values are in hertz. Since training the model twice never ends up in two networks that are exactly the same, all the values in this section are an average of five times of training and testing with the same parameters. The results indicate that both models can predict fairly accurately, however, the LSTM network is clearly better with an MSE of $3.5 \cdot 10^{-7} Hz$ compared to $1.2 \cdot 10^{-6} Hz$ of the NN.

Table 5.1: Prediction accuracy of LSTM and NN baseline model.

	LSTM	**NN**
MAE	0.00041 Hz	0.00080 Hz
MSE	$3.5 \cdot 10^{-7}$ Hz	$1.2 \cdot 10^{-6}$ Hz

As the potential consequences of poor frequency stability control are severe, it is better to have a model that always predicts relatively well over one that is often superior but sometimes fails. Therefore, the largest errors can be more interesting than the average accuracy, which are illustrated as a histogram in Figure 5.1. In frequency nadir prediction, optimistic forecasts are inferior as they can lead to countermeasures that are too small. Therefore, the biggest positive error will presumably be the minimum additional margin for the countermeasures. Figure 5.2 depict the error distribution over the spectrum of frequency nadir target values for the LSTM and NN networks respectively. The graphs show that the errors are well distributed over the range for both models and the biggest positive errors are 0.0016 and 0.0041 Hz for LSTM and NN, respectively.

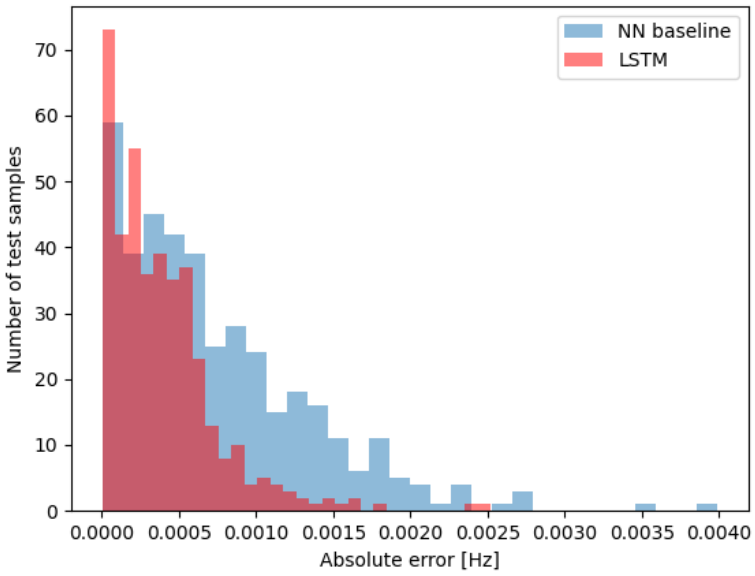


Figure 5.1: Histogram of prediction error of LSTM and baseline model.

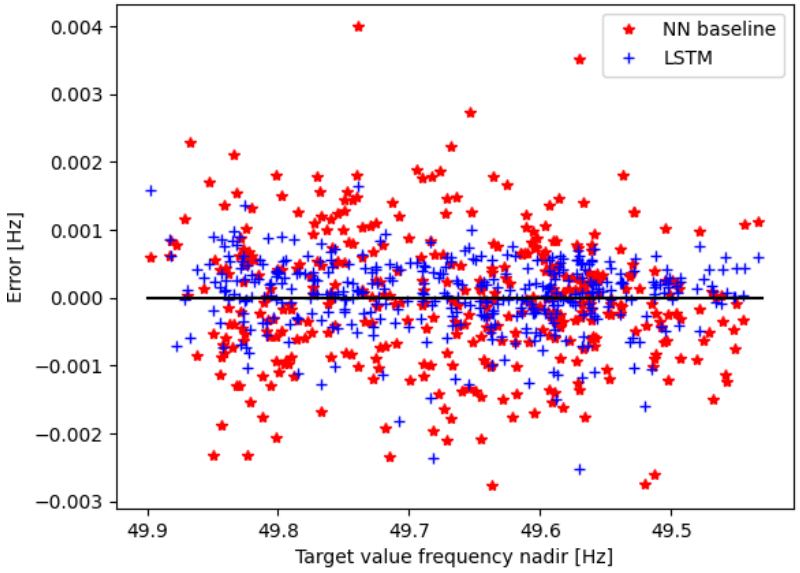


Figure 5.2: Error distribution of LSTM and baseline model.

It is fair to presume that the prediction task will be more difficult with a more realistic data set in a bigger power system. With the lack of historical disturbance data, it is hard to say how accurate the model will be in a real power system. A more realistic data set can be achieved by applying more randomised operational conditions with bigger variations and also by varying kinetic energy in the system. On the other hand, an LSTM network scales up quite good in terms of handling complexity and more data, which can be provided through simulations. Therefore,

an LSTM is a promising network for the task, given the data generation is of high quality.

Even though it is hard to evaluate the model on a real power system, it is essential to understand the potential additional errors when implementing the model online. Primarily, the concept of PMUs is to provide data with a magnitude error less than 0.1%. This aim is often not achieved mainly due to system imbalances or errors from instrumentation channels. Moreover, the evaluation of PMU accuracy is a challenging task [36]. Assuming the PMUs have an error percentage of 0.1%, roughly transferring to an frequency error of 0.05 Hz. It is also certain that the simulations will never be 100% accurate. Furthermore, the cases simulated during data generation will not cover all possible situations in the power system, resulting in additional errors.

5.2 Impact of input data

Models that utilise fast data from PMUs are, from the authors' knowledge, not widely used today. PMUs are an emerging technology expected to play a more significant role in future power systems. Therefore, it is good to specify what kind of data the model requires from the PMUs. There is also a trade-off between high accuracy and ease of implementation, where using fewer data can be preferable over tiny improvements in the predictions. For these reasons, a study of the impact of the data delivered from the PMUs will be presented below. It should be noted that the whole data set is used from the data generation, shown in Section 4.3.1.2, except the parameter studied in each graph that is varied. Further details about how these studies were conducted can be seen in Section 4.3.3.

The results are not aimed at finding the best parameters; instead, at detecting a good range, because if one parameter is changed, it will impact what is best to choose for the other parameters. As an example, the results show that the reporting frequency of 5 Hz and 25% PMU coverage give the best results. Meanwhile, the best combination of them both were found at 17 Hz and 25%. The impact from the reporting frequency of the PMUs is depicted in Figure 5.3. The graph shows that the model can accurately predict with a measurement frequency from 50 Hz down to 2 Hz. A PMU generally has a reporting rate between 30-60 Hz [37]. Hence, the sampling rate of the PMUs is not likely a limiting factor to implement the LSTM model.

An LSTM is a black-box model, meaning it is difficult to prove why it behaves in a certain way. A possible explanation could be that information about the following dynamics is found in the rate of change of the parameters, which can be detected similarly with 2 Hz and 50 Hz, as well as the operational condition before the disturbance. The change in the parameters can probably represent the relationship between the size of the disturbance, the inertia in the system and the frequency nadir. The loads in the system have a significant impact, where the same fault but with high and low load can differ up to 0.1 Hz. Therefore, information about

the system loading is expected to be found in the data prior to the disturbance. Moreover, a higher reporting frequency could result in more noisy data, making the forecasting task harder.

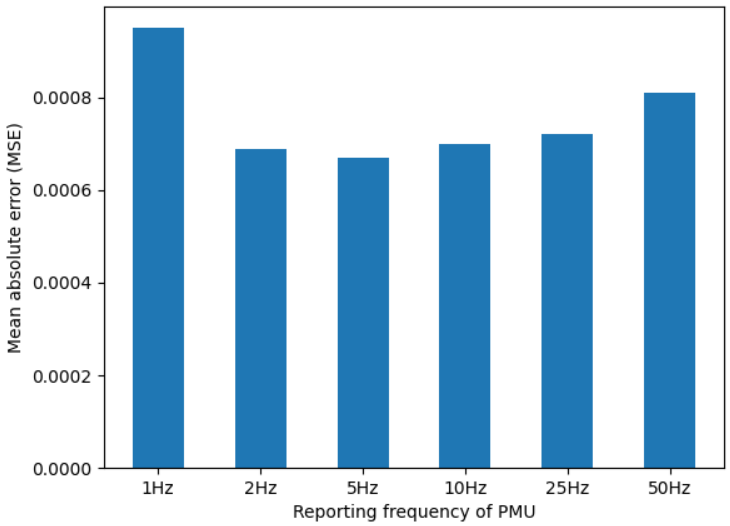


Figure 5.3: Frequency of measurements.

To implement the LSTM model online, a sufficient amount of PMUs need to be available. Figure 5.4 shows how different level of PMU coverage in the Nordic32 test system impacts the prediction accuracy. 100% coverage means that PMUs provide frequency and voltage data from all buses, generators and loads, as well as active power flow data from lines and generators.

As seen from Figure 5.4, the model can provide accurate predictions with down to 10% of PMU coverage. This number can be explained by parameters measured at different locations could possibly give similar information. Furthermore, why 100% PMU coverage is not the best can be related to that additional data can add noise and thereby worsen the predictions. It is not straightforward what the requirement of 10% PMU coverage transfers to in the real system. In the NPS, around 50% of the substations in the transmission network have PMUs, and it is growing [38]. Therefore, the prerequisites in terms of the number of PMUs could already be there but need further studies to confirm.

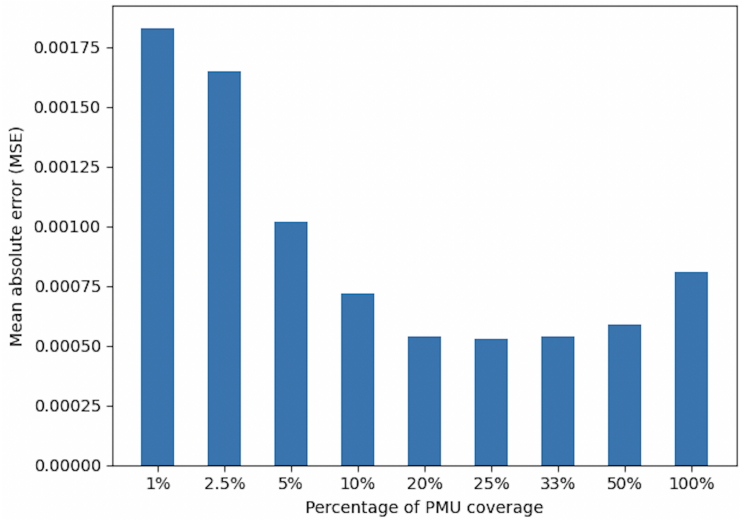


Figure 5.4: PMU coverage.

The data set includes frequency, voltage and active power flow data. Different combinations of these parameters were used as input data to evaluate if the model requires all these parameters to predict with high accuracy. The result of this study is presented in Figure 5.5. The model seems to benefit from using data of all parameters. In a scenario where either voltage or active power flow data are not available, the model can make reasonably good predictions when one of these is missing. Surprisingly, the frequency is the worst single parameter to use. One possible reason for that can be that the frequency can not catch the impact of the load in the system. That is something the voltage can do but is probably worse in detecting the severity of the fault. The fact that the loads in the system are not randomised can act in favour of the voltage, where it is possible the voltage levels prior to disturbance can give precise information about the load level in the system.

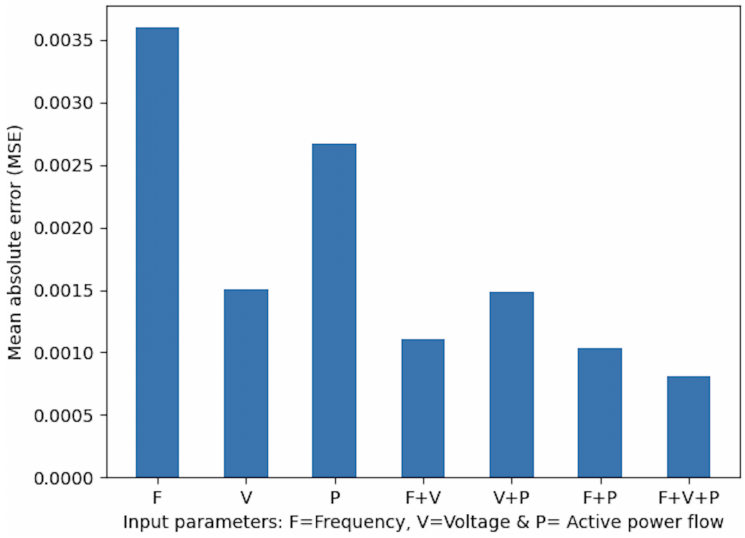


Figure 5.5: Input parameters.

For a frequency nadir forecasting model to be useful, it has to able to predict early to give the control services extra time to operate. To use the LSTM model online, accompanied by a method that can detect the disturbance. This can, for example, be done as in [14], where they continuously observed the RoCoF and a decline greater than 0.035 Hz/s indicated a disturbance. Ideally, with a 50 Hz PMU the method could detect the disturbance in around 0.05s. Even though the reporting rate of the PMUs would be 50 Hz, it does not mean they measure with this frequency. Instead, a part of the samples is calculated. Figure 5.6 clearly shows that the model can predict with no problem as soon as 0.1s after the disturbance if the data is available.

As discussed, it will probably be challenging to detect the fault and provide the data faster than 0.1s, even in an ideal case. Thereby, the time point of the prediction will most likely not be a limiting factor for implementing LSTMs in frequency nadir forecasting. As earlier mentioned, information about the following dynamics can probably be found in the rate of change of the parameters as well as in the operational conditions before the disturbance.

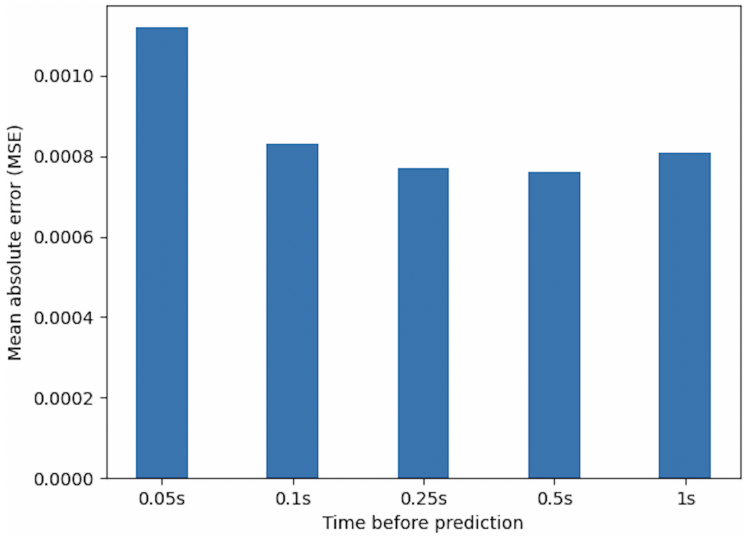


Figure 5.6: Time point of the prediction.

6

Conclusions and future work

In this chapter, a summary of the conclusions, based on the scope and research questions of the thesis, is presented. Furthermore, a discussion on potential future research is introduced.

6.1 Conclusions

The main aim of this thesis has been to develop an ML model that can forecast the FN after a disturbance and thereby support the frequency stability control in the power system. A new methodology that made use of sequence data to forecast the FN was proposed. Studies were further conducted on the developed model to address the required data from the PMUs.

The research question was to develop an ML model to forecast the FN. During the literature review, it was found that a NN benefited from short sequences of data instead of data from a single time point. Apart from that, no other studies were found which used sequence data and consequently, a common ML model for sequence data, LSTM, was chosen. Historical data from disturbances are scarce; hence, the large amount of data required for ML models were gathered from simulations in PSS®E 35.0. The biggest challenges in the project were related to the generation of a realistic data set. Major changes in the test system and large load variations often resulted in stability issues. Therefore, the data generation was a bit limited. The developed model was compared to a NN with data from a single time point. The results showed that the proposed model could predict the FN with high accuracy using simulated data and was significantly better than a NN using data from one time point. Hence, confirming that sequence data can be preferred in FN forecasts.

In the sub-question, the data required from the PMUs to accurately forecast the FN were studied. To the authors' knowledge, similar models that make use of PMU data are not widely used. Since PMUs are an emerging technology in power systems, it is interesting to study what kind of data the model benefits from. The LSTM network seemed to accurately predict even with measurements of low frequency; hence, the reporting frequency of the PMUs will likely not be a limiting factor. The input parameters from the data generation were frequency, voltage and active power flow since these are expected to be easy and fast available. Other studies have used, for

example, inertia or the size of the disturbance, which has a significant impact on the FN but often takes a few seconds to be obtained after a disturbance. The results indicated that the model benefits from data of all the parameters in the data set and also made adequate predictions when either voltage or active power flow were missing. Interestingly, the model could make accurate predictions as soon as 0.1s after the disturbance occurred.

6.2 Future work

The thesis has shown promising results for the proposed FN forecasting model, however there is plenty of research left prior to implementing it online. Future work can be carried out in the following topics:

- The developed model forecasts the FN during a disturbance. Future work should develop methods that automatically can act on this information. For instance, such methods could optimise the activation of frequency reserves based on the forecast.

- The proposed model is strongly dependent on a realistic simulated data set to work online. As mentioned earlier, this comes with some challenges in generating a representative training set. The LSTM model (and ML models in general) can scale up well and handle data that is significantly larger and more complex than what is shown in this thesis, making an extensive and representative data set even more interesting for further research.

- In future research, the developed models should also be tested on a real power system for evaluation and to find the limitations. The availability of data and the time taken to receive it is of considerable interest. Moreover, a sensitivity analysis of the model to measurement errors should be examined.

- The results shown in this thesis indicate that the PMUs in the power system only need to cover a smaller part. Since PMU coverage in the power systems is growing, it would be interesting to study how many PMUs are required and the optimal positioning of the PMUs for effective use in a real power system.

- Further research can also be done on the grid integration of renewables to study the impact of wind penetration and other forms of sustainable energies on the frequency stability and FN.

Bibliography

[1] M. Persson and P. Chen, "Frequency evaluation of the nordic power system using PMU measurements," *IET Generation, Transmission & Distribution*, vol. 11, June 2017.

[2] N. Modig, R. Eriksson, and M. Kuivaniemi, "Online Tool to Predict the Maximum Instantaneous Frequency Deviation during Incidents," in *IEEE Power and Energy Society General Meeting*, vol. 2018-Augus. IEEE Computer Society, dec 2018.

[3] E. Ørum, L. Haarla, M. Kuivaniemi, M. Laasonen, A. Jerkø, I. Stenkløv, F. Wik, K. Elkington, R. Eriksson, N. Modig, and P. Schavemaker, "Nordic Report: Future System Inertia 2," *ENTSOE, Brussels, Tech. Rep*, p. 153, 2018. [Online]. Available: www.entsoe.eu

[4] D. Zografos, T. Rabuzin, M. Ghandhari, and R. Eriksson, "Prediction of Frequency Nadir by Employing a Neural Network Approach," *Proceedings - 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT-Europe 2018*, pp. 1–6, 2018.

[5] Y. Xiao, R. Zhao, and Y. Wen, "Deep Learning for Predicting the Operation of Under-Frequency Load Shedding Systems," in *2019 IEEE PES Innovative Smart Grid Technologies Asia, ISGT 2019*. Institute of Electrical and Electronics Engineers Inc., may 2019, pp. 4142–4147.

[6] H. Li, C. Li, and Y. Liu, "Maximum frequency deviation assessment with clustering based on metric learning," *International Journal of Electrical Power and Energy Systems*, vol. 120, no. March, p. 105980, 2020. [Online]. Available: https://doi.org/10.1016/j.ijepes.2020.105980

[7] T. C. Njenda, M. E. Golshan, and H. H. Alhelou, "WAMS based under Frequency Load Shedding Considering Minimum Frequency Predicted and Extrapolated Disturbance Magnitude," in *Proceedings - 2018 Smart Grid Conference, SGC 2018*. Institute of Electrical and Electronics Engineers Inc., nov 2018.

[8] H. Hagmar, L. Tong, R. Eriksson, and L. A. Tuan, "Voltage instability prediction using a deep recurrent neural network," *IEEE Transactions on Power*

Systems, vol. 36, no. 1, pp. 17–27, 2021.

[9] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, 2004.

[10] P. Kundur *et al.*, *Power system stability and control*. McGraw-Hill, 1994.

[11] N. Hatziargyriou, J. V. Milanovic, C. Rahmann, V. Ajjarapu, C. Canizares, I. Erlich, D. Hill, I. Hiskens, I. Kamwa, B. Pal, P. Pourbeik, J. J. Sanchez-Gasca, A. M. Stankovic, T. Van Cutsem, V. Vittal, and C. Vournas, "Definition and classification of power system stability revisited & extended," *IEEE Transactions on Power Systems*, p. 3, 2020.

[12] F. Teng and G. Strbac, "Assessment of the Role and Value of Frequency Response Support from Wind Plants," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 586–595, 2016.

[13] H. Saadat *et al.*, *Power system analysis*. McGraw-hill, 1999, vol. 2.

[14] Erik Ørum, Mikko Kuivaniemi, Minna Laasonen, Alf Ivar Bruseth, Erik Alexander Jansson, Anders Danell, Katherine Elkington, Niklas Modig, "Nordic Report: Future System Inertia," *Entso-E*, pp. 1–58, 2018. [Online]. Available: https://www.entsoe.eu/Documents/Publications/SOC/Nordic/Nordic_report_Future_System_Inertia.pdf

[15] T. Kristiansen, "The nordic approach to market-based provision of ancillary services," *Energy Policy*, vol. 35, no. 7, pp. 3681–3700, 2007.

[16] Svenska Kraftnät, "Nätutvecklingsplan 2016–2025," *Svenska Kraftnät, Sundbyberg*, 2015.

[17] N. Modig, R. Eriksson, and Svenska Kraftnät, "FCR-D," Mar 2021. [Online]. Available: https://www.svk.se/aktorsportalen/systemdrift-elmarknad/information-om-stodtjanster/fcr-d/

[18] L. Saarinen, "The frequency of the frequency: on hydropower and grid frequency control," Ph.D. dissertation, Acta Universitatis Upsaliensis, 2017.

[19] Svenska Kraftnät, "Systemutvecklingsplan 2018-2027," *Stockholm, Sweden*, vol. 58, 2017.

[20] ENTSO-E, "Fast Frequency Reserve – Solution to the Nordic inertia challenge," no. December, pp. 1–22, 2019. [Online]. Available: https://www.svk.se/siteassets/aktorsportalen/elmarknad/information-om-reserver/ffr/ffr-stakeholder-report-20191213.pdf

[21] A. Akhil, G. Huff, A. B. Currier, B. C. Kaun, D. M. Rastler, S. B. Chen, A. L. Cotter, D. Bradshaw, and W. Gauntlett, "Electricity storage handbook," *Sandia Report United States Department of Energy*, 2013. [Online]. Available: https://www.sandia.gov/ess-ssl/lab_pubs/doeepri-electricity-storage-handbook/

[22] C. C. Aggarwal *et al.*, "Neural networks and deep learning," *Springer*, 2018.

[23] K.-L. Du and M. N. S. Swamy, *Neural Networks and Statistical Learning, Second Edition.* Springer, 2019. [Online]. Available: https://doi.org/10.1007/978-1-4471-7452-3_12

[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2016, http://www.deeplearningbook.org.

[25] A. Mohri, Mehryar; Rostamizadeh, Afshin; Talwalkar, *Foundations of Machine Learning*, 2nd ed. MIT Press, 2018, vol. 2. [Online]. Available: https://cs.nyu.edu/$\sim$mohri/mlbook/

[26] A. Gupta, G. Gurrala, and P. S. Sastry, "Instability prediction in power systems using recurrent neural networks," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 0, pp. 1795–1801, 2017.

[27] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, "Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network," *Applied Energy*, vol. 282, no. PA, p. 116177, 2021. [Online]. Available: https://doi.org/10.1016/j.apenergy.2020.116177

[28] X. Yin, Z. Jiang, and L. Pan, "Recurrent neural network based adaptive integral sliding mode power maximization control for wind power systems," *Renewable Energy*, vol. 145, pp. 1149–1157, 2020. [Online]. Available: https://doi.org/10.1016/j.renene.2018.12.098

[29] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "Recurrent neural networks for short-term load forecasting," *SpringerBriefs in Computer Science*, 2017. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-70338-1

[30] P. J. Werbos, "Backpropagation Through Time: What It Does and How to Do It," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[31] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[32] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[33] Q. Wu, K. Ding, and B. Huang, "Approach for fault prognosis using recurrent neural network," *Journal of Intelligent Manufacturing*, vol. 31, no. 7, pp. 1621–1633, 2020. [Online]. Available: https://doi.org/10.1007/s10845-018-1428-5

[34] L. Lima, T. Van Cutsem, M. Glavic, W. Rosehart, J. Santos, C. Cañizares, M. Kanatas, F. Milano, L. Papangelis, R. Ramos, B. Tamimi, G. Taranto, and C. Vournas, "Test systems for voltage stability analysis and security assessment," *IEEE*, 08 2015.

[35] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.

[36] A. Gomez-Exposito, A. Abur, P. Rousseaux, A. de la Villa Jaen, and C. Gomez-Quiles, "On the use of PMUs in power system state estimation," *Proceedings of the 17th PSCC*, 2011.

[37] L. Vanfretti, M. Baudette, and A. D. White, "Monitoring and control of renewable energy sources using synchronized phasor measurements," in *Renewable Energy Integration*. Elsevier, 2017, pp. 419–434.

[38] R. Eriksson, Private communication with Svenska Kraftnät, May 2021.