

Machine Learning Tool for Transmission Capacity Forecasting of Overhead Lines based on Distributed Weather Data

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der KIT-Fakultät für
Elektrotechnik und Informationstechnik,
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Ing. Gabriela Alejandra Molinar Torres

geb. in Caracas, Venezuela

Tag der mündlichen Prüfung:

19.11.2020

Hauptreferent:

Prof. Dr. rer. nat. Wilhelm Stork

Korreferent:

Prof. Dr. Christof Weinhardt

Zusammenfassung

Die Erhöhung des Anteils intermittierender erneuerbarer Energiequellen im elektrischen Energiesystem ist eine Herausforderung für die Netzbetreiber. Ein Beispiel ist die Zunahme der Nord-Süd Übertragung von Windenergie in Deutschland, die zu einer Erhöhung der Engpässe in den Freileitungen führt und sich direkt in den Stromkosten der Endverbraucher niederschlägt. Neben dem Ausbau neuer Freileitungen ist ein witterungsabhängiger Freileitungsbetrieb eine Lösung, um die aktuelle Auslastung des Systems zu verbessern. Aus der Analyse in einer Probeleitung in Deutschland wurde gezeigt, dass einen Zuwachs von einem Medianwert von 28% der Stromtragfähigkeit eine Reduzierung der Kosten für Engpassmaßnahmen um ca. 55% bedeuten kann. Dieser Vorteil kann nur vom Netzbetreiber wahrgenommen werden, wenn eine Belastbarkeitsprognose für die Stromerzeugungsplanung der konventionellen Kraftwerke zur Verfügung steht. Das in dieser Dissertation vorgestellte System prognostiziert die Belastbarkeit von Freileitungen für 48 Stunden mittels Maschinelles Lernens, mit einer Verbesserung der Prognosegenauigkeit im Vergleich zum Stand-der-Technik von 6,13% im Durchschnitt. Der Ansatz passt die meteorologischen Vorhersagen an die lokale Wettersituation entlang der Leitung an. Diese Anpassungen sind aufgrund von Veränderungen der Topographie entlang der Leitungstrasse und Windschatten der umliegenden Bäume notwendig, da durch die meteorologischen Modelle diese nicht beschrieben werden können. Außerdem ist das in dieser Dissertation entwickelte Modell in der Lage die Genauigkeitsabweichungen der Wettervorhersage zwischen Tag und Nacht abzugleichen, was vorteilhaft für die Strombelastbarkeitsprognose ist. Die Zuverlässigkeit und deswegen auch die Effizienz des Stromerzeugungsplans für die nächsten 48 Stunden wurde um 10% gegenüber dem Stand der Technik erhöht. Außerdem wurde in Rahmen dieser Arbeit ein Verfahren für die Positionierung der Wetterstationen entwickelt, um die wichtigsten Stellen entlang der Leitung abzudecken und gleichzeitig die Anzahl der Wetterstationen zu minimieren. Wird ein verteiltes Sensornetzwerk in ganz Deutschland umgesetzt, wird die Einsparung von Redispatchkosten

eine Kapitalrendite von ungefähr drei Jahren bedeuten. Die Durchführung einer transienten Analyse ist im entwickelten System ebenfalls möglich, um Engpassfälle für einige Minuten zu lösen, ohne die maximale Leitertemperatur zu erreichen. Dieses Dokument versucht, die Vorteile der Freileitungsmonitoringssysteme zu verdeutlichen und stellt eine Lösung zur Unterstützung eines flexiblen elektrischen Netzes vor, die für eine erfolgreiche Energiewende erforderlich ist.

Abstract

The increment of intermittent renewable energy sources in the electrical power system is a challenge for grid operators. One example is the increase in congested overhead lines in Germany, because of the increment of wind energy transport from north to south. This problem reflects directly into the electricity costs of the end-users. Besides the construction of new overhead lines, a weather-dependent system operation is a short-term solution to improve the current utilization of the system. The analysis of a sample line in Germany presents a median increment of about 28% of the transmission capacity, which can mean a reduction of congestion measures costs in around 55%. The system presented in this dissertation forecasts the transmission capacity of overhead lines for 48 hours using machine learning algorithms, giving transmission system operators the possibility to create ahead a more reliable power generation plan. This system improves the transmission capacity prediction with respect to the state-of-the-art reference model in 6.13% in average. The approach adjusts the meteorological forecasts to the local weather situation along the line. These adjustments are necessary due to changes in topography along the line route and wind-shadows from the surrounding trees, which cannot be described by the meteorological models. Moreover, the models developed in this dissertation are able to compensate the prediction accuracy deviations from day to night hours, which corresponds to an improvement in the current capacity forecast. Thus, an increment of 10% in the efficiency of the power transmission plan compared to the state-of-the-art has been achieved. Additionally, the positioning of the weather stations has been optimized to cover those spots that are more susceptible to experiencing the highest conductor temperatures of the line. The installation costs of a distributed sensor network covering Germany were compared to the reduction in redispatching costs, giving a return-of-investment of three years. A tool has been developed for the visualization of the weather measurements and the capacity predictions once the system is installed and running. It can also execute a transient analysis, to determine the possibility to free congestion cases for some minutes without

reaching the maximum conductor temperature. This document attempts to clarify the benefits of dynamic line rating forecasting systems. It presents a solution to support a flexible electrical grid, which is necessary for a successful energy transition plan.

Contents

Zusammenfassung	i
Abstract	iii
Acknowledgments	xv
1 Introduction	1
1.1 Research questions	3
1.2 Dissertation structure	6
2 The Electrical Power System	7
2.1 The traditional electrical network	7
2.2 The energy transition and its challenges	10
3 Weather-dependent Overhead Line Operation	15
3.1 Thermal rating of overhead lines	16
3.1.1 Joule heating	17
3.1.2 Magnetic heating	18
3.1.3 Solar heating	18
3.1.4 Radiative cooling	19
3.1.5 Convective cooling	19
3.1.6 Current-carrying capacity	20
3.1.7 Transient state	20
3.1.8 Influence of weather in the current-carrying capacity	22
3.2 State-of-the-art of overhead line monitoring systems	25
3.3 Current-carrying capacity forecasting	29
4 Artificial Intelligence in Power Systems	33

- 4.1 General concepts of machine learning 34
 - 4.1.1 Learning methods 35
 - 4.1.2 Over- and underfitting 36
 - 4.1.3 Hyperparameter optimization 36
- 4.2 Learning algorithms 37
 - 4.2.1 Artificial Neural Networks 37
 - 4.2.2 Recurrent Neural Networks 40
 - 4.2.3 Convolutional Neural Networks 43
 - 4.2.4 Quantile Regression Forests 45
- 4.3 Evaluation metrics 47
- 5 The PrognONetz System 49**
 - 5.1 Investment analysis of dynamic line rating in Germany 51
- 6 Case Study 57**
 - 6.1 NOAA INL Mesonet 58
 - 6.2 Numerical Weather Prediction database from NOAA 59
 - 6.3 Virtual overhead line 64
 - 6.4 The PrognONetz dataset 64
 - 6.5 Software development framework and data loader 69
- 7 The PrognONetz Machine Learning Models 71**
 - 7.1 Baseline model 73
 - 7.2 Models based on weather observations 74
 - 7.2.1 Feature engineering 75
 - 7.2.2 Quantile Regression Forests 78
 - 7.3 Models based on numerical weather predictions 82
 - 7.3.1 Feedforward Neural Networks 84
 - 7.3.2 Convolutional Neural Networks 85
 - 7.3.3 Comparison and discussion 88
 - 7.4 Ensemble and final solution 90
 - 7.4.1 Comparison to the baseline and discussion 93
 - 7.4.2 Prediction safety factor for system operations 94
- 8 Integration of PrognONetz into System Operations 99**
 - 8.1 Implementation of a distributed sensor network 99

8.2	Software tool for system operators	105
9	Conclusion and Outlook	109
A	Appendix	113
A.1	The Nusselt number in convective cooling of overhead lines .	113
A.1.1	Forced convection	113
A.1.2	Natural convection	115
A.2	Relationship between sag, tension and conductor temperature	116
B	Acronyms	121
	References	123
	Personal publications	133
	Journal articles	133
	Conference proceedings	133
	Invited talks	134
	Supervised Student Work	135

List of Figures

2.1	Structure of the traditional power system	8
2.2	Amount of congestion management in Germany from 2013 to 2018	13
3.1	Example of current-carrying capacity over time	16
3.2	Capacity derivative with respect to ambient temperature	23
3.3	Capacity derivative with respect to solar radiation	23
3.4	Capacity derivative with respect to wind speed and direction . .	24
3.5	Capacity derivative with respect to wind speed and direction . .	24
3.6	Capacity derivative with respect to ambient temperature for differ- ent wind speeds	25
3.7	CAT-1	26
3.8	EMO System from Micca	26
3.9	OTLM	26
3.10	Ampacimon	27
3.11	Astrose	27
3.12	LiDAR System from the Lindsey company	28
3.13	LineVision	28
4.1	Venn diagram for AI and its derivations	34
4.2	Example for a Feedforward Neural Network	38
4.3	Common activation functions	39
4.4	Example of a Recurrent Neural Network	41
4.5	Example of a 2D-convolution	43
4.6	Explanation of 3D convolution	44
4.7	Example of a binary tree	46
4.8	Example of a multidimensional binary tree	47
5.1	PrognoNetz System Diagram	49

5.2	Case study for the benefit analysis of OLM systems	52
5.3	Additional current capacity over SLR	53
5.4	DLR cases under SLR	54
6.1	Case study - selected weather stations from the INL Mesonet . .	59
6.2	MAE for seven days of surface temperature forecast	61
6.3	MAE for seven days of wind speed and direction forecast	62
6.4	MAE for 48 hours NWP in the INL-Mesonet region	63
6.5	NWP grid at the INL Mesonet Region and virtual overhead line	65
6.6	Example diagram of temporal alignment of the NWP dataset . .	67
7.1	Process for performance improvement of a machine learning fore- cast model	72
7.2	Accuracy plot of NWP-nearest neighbor model	74
7.3	Feature engineering standard model	76
7.4	Geostrophic Wind Approximation at the INL-Mesonet area . . .	78
7.5	QRF input-output structure	79
7.6	Accuracy plot from experiments at different amounts of past weather observations	80
7.7	MAPE for Quantile Regression Forests	81
7.8	Map of NWP grid sizes for evaluation with CNN	83
7.9	Accuracy comparison of FFNN model to the NWP-nearest neigh- bor approach	85
7.10	Comparison of CNN models trained for different NWP grid sizes and inclusion of an age matrix	87
7.11	Comparison of the prediction models based on NWP	89
7.12	Block diagram of the ensemble of the best performing models .	91
7.13	MAPE of model fusion compared to the accuracy of the original models	92
7.14	Statistics of final model compared to QRF and FNN at the 1st forecast hour	93
7.15	Statistics of final model compared to QRF and FNN at the 48th forecast hour	94
7.16	MAPE comparison of the ensemble model to the NWP-nearest neighbor approach	95

7.17	Statistics of final model compared to NWP at the 48th forecast hour	96
8.1	Distribution of current-carrying capacity minima along the overhead line	100
8.2	Example for the determination of the filtering threshold for the hotspot analysis	101
8.3	Clustering example of hotspots along the overhead line	102
8.4	Influence of threshold value on the number of hotspots clusters .	103
8.5	Example for mean current-carrying capacity error for different number of weather stations covering the overhead line	104
8.6	PrognoNetz software tool: view of the overhead line and its ampacity	105
8.7	PrognoNetz software tool: transient analysis of the conductor temperature	106
8.8	Example of a transient management measure	107
8.9	PrognoNetz software tool: transient management	108
A.1	Inclined span diagram for calculation of line length	117

List of Tables

5.1	Comparison of the German redispatching situation under SLR and DLR	56
7.1	Comparative table for feature engineering	76
7.2	Summary of key features of final QRF model	82
7.3	Hyperparameters of CNN models	88
7.4	Summary of key features of final FNN model	90
7.5	Comparison of the German redispatching situation under SLR and DLR with a prediction safety factor	96
A.1	Values of the coefficients B and n for the calculation of the Nusselt number	114
A.2	Values of the coefficients k_1 and k_2 for the calculation of the Nusselt number	115
A.3	Values of the coefficients A and m for the calculation of the Nusselt number	116

Acknowledgments

Everything began with very careful and motivating parents, which keep giving me energy today, even though they are around 8000 km away. My siblings are the ones who bring joy to my life. My best friend and husband has been there for the past nine years and still gives me courage to follow my dreams.

My tutor, Prof. Dr. rer. nat. Wilhelm Stork, has taught me how to work autonomously, one of the most important values I take with me from the PhD.

I was surrounded by very creative, intelligent, and motivated colleagues at KIT-ITIV, which have given color to the working place. Specially I have to thank the LAMA-Team, Simon Stock, Marco Stang, Daniel Grimm, Christopher Meier and Tim Hotfilter, for the development of one of the nicest projects I had worked on.

Christos Klamouris and Franz Wegh are the pioneers of the research project PrognNetz. Their constant advice was invaluable.

During my PhD I had the pleasure to work with 16 bachelor, 12 master students, and 4 assistants. Two of them, Na Shen and Lintao Toni Fan, are now the main players of our PrognNetz-team. Amazing and hard-working people.

I was financially supported by the Nagelschneider Stiftung during my first three years of PhD. My codes were mostly executed on the supercomputer ForHLR II, at the Steinbuch Centre for Computing at KIT, funded by the Ministry of Science, Research and Arts from Baden-Württemberg and by the Federal Ministry of Education and Research.

The research project PrognNetz is supported by the German Federal Ministry of Economic Affairs and Energy (BMWi) and is a cooperation between the KIT and the companies UBIMET GmbH, Wilmers Messtechnik GmbH, unilab AG, Transnet BW GmbH, and GWU-Umwelttechnik GmbH. For the excellent cooperative work, I am grateful.

1 Introduction

Germany has the goal to achieve 60% of the final gross consumption of energy and 80% of the gross electricity consumption made up by renewable sources by 2050. A nuclear phase-out is also planned by 2022 because of the potential risks of nuclear power plants and the uncertain long-term handling of their residues [1]. Besides the benefits that this has against global climate change, it also represents a challenge for electrical network operators.

The increasing number of wind parks in the north of Germany and the nuclear power phase-out have unbalanced the spatial distribution of the generation centers [2]. The transmission network is prepared for a traditional structure, where the power plants are near to the consumption areas. Therefore, the amount of congestion (episode at which an electrical equipment cannot transport the power required to fulfill the load) has been increasing in the past years, as reported by the German Federal Network Agency [3].

The Transmission System Operators (TSOs) have a sequence of steps to follow, when congestion occurs. First, they are allowed to do redispatching, i.e., change the generation schedule of the conventional power plants. Thus, reducing the power flow in some lines and increasing it in others. If the congestion could not be solved using this method, then they can apply feed-in management measures, i.e., curtailment of Intermittent Renewable Energy Sources (IRES's) [4]. The application of these measures has been increasing from the beginning of the energy transition plan, called *Energiewende* in German. In 2018 5.4 TWh of IRES's were curtailed (9.7 times more compared to 2013), which meant a cost of 635 Million Euros [5].

Therefore, a power grid expansion is necessary to achieve the climate and energy targets. Plans to build new overhead lines or underground cables have been discussed at least from a decade ago, and only some of them have been approved. One of the main reasons for this slow process is the lack of the population's acceptance of this infrastructure [6]. The German Federal

Network Agency states the NOVA-Pinciple to cope with the slow growth of the electrical network. It prioritizes short-term alternatives, as optimization measures, over the construction and reinforcement of transmission lines [7]. Network optimization consists of using the maximum transmission capacities of the overhead lines, instead of the conservative limits considered nowadays. Overhead Line Monitoring Systems (OLMS's) are a grid optimization method, which retrieve the conductor's maximum power transmission capacity, also called Dynamic Line Rating (DLR) [8].

The DLR is determined by the maximum permissible conductor temperature and the minimum distance allowed from conductor to ground [9]. Consequently, the current-carrying capacity depends on the weather conditions, since high wind speeds, low ambient temperatures, and no solar radiation can reduce the conductor temperature, and vice versa. When monitoring systems are not available, the maximum transmission capacities are not known. Instead, they are approximated as a conservative limit (the so-called static line rating, described in detail in Section 3), which is determined for the worst-case weather conditions: high ambient temperature, full solar radiation, and low wind speed. This maximum limit is fixed independently of the actual weather scenario, which means that the electrical network is most of the time operating under sub-optimum levels.

Most of the DLR systems available in the market provide real-time current-carrying capacities. However, many decisions in system operations are taken one to two days ahead. The calculations of the operational limits, as the capacity allocations (Net Transfer Capacity) for cross-border energy markets, are carried out two days in advance. Network security calculations are executed one day ahead after the electricity market trade is closed. Therefore, real-time DLR is difficult to exploit by the TSO, while DLR predictions are necessary. That is the main reason why OLMS's have not been completely adopted, although the technology exists since the 1950s [10].

The straightforward solution to calculate the forecast for the current-carrying capacity corresponds to the direct utilization of the existing weather prediction models. Weather predictions for two days ahead are found with a spatial resolution of down to 1 km [11]. Nevertheless, irregular topographies, as mountainous places, can show dramatic changes in the weather conditions even at scales below that grid size. Moreover, forest routes, as those followed by overhead lines, produce wind turbulence, which is challenging to be de-

scribed by atmospheric models [10]. The improvement in the spatial resolution of the numerical solution of the weather prediction models is a trade-off for its forecast scope. The limitation is the total calculation time required to solve the set of differential equations numerically [12]. Downscaling is a meteorological method, which interpolates the weather models solution horizontally using topographical features of the area. The method uses distributed measurements of the weather conditions as calibration points, to improve the spatial resolution of the weather predictions [13]. However, a vertical interpolation to the overhead line-height, as expected for OLMS's, is still a challenge nowadays.

This dissertation presents a solution for the problems of the state-of-the-art, as a tool for TSOs with four main characteristics. First, weather observations are collected along the overhead line routes at conductor height. Thus, aiming to a description of the weather conditions along the line as accurate as possible. Second, the DLR is calculated for real-time monitoring, but also it is predicted for 48 hours. The DLR forecast is based on an adjustment of the meteorological predictions to the weather conditions along the line, using machine learning methods because of the possibility to optimize directly the DLR prediction accuracy. Third, a pre-installation procedure for the location of the weather stations along the overhead lines has been developed, which reduces the investment costs while still covering the conductor hotspots. Fourth, a thermal transient analysis, which tells TSOs the time, if any, they have to transmit more than the current-carrying capacity without overheating the conductor (normally around several minutes), thus coping with short-term congestion scenarios.

The solution has been developed in the framework of the research project PrognNetz, a cooperation between the Karlsruhe Institute of Technology, four companies and a TSO. The project is financed by the German Federal Ministry of Economic Affairs and Energy, for the period 2019 to 2021.

1.1 Research questions

Redispatching and feed-in management measures have nowadays a high economic impact in Germany [14]. Their use has increased over the past years due to the growth of congestion cases in the electrical network. This situation is related to the increasing amount of IRES's, but a slow grid expansion. The initial

research question of this dissertation studies the economic influence of DLR as a grid optimization technique on the congestion problem in Germany [8].

Research Question 1 *What is the potential of DLR to reduce congestion in the power grid and its related costs in Germany?* (Section 5.1)

The economic impact of DLR in operations has been analyzed on a sample overhead line in Germany. First, the increase in transmission capacity due to the availability of monitoring systems is estimated. Then, this extra-capacity is translated into the amount of power that could have been transmitted, instead of redispatched or curtailed. Finally, the potential savings in congestion management were estimated, which confirm the benefit on DLR investment.

Despite of their economic importance, existing OLMS's have faced a difficult integration and non-extensive use in grid operations [10]. The examination of the state-of-the-art (Section 3.2) shows that most of those systems focus on real-time measurements, which are hard to exploit when the decisions in systems operations are taken some days in advance. The systems for current-carrying capacity forecasting available in the market (Section 3.3) are nowadays a direct translation from weather predictions, which do not consider the influence of the topographical characteristics surrounding the line routes. This dissertation explores the improvement of the current-carrying capacity predictions when machine learning methods are trained using the local meteorological conditions along the overhead lines.

Research Question 2 *How do machine learning models improve the current-carrying capacity forecasting accuracy based on distributed weather measurements?* (Chapter 7)

The first point to address before answering this question is the definition of an accurate current-carrying capacity prediction. It must be operationally reliable enough for transmission system operators. N. Doban and Prof. Nordström [15] work in their research work with the Swedish operator Vattenfall with a maximum limit for the mean absolute percentage error of the current-carrying capacity prediction of 20%. This value was considered as the standard metric during the development of this dissertation. Moreover, the project partner TSO, Transnet BW, considers the forecast of the first ten hours the most important for accuracy improvement. The reason is that a change in the generation schedule of traditional power plants requires between six to eight hours to execute.

To answer the second research question of this dissertation, an exploration of different machine learning algorithms and their input feature set was performed. During this process, the dataset was divided into weather observations and predictions, and a separate analysis was carried out (Sections 7.2 and 7.3, respectively). This distinction helped to determine the contribution of each dataset, thus optimizing the information extraction by their later combination into the final model (Section 7.4).

The resulting system makes use of data from a distributed meteorological network. An installation of a weather station at every electrical tower is ideal but expensive¹. Therefore, the third research question studies how to reduce the installation costs of the system, while maintaining the accuracy of the current-carrying capacity calculation.

Research Question 3 *How can the installation costs of the system be reduced, while maintaining the accuracy of the current-carrying capacity?* (Chapter 8)

The installation costs are directly proportional to the number of weather stations in the system. The amount of measuring points influences the current-carrying capacity calculation accuracy. Moreover, the position of the weather stations also influences the results, since measuring at the line sections, where the probability of high conductor temperatures is higher, gives more information than measuring in other places. This dissertation presents a method for weather station positioning along the overhead lines, which considers these aspects.

¹ The installation procedure takes around a day per weather station. A line has around 3 towers per kilometer, which means around 150 towers for a normal line length of 50 km.

1.2 Dissertation structure

The structure of this dissertation is oriented along its contribution and is divided into three parts:

- First, the foundations are laid to provide context and an extensive understanding of the solution development process. In Chapter 2 the electrical power system is defined, its traditional structure is described, and the challenges of the integration of renewable energy sources are discussed. The weather-dependent overhead line operation is presented in general terms in Chapter 3. It describes the fundamental equations for the calculation of the current-carrying capacity depending on the weather conditions and discusses the current situation of the state-of-the-art. Afterward, Chapter 4 explains the machine learning algorithms used in the final solution presented in this dissertation.
- Chapter 5 presents the research project PrognoNetz as the framework for the development of this dissertation. The benefits of the system are motivated by an economical analysis of operating the electrical network using dynamic line rating. The study is based on a sample overhead line in Germany and an abstraction of the German electrical grid.
- Chapters 6 and 7 explain the development process of the current-carrying capacity forecasting models. First, the case study is described. Afterward, the machine learning model performance is evaluated.
- Finally, the integration of the solution into system operations is described in Chapter 8, giving emphasis to the positioning of the weather stations along the overhead line. Moreover, a tool for short-term congestion avoidance using a transient analysis is also presented.

A final summary, a discussion about the achievements of this dissertation and the further steps to make to achieve a constant growth of systems supporting the energy transition plan in the years to come, is presented in Chapter 9.

The external literature is marked with increasing numbers until [114]. The personal publications begin at [115]. The supervised student work will be referenced with the numbers starting on [126].

2 The Electrical Power System

The goal of this chapter is to understand the differences between the traditional electrical power system and the future environmental-friendly smart grid. The challenges of including IRES are outlined and analyzed. This analysis lays a solid basis to explain in Chapter 3, how OLM and ampacity forecasting systems are essential tools to succeed in the implementation of the network of the future.

In the following sections, the electrical network is theoretically defined, then the German energy transition plan is explained in detail, and finally, the congestion management measures (redispatching and feed-in management) are reviewed.

2.1 The traditional electrical network

Power systems are traditionally organized in a vertical structure, from generation to consumers, going through a transmission and distribution electrical network, as shown in Figure 2.1. Each of these subsystems work at a different nominal voltage level. In order to reduce losses by conductor heating, the voltage is increased at the transmission network, which works as "the electricity freeways" [16], allowing to move bulk electrical energy from production centers to consumption areas, usually over hundreds of kilometers. Distribution networks are "the highways" [16] that move electricity to the consumers. They are generally meshed circuits, but they are usually operated radially for protection reasons. Substations connect the transmission and distribution subsystems, where the transformers, protections, and metering systems are located [16].

The increment of small generating facilities, as solar- and wind-based power units, located in distribution networks, has changed the usual operation of the electrical system. From a single direction of power transmission, as explained

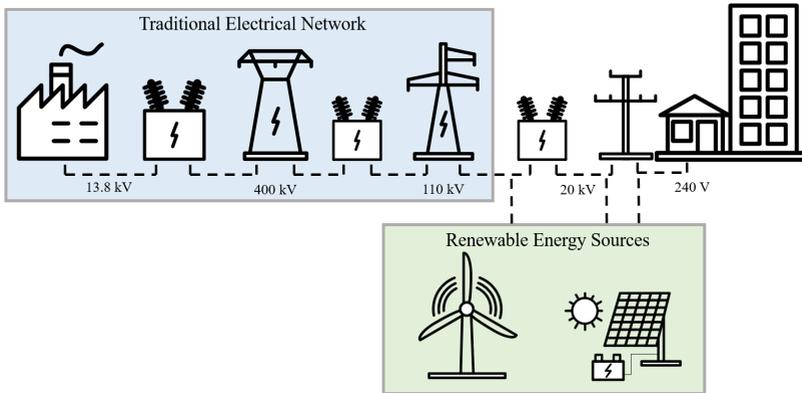


Figure 2.1: Structure of the traditional power system: Generation, transmission, distribution and supply.

above, now distribution networks are net electricity producers during some hours per day. It means that now the power can flow in reverse, from distribution to the transmission network.

Power systems are generally based on three-phase Alternating Current (AC) circuits. They work at a fixed frequency (in Europe, 50 Hz) and at different voltages, ranging from 10 kV to 1150 kV [17], [18]. In principle, AC systems have usually no direct regulating mechanisms. The network is governed by Kirchhoff Voltage Law, which means, that the power flows along transmission lines are distributed according to the system impedance.

In 1988 the concept of Flexible AC Transmission System Controllers (FACTS) was introduced, which constituted new mechanisms to control the power flow on the electrical grid [19]. An example for these systems are the Phase Shifting Transformers, which adjust externally the phase angle, thus controlling the maximum power flow of a line, P_{max} . In general, FACTS enhance power transfer capability, increase the controllability of the system by using power electronics, which therefore allow a flexible grid.

The calculation of the power through a single AC line can be described by Equation (2.1). It considers a lossless line of series reactance, X , connecting two nodes with voltages $V_i \angle \theta_i$ and $V_k \angle \theta_k$ [19].

$$P = \frac{V_i V_k}{X} \sin(\theta_i - \theta_k) \quad (2.1)$$

The maximum power flow of a line, P_{max} , depends on the angle difference in Equation 2.1, which is normally called $\delta_{max} = \theta_i - \theta_k$. The current-carrying capacity for short lines is determined by the thermal limit. However, as the line length increases, X grows linearly, and P_{max} reduces following $\frac{1}{X}$. That means, that there is a maximum length, at which the thermal limit does not determine anymore P_{max} . This phenomenon is called voltage drop. Kundur in his book "Power system stability and control" [20] generalizes this, stating that line lengths under 80 km remain under the thermal limit, lines between 80 and 320 km are governed by the voltage drop limit, and the longer lines present stability limits.

Equation 2.1 can be generalized for an interconnected grid with L lines and N nodes. It can also be linearized assuming that the angle difference $(\theta_i - \theta_k)$ is small. Then, Equation (2.2) is obtained, considering that the voltage amplitude is equal for all nodes in the grid, where B_{ik} is the per line susceptance diagonal matrix, equivalent to X_{ik}^{-1} . The subindices ik represent the line from node i to node k [21].

$$P_{ik} = B_{ik}(\theta_i - \theta_k) \quad (2.2)$$

The angle difference vector can be written in terms of the incidence matrix, A , which is a $L \times N$ -matrix and describes the topology of the network, as can be seen in Equation (2.3).

$$P_{ik} = B_{ik} A \theta_i \quad (2.3)$$

The translation of the line flows, P_{ik} , into the nodal active power balance, P_N , is shown in Equation (2.4), which describes the power injections and demand at the node.

$$P_N = A^T B_{ik} A \theta_i \quad (2.4)$$

After inserting Equation (2.4) in (2.3), the DC power flow equation is obtained (Equation (2.5)), where the term multiplying P_N is called the matrix of power transfer distribution factors, $D = B_{ik}A(A^T B_{ik}A)^{-1}$. It describes the relationship between the nodal power injections, P_N , and the active power flows through the transmission lines, P_{ik} [21].

$$P_{ik} = B_{ik}A(A^T B_{ik}A)^{-1}P_N \quad (2.5)$$

To solve Equation (2.4) one node has to be the reference, which means it has to be removed from the matrices and its voltage angle set to zero. The solution corresponds to the power flow DC and has to meet the following conditions [21]:

- The sum of all nodal injections equals zero, $\sum_N P_N = 0$.
- The active power flows per line are limited by their transmission current-carrying capacities, $P_{ik} \leq P_{ik,max}$.
- There is no negative generation and it cannot exceed the maximum generation capacity per unit.
- The demand at each node has to be completely covered.
- The dispatching solution follows the order of merit, with the renewable sources with the highest priority.

The power flow calculations are used to plan the generation of power plants and for the calculation of redispatching, depending of the expected demand, generation and transmission capacities.

2.2 The energy transition and its challenges

Carbon dioxide is a minor but critical component of the atmosphere. It is released through natural processes such as respiration, volcano eruptions, and due to human activities such as deforestation, land-use changes, and burning fossil fuels. From the Climate Change report 2014 [22], "about half of cumulative anthropogenic CO₂ emissions between 1750 and 2010 have occurred in

the last 40 years. In 1970, cumulative CO₂ emissions from fossil fuel combustion, cement production and flaring since 1750 were 420±35 GtCO₂; in 2010, that cumulative total had tripled to 1300±110 GtCO₂. Cumulative CO₂ emissions from Forestry and Other Land Use (FOLU) since 1750 increased from 490±180 GtCO₂ in 1970 to 680±300 GtCO₂ in 2010."

The Paris Agreement, signed by 181 countries in December 2015, is the first global climate change agreement. The main goal corresponds to keep the increase of ambient average temperature below 2 °C over pre-industrial levels. The year 2015 was the first time that at least 1 °C over the pre-industrial period was achieved¹.

As a measure to work against the greenhouse effect, several countries have implemented the so-called "energy transition plan". The International Renewable Energy Agency (IRENA) defines the energy transition as the "pathway toward transformation of the global energy sector from fossil-based to zero-carbon by the second half of the 21st century". IRENA states that renewable energy and energy efficiency measures can potentially achieve 90% of the required carbon reductions [24].

Germany has set the goal to achieve 60% of the final gross consumption of energy and 80% of the gross electricity consumption made up by renewable sources by 2050. A nuclear phase-out is also planned by 2022 because of the potential risks of nuclear power plants and the uncertain long-term handling of their residues [1].

The German energy transition, called *Energiewende* in German, also aims to increase energy efficiency. The goal is a 20% reduction in primary energy consumption by 2020, and a 50% reduction by 2050, compared to 2008. Power-heat coupling is one important improvement for energy efficiency, where heat, as a byproduct of electricity production, is used for buildings acclimatization or industry processes. All these measures should lead to a 55% reduction of greenhouse gas emissions by 2030 (from 1990 levels) [25].

¹ In [23] the pre-industrial period is defined. They suggest 1720-1800 as a suitable period in physical terms, since the major natural phenomena that also affect the climate (solar and volcanic activity) are similar to the levels of today. However, there are not enough observations. Therefore, the 1850-1900 period has been considered as a baseline to calculate the temperature anomalies of today, although it had several large volcanic eruptions and the greenhouse effect is considered to have started already.

The integration of more IRES corresponds to a challenge as well. It makes the system more complex by reducing the controllability of the power generation (neither the wind nor the sun can be controlled). This forces a review of the methods and processes of the former power system operation and planning.

In order to achieve an electrical system operating exclusively on IRES, there is a need for long term storage, to balance the load. The most used technology is the pumped hydroelectric storage because of its costs and efficiency. The system changes electricity into gravitational potential energy, i.e., water is pumped from a lower to a higher level. When energy has to be produced, the water is released, producing electricity by spinning the turbines [26].

Other energy storage mechanisms are, for example, power-to-gas, which converts electrical energy to a gas fuel (as hydrogen), or batteries (Lithium-ion, lead-acid, solid-state). According to the Electric Power Research Institute in the US, the installed cost for pumped hydroelectric storage varies between \$1700 and \$5100/kW, compared to \$2500/kW to 3900/kW for lithium-ion batteries. Pumped-storage hydropower is more than 80% energy efficient through a full cycle. However, the challenge of pumped hydroelectric is, first, that they are long-term projects needing 3 to 5 years of construction time, and second, that it is not always topographically possible to create a reservoir of water big enough to cover the storage requirements [26].

Another problem of the increment of IRES in the electrical system is the uneven spatial distribution of high wind speeds and solar radiation. This situation is experienced nowadays in Germany, where the wind power capacity is mostly located in northern Germany but the consumption centers are situated in the south [2]. The transmission network is prepared for a traditional structure, where the power plants are near to the consumption areas. Therefore, the congestion management measures² have increased around three times compared to 2014, as shown in Figure 2.2 [8]. The Germany's Electric Future Report, part 2, from the WWF [2] studies energy transition scenarios until 2050. It states that the mean value of all overloads is around 150% over the current transmission capacities. A large portion of those grid overloads may occur from the increase of offshore wind power.

² Congestion in power systems is defined as the episode at which an electrical equipment (as a transformer or an overhead line) cannot transport the power required to fulfill the load [8].

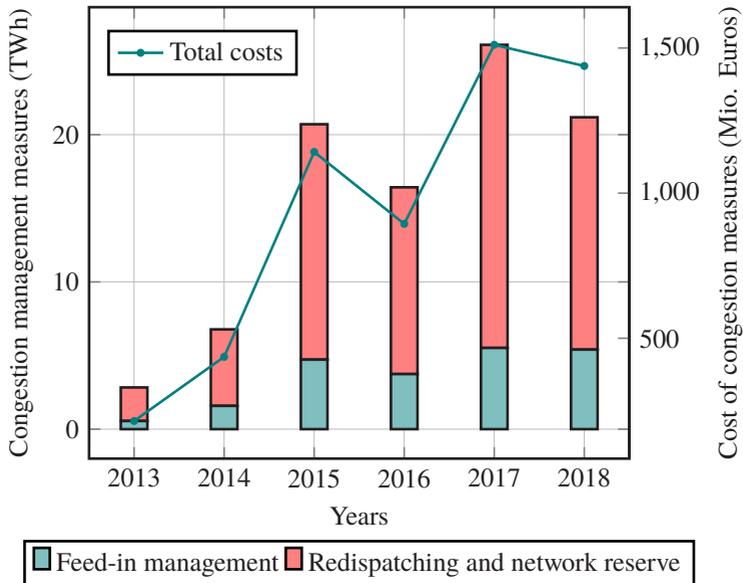


Figure 2.2: Amount of congestion management measures taken in Germany from 2013 to 2018 and their monetary impact for the country (data from [4], [5] and [14]).

Based on the German Energy Industry Act (Energiewirtschaftsgesetz, EnWG), if a TSO realizes a congestion on part of its infrastructure, which very often is on an overhead line, it has a sequence of steps to follow in a precise order. First, they are allowed to do redispatching or disconnect loads. As second step, if the congestion is not solved, they can bring conventional power plants to the minimum. At last, they are allowed to curtail IRES [4].

The process, at which power plants are instructed to increase or decrease their generation with respect to their original plan, is called redispatching. If the plant is a renewable source, then it is called feed-in management. The fact that the regulation of the latter is left as the last measure can also be seen in Figure 2.2, since the total feed-in management in TWh is always less than the yearly redispatching level [4].

Redispatching in Germany works as a costs-based mechanism, i.e., the redispatched power plants receive only their operating costs. This does not motivate

investors to optimize the grid, expanding local generation, or building more storage capacity [27].

Today, the incentive to avoid congestion in Germany relies on the NOVA Principle, a measure from the German Grid Development Plan. It gives priority to grid optimization, then to grid reinforcement before any further grid expansion (in German: *NetzOptimierung, vor -Verstärkung, vor -Ausbau*) [7]. Network optimization consists of using the maximum transmission capacities of the overhead lines, instead of the conservative limits considered nowadays. Overhead Line Monitoring Systems (OLMS's) are a grid optimization method, which retrieve the conductors maximum power transmission capacity, also called Dynamic Line Rating (DLR) [8].

The transmission capacity of overhead lines, for example, is limited thermally or by stability criteria, as explained in Section 2.1. System operators avoid reaching the limits for safety reasons. A long exposure of the conductor to its maximum temperature causes annealing, which does not allow the material to return to its original elongation at ambient temperature (usually 20 °C). An elongated conductor can be dangerous to objects, people or animals below the overhead line, when the safety distance cannot be maintained.

3 Weather-dependent Overhead Line Operation

The maximum current-carrying capacity of a transmission line is defined by many utilities as a static value, the so-called, Static Line Rating (SLR). It is calculated as a conservative limit to avoid, under severe weather conditions, overheating and dangerous elongation of the conductor [28]. The standard DIN EN 50341 recommends the consideration of 35°C ambient temperature, 0.6 m/s of perpendicular wind speed, and 900 W/m² of solar radiation for the calculation of SLR [29]. Other TSOs regard seasonal average weather conditions for the calculation of the SLR. However, these conservative assumptions usually apply just a few days a year. [30]

Instead of an SLR, a Dynamic Line Rating (DLR) can be implemented. It consists of a continuous real-time monitoring of the thermal rating of the line. That can be achieved by temperature sensors installed directly on the conductor or by weather stations located in its surroundings, as explained in detail in Section 3.2. With this information, the utilities can calculate the dynamic change of the current-carrying capacity of an overhead line. System operations based on DLR guarantee a safe conductor temperature, enough clearance to ground, and prevent a faster aging of the conductor.

The substantial temporal variations of the thermal rating make it difficult for system operators to exploit its benefits in real-time [10]. Figure 3.1 shows the current-carrying capacity over time for two sample days for the overhead line described in the case study of this dissertation. As seen in the graphic, the transmission capacity can double in a period of twelve hours.

Transmission capacity forecasts are necessary to operate efficiently a flexible electrical network. It is achieved by improving the dispatching plans of one- and two-days-ahead. In fact, the net transfer capacity allocations for cross-border energy markets are carried out two days in advance [10]. Therefore,

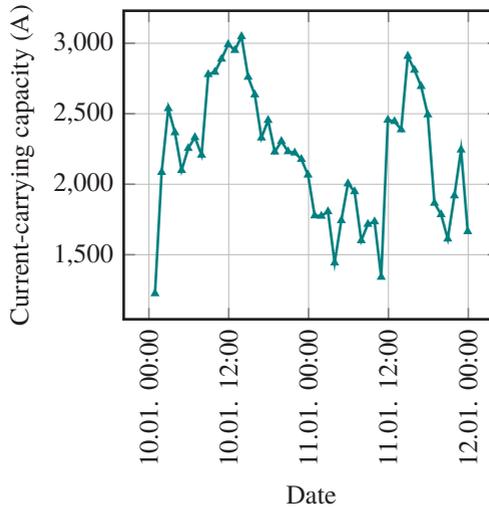


Figure 3.1: Example of current-carrying capacity over time, from January 10th and 11th, 2013, at the overhead line of the case study of this dissertation.

the developing efforts from the last years have turned into the current-carrying capacity forecasts.

The following sections describe, first, the equations necessary to calculate the current-carrying capacity of conductors from the weather conditions (Section 3.1); second, the State-of-the-Art of OLM systems (Section 3.2), and third, the advantages of including DLR forecasting to the existing monitoring devices (Section 3.3).

3.1 Thermal rating of overhead lines

Note: this section is based on the Cigré Standard 601 [9].

As every thermodynamic system, the temperature change of an overhead line depends on the heat exchange between the body and the environment surrounding it. This heat balance can be described, as shown in Equation (3.1), by the

balance between the heat gain from the Joule effect (P_J), the magnetic heating (P_M) and the solar radiation (P_S), as well as by the heat losses corresponding to the radiative (P_r) and convective (P_c) cooling (all of them in W m^{-2}). This balance determines the change in conductor temperature¹ (T_c , °C) over time, which is also influenced by the conductor mass per unit length (m , kg m^{-1}) and its heat capacity (c , $\text{J kg}^{-1} \text{K}^{-1}$).

$$mc \frac{dT_c}{dt} = P_J + P_M + P_S - P_r - P_c \quad (3.1)$$

The solution to Equation (3.1) results in the unsteady state or transient analysis of the conductor temperature. A steady-state solution ($\frac{dT_c}{dt} \rightarrow 0$) may be useful to analyze the maximum temperature level, which the conductor can reach, given its surrounding conditions are fixed over a long time.

3.1.1 Joule heating

The Joule heating, also known as Ohmic heating, is the thermal power generated by the current flowing in the conductor. It can be described using Equation (3.2), where I is the electrical current (A), R_{dc} is the DC resistance per unit length (Ωm^{-1}) and K_{sk} is the skin effect² factor (unitless, usually less than 1.02).

$$P_J = I^2 R_{dc} K_{sk} \quad (3.2)$$

The DC resistance depends on the conductor temperature, as Equation (3.3) shows, where ρ_{20} is the resistivity of the conductor at 20 °C (Ωm), A is the cross sectional area (m^2), α_{20} the linear temperature coefficient at 20 °C (K^{-1}) and ζ_{20} the quadratic temperature coefficient at 20 °C (K^{-2}). The quadratic term in Equation (3.3) is significant for $T_c > 130$ °C. Since the maximum

¹ In this document, the conductor temperature, T_c , refers to the average between its surface and its core temperatures.

² The skin effect is a phenomenon of migration of the AC current towards the surface of the conductor. The result is an increment of the effective conductor resistance [9].

conductor temperature is often under 100 °C, the literature normally refers only to the linear relationship between resistance and temperature.

$$R_{dc} = \frac{\rho_{20}}{A} [1 + \alpha_{20}(T_c - 20) + \zeta_{20}(T_c - 20)^2] \quad (3.3)$$

3.1.2 Magnetic heating

In the case of a steel-cored conductor, as Aluminium Conductor Steel-Reinforced (ACSR)³, the axial alternating magnetic flux produced by the spiraling conductor layers causes heating in the steel core. The transformer effect also takes place, i.e., heating in the layers of the non-ferrous wires due to the redistribution of the current densities.

For many cases in transmission lines, the magnetic heating can be neglected. Therefore, it will not be considered in further analyses.

3.1.3 Solar heating

The sun heats the outer layer of the conductor. Therefore, Equation (3.4) depends on the conductor diameter (D , m), the absorptivity factor (α_S , with values between 0.2 and 0.9) and the solar radiation (S , W m^{-2}).

$$P_S = \alpha_S D S \quad (3.4)$$

The albedo and angle of incidence of radiation can be considered in this calculation [9]. For simplicity, they have been omitted from further analysis. However, in the actual application in the field, they have to be included in the calculations.

³ ACSR conductors are composed of stranded aluminum wires layered around a steel-core. They constitute a good example, because of their common use on transmission lines.

3.1.4 Radiative cooling

Based on the Stefan-Boltzmann law, the thermal radiation heat loss can be described as the power radiated from the conductor in terms of the conductor's diameter (D , m), Stefan-Boltzmann constant (σ_B), the conductor's surface emissivity (ϵ_S), the conductor temperature (T_c , °C) and the ambient temperature (T_a , °C) to the fourth power (see Equation (3.5)).

$$P_r = \pi D \sigma_B \epsilon_S [(T_c + 273)^4 - (T_a + 273)^4] \quad (3.5)$$

3.1.5 Convective cooling

Convection is a phenomenon produced by airflow around the line. It consists of heat transfer from the conductor surface to the surrounding air. It is the most important cooling factor for overhead lines, even at calm wind conditions.

The general formula corresponds to the relationship between the thermal conductivity of the air (λ , $\text{W m}^{-1} \text{K}^{-1}$), the difference from conductor to ambient temperatures, and the dimensionless Nusselt number (N_u), as shown in Equation (3.6).

$$P_c = \pi \lambda (T_c - T_a) N_u \quad (3.6)$$

The thermal conductivity of the air can be calculated from the empirical Equation (3.7), which depends on the temperature of the film of air in contact with the conductor surface (T_f , °C).

$$\lambda = 2.368 \cdot 10^{-2} + 7.23 \cdot 10^{-5} \cdot T_f - 2.763 \cdot 10^{-8} \cdot T_f^2 \quad (3.7)$$

$$\text{where } T_f = \frac{1}{2}(T_c - T_a) \text{ for } T_f < 300^\circ \text{C}$$

The Nusselt number depends on the wind conditions, i.e., natural or forced convection scenarios. Further details about the calculation of the Nusselt number can be found in Appendix A.1.

3.1.6 Current-carrying capacity

The current carrying capacity of an overhead line, also found in the literature as *ampacity*, can be calculated from the heat balance equation considering that the steady state condition applies, i.e., the conductor has reached its maximum permissible temperature ($T_c = T_{c,max}$). Inserting the Joule heating Equation (3.2) into Equation (3.1) and considering $\frac{dT_c}{dt} \rightarrow 0$ (steady state), Equation (3.8) is obtained.

$$0 = I^2 R_{dc} K_{sk} + P_M + P_S - P_r - P_c \quad (3.8)$$

Equation (3.8) can be reorganized, neglecting the magnetic heating, to obtain Equation (3.9), which describes the current-carrying capacity, I_{max} . It calculates the individual capacities at each line section. Then, the current-carrying capacity of an overhead line corresponds to the minimum value between all spans.

$$I_{max} = \sqrt{\frac{P_r(T_{c,max}) + P_c(T_{c,max}) - P_S}{R_{dc}(T_{c,max}) K_{sk}}} \quad (3.9)$$

3.1.7 Transient state

The knowledge of the conductor temperature over time gives TSOs the flexibility to solve momentary congestion cases, without reaching the maximum conductor temperature. The time-dependent conductor temperature is obtained from the heat Equation (3.1) given the electrical current and the weather conditions during the period of study. It is a non-linear ordinary differential equation without a simple, closed-form solution. It can be solved either by numerical integration or by linearizing the radiative cooling term. Since the operation temperature of the conductors is limited to a maximum value of around 100 °C, the linearization entails an acceptable calculation error [31].

Considering constant weather conditions during a short period of time and a step-change in the electrical current, I , then the Equation (3.1) can be rewritten as Equation (3.10), where $(A + BT_c)$ corresponds to $R_{dc} K_{sk}$ from the Joule

heating, $F(T_c - T_a)$ is the linearized term of the radiative cooling, $G(T_c - T_a)$ corresponds to $\pi\lambda(T_c - T_a)N_u$ of the convective cooling term, the magnetic heating is neglected, and the solar heating P_S is constant.

$$mc \frac{dT_c}{dt} = I^2(A + BT_c) + P_S - F(T_c - T_a) - G(T_c - T_a) \quad (3.10)$$

The constant F from the linear radiative cooling comes from the multiplication of $\pi D\sigma_B\epsilon_S$ by a linearization factor E , which can be determined by the intersection points between the fourth order term $(T_c^4 - T_a^4)$ and the linear term $(T_c - T_a)$ [31].

By reorganizing and defining the constants $c_1 = I^2B - F - G$ and $c_2 = I^2A + P_S + FT_a + GT_a$ the simplified linear differential equation (3.11) is obtained.

$$mc \frac{dT_c}{dt} = c_1T_c + c_2 \quad (3.11)$$

The solution is the exponential equation (3.12), which describes the change in the conductor temperature over time.

$$T_c(t) = \left(T_c(t_0) + \frac{c_2}{c_1} \right) \exp\left(\frac{c_1}{mc} t \right) - \frac{c_2}{c_1} \quad (3.12)$$

Thus, it is possible to derive a time constant, τ , shown in Equation (3.13). It expresses the range in time, at which $T_c(t)$ reaches $(1 - 1/e) \approx 0.632$ of its steady state solution. It spans between 1 and 15 minutes, depending on the conductor characteristics, the electrical current, and the weather conditions. Emergency ratings can be executed inside this margin, i.e., taking advantage of the slow change of conductor temperature to transmit a higher electrical current than the steady state carrying capacity.

$$\tau = -\frac{mc}{c_1} \quad (3.13)$$

3.1.8 Influence of weather in the current-carrying capacity

Equations (3.2) to (3.9) show a non-linear relationship between the weather parameters and the current-carrying capacity. An influence analysis is necessary to understand how each weather variable affects the thermal rating of the conductor. It is essential to determine the requirements of the weather stations, as measurement accuracy and resolution. This analysis has been done considering the change in the capacity value for a change of the weather parameter under study. It was calculated as the discrete second-order approximation of the partial derivative of the current-carrying capacity, I_{max} , to the studied weather variable, x , as described in Equation (3.14), with Δx is the calculation step of the weather variable. The smaller it is, the smaller the approximation error.

$$\frac{\partial I_{max}}{\partial x} \approx \frac{I_{max}(x + \Delta x) - I_{max}(x - \Delta x)}{2\Delta x} \quad (3.14)$$

Each weather parameter was derived separately. The not-studied weather variables were kept constant with the following values: ambient temperature 35°C, solar radiation 1000 W/m², wind direction 90° (perpendicular to overhead line) and wind speed 1.2 m/s.

The conductor type considered for the current-carrying capacity calculation is a Drake 26/7 ACSR (outer diameter 28,1 mm, steel core diameter 10,4 mm) with emissivity 0.8, solar absorptivity 0.8 and maximum allowable conductor temperature 100°C. These characteristics allow the direct validation of the current-carrying capacity calculations with the Cigre Standard 601 [9].

The results are summarized in Figures 3.2 to 3.5. They show to the left side the absolute value of the current-carrying capacity, I_{max} , and to the right side the derivative of the capacity with respect to the respective weather parameter in percentage relative to I_{max} . That allows the comparison independent of the ampacity levels.

The negative values for the derivatives with respect to temperature (Figure 3.2) and solar radiation (Figure 3.3) describe the reduction in the current-carrying capacity when those weather parameters increase in value. A significant influence of solar radiation occurs when the parameter changes in the order

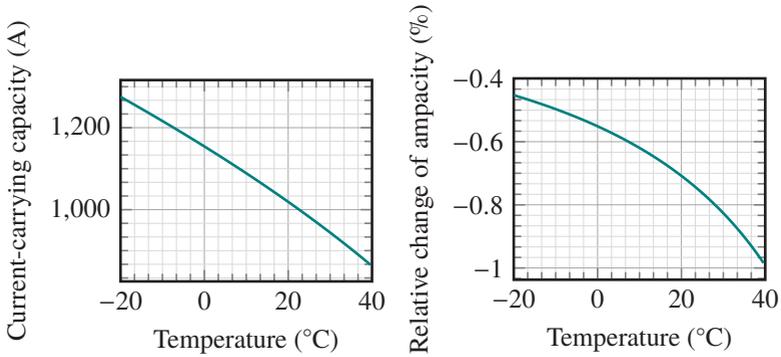


Figure 3.2: Capacity derivative with respect to ambient temperature relative to I_{max} absolute value.

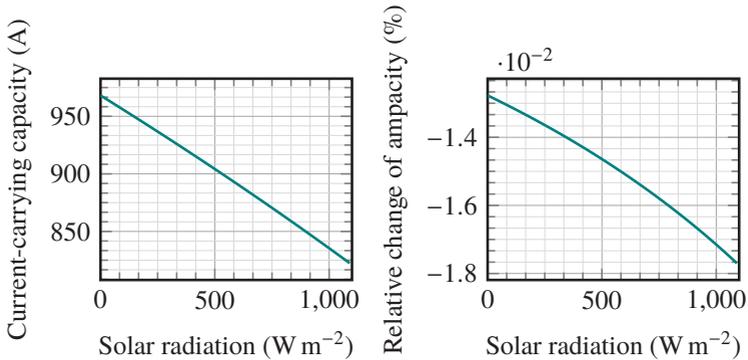


Figure 3.3: Capacity derivative with respect to solar radiation relative to I_{max} absolute value.

of hundreds of $W m^{-2}$, which normally occurs in a sunny day, reaching about $1000 W m^{-2}$.

The most influential weather parameter is the wind speed (Figure 3.4). Changes of slow wind speeds have a significant impact on the current-carrying capacity. The discontinuities seen at 0.15 m/s and 1.85 m/s are due to the calculation of the Nusselt number in the convective cooling term of the thermal model equation, which has different definitions depending on the wind strength.

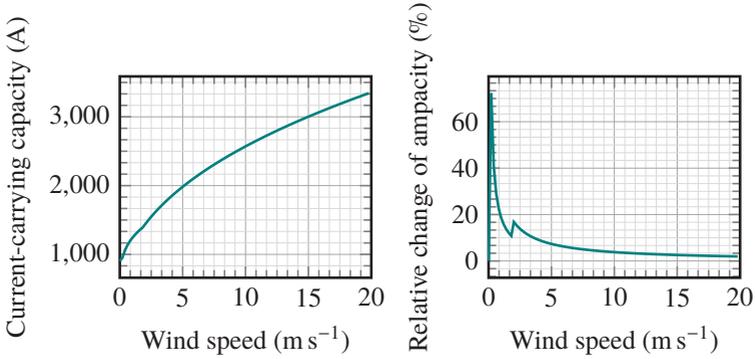


Figure 3.4: Capacity derivative with respect to wind speed (left) and wind direction (right) relative to I_{max} absolute value.

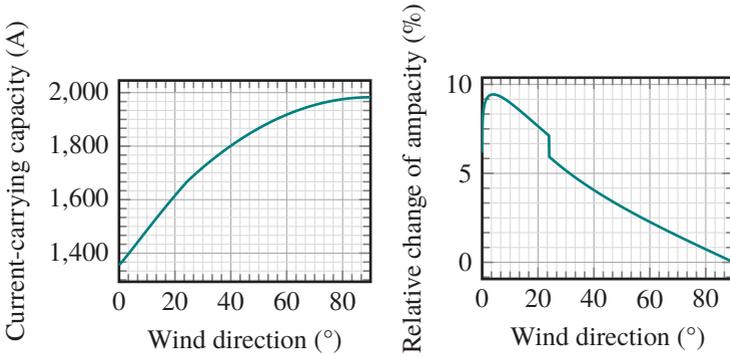


Figure 3.5: Capacity derivative with respect to wind speed (left) and wind direction (right) relative to I_{max} absolute value.

The small influence of the ambient temperature on the current-carrying capacity presented in Figure 3.2 motivated experiments under different weather conditions. As seen in Figure 3.6, changes in the ambient temperature have a greater impact on the values of the current-carrying capacity when the wind is near to 0 m s^{-1} (natural convective cooling).

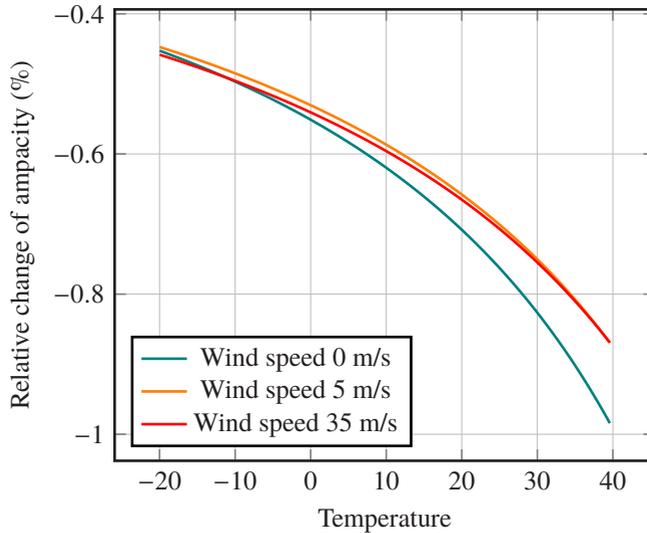


Figure 3.6: Capacity derivative with respect to ambient temperature, considering different wind speeds. The impact of changes of ambient temperature on the value of the current-carrying capacity is larger on natural convection.

3.2 State-of-the-art of overhead line monitoring systems

The first article about the topic current-carrying capacity of overhead conductors was published in 1943 by H. A. Enos of the American Gas and Electric Service Corporation [32]. He described the heat equations for ACSR Conductors, which afterward were used in 1958 for the first time in a case study on an ACSR 220 kV line, lead by Jack Roth, as a project engineer in the System Planning Division of Pennsylvania Power & Light Company [33]. These studies were motivated by the length reduction of overhead transmission lines in the U.S., which changed the limit of transmission capacity from the stability and voltage drop limitation to the thermal limit [34].

Dynamic Line Ratings were calculated in the beginning using only weather information. The first sensor system installed directly on the conductor was the

CAT-1 (see Figure 3.7), introduced in 1991 by Virginia Power in the U.S. [35]. This device measured the mechanical tension between the conductor and its support at the electrical tower. This force is related to the length of the line, and this one on the material temperature.

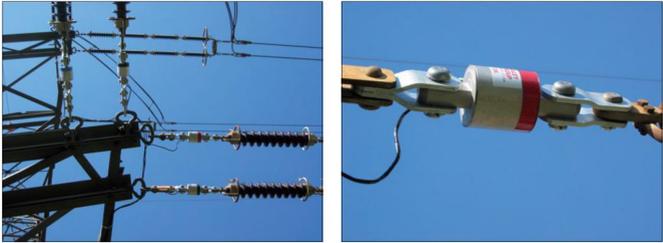


Figure 3.7: CAT-1 was the first OLM system, which was installed for the first time in 1991 by Virginia Power in the U.S. [36].

Since the introduction of CAT-1 in the market, different sensor systems have been developed to measure the conductor temperature in real-time, either directly or indirectly. Figures 3.8 and 3.9 show two examples from the first type, the Emo System from Micca and the OTLM sensor system. This kind of system has the disadvantage of measuring the conductor temperature from a single spot, which is covered from the wind cooling influence by its own body. A calibration process usually corrects this problem.



Figure 3.8: EMO System [37].



Figure 3.9: OTLM [38].

There are several indirect temperature measurement systems in the market. Ampacimon (Figure 3.10) is mounted on the conductor and measures low-frequency mechanical vibrations of the line (induced by wind or thermal convection). The line sag can be derived from its vibration frequency, which relates to the conductor temperature [39], [40].

Astrose is another system, which is installed directly on the conductor (see Figure 3.11). It measures the angle at which the sensor node is at each time step. If the angle changes towards the vertical, the conductor has elongated, and that corresponds to an increment in temperature. The system also has a temperature sensor, which is used as a reference observation [41].



Figure 3.10: Ampacimon [40].



Figure 3.11: Astrose [42].

Optical systems also play a role in OLM. The company Lindsey has a LiDAR⁴ system in the market (Figure 3.12), which measures the distance from the sensor to ground. It is generally installed on the deepest point along the catenary. Therefore, the measurement represents the difference between the height of the line span and the sag [44].

There are OLM systems, which are not mounted directly on the conductor, but at the electrical tower or on the ground. Those positions simplify the mounting process and give flexibility to maintenance procedures. For many TSOs and DSOs, a power outage of the line is a requirement to install any equipment on the conductor. That not only increases the costs but also means a more careful planning process for the installation. However, systems which are not directly

⁴ LiDAR stands for Light Detection And Ranging. It is a remote sensing method that uses light in the form of a pulsed laser to measure distances to objects or surroundings [43].



Figure 3.12: LiDAR System from the Lindsey company [44].



Figure 3.13: LineVision [45].

measuring the conductor temperature or sag, have to be carefully calibrated, and the models have to be checked to guarantee correct observations.

As an example of indirect measuring devices, Figure 3.13 shows the Line Vision system. It consists of an electromagnetic field sensor, which tracks the conductor position. From that information, the line elongation, and hence its temperature are calculated [46].

Weather measurements are also used for DLR calculations. The first application of this approach utilized observations from existing measuring devices. These are normally located at airports or interesting places from the meteorological perspective. Those positions are not necessarily near the overhead line. Hence the observations have to be treated carefully.

Other TSOs install weather stations at the substations, where power supply, and in some cases also wired communications, are available. In Germany, Tennet has at least 20 weather stations distributed along 900 km overhead lines [47]. After a climatology study, the area was divided into climate zones, at which a measuring system was installed. From these observations, an estimation of the current-carrying capacity is done, leaving a safety margin in the calculation to consider spatial measurement errors.

Different kinds of OLM systems are in the market from decades ago. However, they have not seen a broad application in the electrical system operations. The temporal variations of the thermal rating make it difficult for system operators to exploit its benefits in real-time [10]. Therefore, the developing efforts from the last years have turned into the current-carrying capacity forecasts, which allow TSOs to plan the dispatching of the power plants two-days-ahead.

3.3 Current-carrying capacity forecasting

Dynamic Line Rating is an asset for system operations only if operators have enough time to react to the current-carrying capacity changes over time. For this, the two-days-ahead forecast is necessary to operate a flexible network efficiently. In this section, the state-of-the-art systems for DLR forecasts are presented.

The first trials to forecast the rating of overhead lines began at the end of the 80s. The Niagara Mohawk Power Corporation built a database in the period July 1988 until December 1989 with the weather, the conductor temperature, and current measurements for two parallel power lines located in eastern New York State. The sensor nodes consisted of thirteen on-line conductor monitors and five weather stations installed along the power lines. The researchers developed a four-hour forecast based on probabilistic methods. The results showed that 90% of the forecast ratings maintained conservatively critical span temperatures below the maximum allowable conductor temperature [48]. Similar solutions were developed by J. F. Hall et al. [49] and D. A. Douglass [50], both articles published in 1988. All these studies seemed to remain untouched until the beginning of the 21st century when the integration of IRES in the electrical network increased, and the need for optimization systems raised.

From 2010 to 2013 the topic was studied as part of the European project Twenties. The study, called NETFLEX Demo, applied the forecast models to two parallel 150 kV overhead lines of 18 km length in Belgium. They developed a system to forecast the DLR for 24 to 48 hours ahead [51]. The prediction was calculated based on NWP from the Weather Research and Forecasting model. Ten Ampacimon systems were distributed along the pilot lines at a mean distance of 3-4 km as references. The system achieved an average gain

of 10% to 15% over static, seasonal ratings with 98% of confidence [52], [10]. From the reports and articles, it is not clear what methods and algorithms were utilized.

The TSO Tennet from the Netherlands has also reported a DLR forecasting system for 94 km of overhead lines. The system is based on statistical post-processing of NWP, based on local weather observations. The methods and the NWP models used for this are not specified. A roughness map of 100 x 100 m resolution is utilized to account for the effect of terrain roughness on the wind values and adjust them accordingly [53].

Other studies apply probabilistic methods to forecast the current-carrying capacity. T. Ringelband from the Aachen University in Germany [54] developed a method to predict the probability density function of the current-carrying capacity. The algorithm used a Monte Carlo simulation for prediction of the probability density function of each weather parameter, which used the parameter estimation by the expectation-maximization algorithm. The input parameters of the system are weather observations and NWP. It predicts the current-carrying capacity for the following 24 hours.

Aznarte and Siebert (Ecole des Mines de Paris, [55]) implemented three machine learning algorithms (Generalized Linear Models, Multivariate Adaptive Regression Splines and Random Forest) to forecast the weather. From the results they predict 27 hours of the current-carrying capacity as a second step. A similar approach was done by Doban [15] from the KTH Royal Institute of Technology in Stockholm, Sweden. Weather predictions were generated from Feedforward Neural Networks and Support Vector Machines. From the model output the current-carrying capacity was calculated for 48 hours.

Companies as Ampacimon [56] and Lindsey [57] sell today dynamic line rating forecasting software packages, which are based on the measurements of their sensors. However, they have not been completely adopted in Germany because of the high installation costs they carry. German TSOs have the requirement of shutting down the overhead line for mounting any sensor directly on the conductor, which means a clear disadvantage for the application of those systems. On the other hand, the meteorological company Energy & Meteo [58] offers current-carrying capacity predictions based on weather models. Simulations are used to describe the terrain along the line, while actual local

measurements are not gathered, which makes an error estimation difficult to execute.

Dynamic line rating forecasting is still a new field in research. This dissertation focuses on a different approach compared to the state-of-the-art. Instead of concentrating the efforts of modeling a weather forecast, the machine learning models developed in this study optimize directly the prediction error of the current-carrying capacity. Meteorologists do already a good job by describing the atmosphere using physical equations. The approach of this dissertation takes advantage of the existing meteorological models for weather forecast and integrates the distributed local weather observations along the overhead lines to the system.

4 Artificial Intelligence in Power Systems

Artificial Intelligence can be defined as a program or set of programs, which can augment, replicate, and, in some cases even, replace humans [59]. The program can be as simple as a set of rules expressed in a formal language, as the so-called knowledge-based systems. One of the first proposals to include artificial intelligence in power system operations was given already in the 80s, to automate some of the features of the energy management system using knowledge-based or expert systems. The "intelligent" machines were programmed to embed sophisticated heuristics into conventional application software, in order to help system operators to react quicker in emergencies. A knowledge-based system is a collection of information from operation manuals and experience of system operators. It intercepts alarm messages and presents the operator with a summary of the most pertinent information to make quicker decisions [60].

This kind of hard-coded intelligent system faces the problem of inconsistency or undefined rules very often. The ability to acquire knowledge by extracting patterns from data is necessary and that is the core of the machine learning algorithms. The performance of these programs depends on the data preprocessing quality, its representation, and on the selection of the correct input variables, the so-called features. The job of a machine learning algorithm is to learn how the features correlate with different expected outputs. For example, the relationship between a future value of the current-carrying capacity and past weather conditions.

Deep learning algorithms are a type of machine learning method, which builds complex concepts out of simpler representations. The more straightforward the representation, the more detailed (better resolution) the final output can be. The typical example is the object recognition, which breaks down an image into corners and contours, defined in terms of edges [61].

The general relationship between artificial intelligence, machine learning and deep learning is illustrated in Figure 4.1. These three can be found nowadays in applications in the energy sector, as the RES supply prediction and demand forecasts based on weather and smart meter measurements [62].

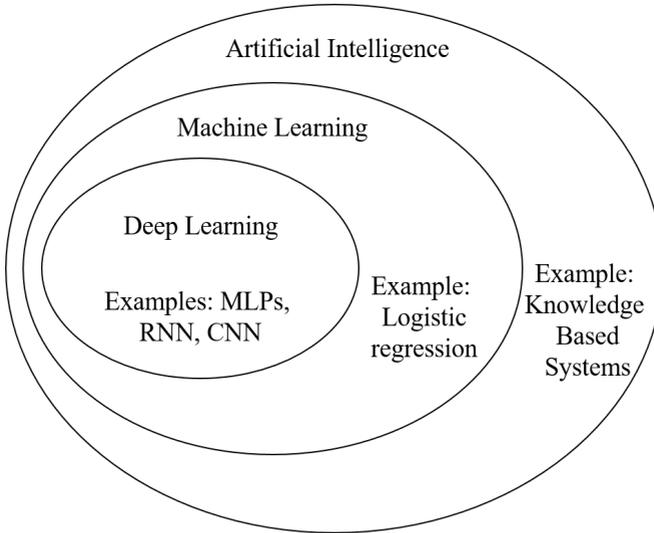


Figure 4.1: Venn diagram for AI and its derivations: machine learning and deep learning [61].

This dissertation focuses on machine learning approaches to solve the problem of current-carrying capacity forecasting. Therefore, the required concepts to develop this kind of solutions are explained in detail in the following sections.

4.1 General concepts of machine learning

Note: if not specified, the following sections are based on the book "Deep Learning" of Goodfellow et al. [61].

The book *Machine Learning* from Tom Mitchell (1997) begins explaining the meaning of *learning* from the perspective of computer programs: "A computer program is said to learn from experience E with respect to some class of tasks

T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" [63]. Hence, to have a well-defined machine learning problem, it is necessary to identify the set of tasks the computer has to fulfill (current-carrying capacity forecast, for example), a metric to score the performance of the computer (mean absolute percentage error) and a source of experience, which can be directly related to the task (historical current-carrying capacities), or indirectly (weather measurements from the past and weather forecasts from numerical models).

4.1.1 Learning methods

Depending on the task and the sort of experience examples available, most of the problems can be divided into supervised, unsupervised, or reinforcement learning paradigms. As the name suggests, a supervised learning problem can be seen as an instructor teaching the machine what to do. The machine receives several examples, as a dataset containing features, x , and the corresponding labels to each example, y . The computer has to learn to predict y from x , usually by estimating $p(y|x)$. These input and output variables can be a time series (as the case of this dissertation), but they can also represent images, text or object classes.

Unsupervised learning algorithms receive a dataset containing the features, x , and are trained to learn the structure of this data set by reconstructing implicitly or explicitly its probability distribution, $p(x)$.

Reinforcement learning algorithms have, on the other side, the possibility to interact with the environment. That creates a feedback loop between the learning system and its experiences. The training process for these models uses a reward (or penalty) function to optimize the output of the system given a series of past events.

Since this dissertation makes use of supervised learning methods, the following sections are focused on them.

4.1.2 Over- and underfitting

The examples used to train the machine learning model are called the training data set. It consists of features, x , and labels, y . The error of the model's outputs during training is called the training error, which has to be minimized. Afterward, the model is measured again using a separate data set containing examples similar to reality, which is called the test set. The goal is to minimize the error of the model on this unseen data set as well. If this is achieved, it is called model generalization. One of the central challenges in machine learning is to achieve a generalized model.

If the model is not able to obtain a sufficiently low error value on the training set, then it is said that the model is underfitting. On the other hand, if the difference between the training and the test error is significant, then it is called overfitting. Neither of them are desirable and have to be avoided while solving a machine learning problem.

4.1.3 Hyperparameter optimization

Most machine learning algorithms have hyperparameters. These are settings that can control the algorithm's performance. Their values are usually not adapted by the learning algorithm itself, but they have to be adjusted manually or automatically by an external searching program. In both cases, a range of options has to be defined. An example of hyperparameters is the number of neurons and layers a neural network has.

A manual search requires understanding what the hyperparameters do and how machine learning models achieve good generalization. Automatic hyperparameter selection algorithms reduce the need to understand the algorithm and the influence of the hyperparameters on them. However, an automatic search can have high computational costs.

Between all hyperparameter optimization techniques, three can be named as the most used nowadays:

- Grid search: the computer tries all possible hyperparameter combinations from the predefined range of options.

- Random search: the program selects a subset of hyperparameters from the predefined range of options randomly, reducing the amount of training processes to run.
- Bayesian optimization: This algorithm is a popular optimization strategy "for finding the extremes of objective functions that are expensive to evaluate" [64]. At each iteration, it makes use of the results of previous iterations employing the Bayes' theorem. Then it generates a set of plausible objective functions based on the previously sampled values using Gaussian processes, which normally reduces the optima searching time.

4.2 Learning algorithms

This section is a compilation of the theoretical description of the main machine learning algorithms applied in the development of this dissertation. Artificial neural networks are the main building block for the explanation of other neural networks; quantile regression forests and recurrent neural networks were the main forecasting models in the project; and convolutional neural networks were used to integrate spatially distributed weather information (as the numerical weather forecasts) into the prediction models.

4.2.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) is the generic term for machine learning models with brain-like structures, i.e., neurons (cells or nodes) and axons (connections between cells). Feedforward Neural Networks (FFNNs), also known as Multilayer Perceptrons, are the most common approach of ANNs, and their main goal is to approximate a function f^* as $y = f^*(x)$, where x is the set of input features of the model.

FFNNs are described graphically as directed acyclic graphs. Figure 4.2 shows a simple example, which will help to explain FFNNs functionality.

A FFNN is a function composition network, i.e., each neuron represents a single function, f_a , also called activation function, of the n inputs features,

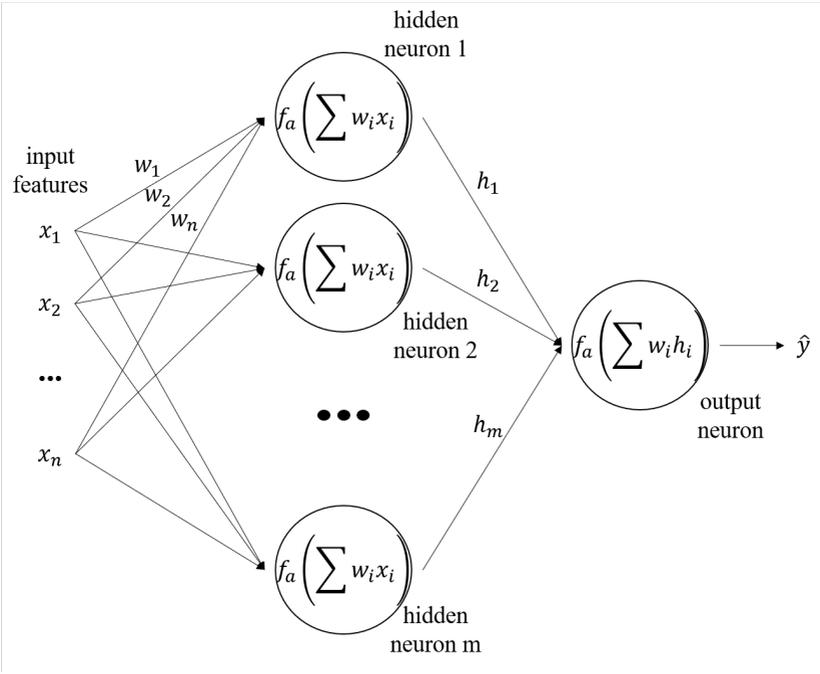


Figure 4.2: Example for a Feedforward Neural Network with a single hidden layer and no bias.

x_1, x_2, \dots, x_n . The output of each neuron corresponds to the input set for the neurons in the next layer, which can be a hidden layer (forming the body from the network) or the output layer. The number of neurons per layer and the total number of layers are hyperparameters, which have to be adjusted to optimize the output error.

The activation function f_a of neuron i is a function of a linear combination of its inputs, as follows, where w_1, w_2, \dots, w_n are the weighting coefficients between neurons, and b is a bias.

$$f_a(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) \tag{4.1}$$

Examples of typical activation functions are the sigmoid, hyperbolic tangent, and the Rectified Linear Units (ReLUs) functions, shown in Figure 4.3. The first two are bounded, normalizing the output of each neuron to the range where the function is defined. However, this characteristic has the disadvantage that for very high or very low input values, there is almost no change to the prediction, causing a vanishing gradient problem. The ReLU function avoids this problem, also having a lower run time than the other two. Nevertheless, this function has the disadvantage of gradients equal to zero for negative inputs.

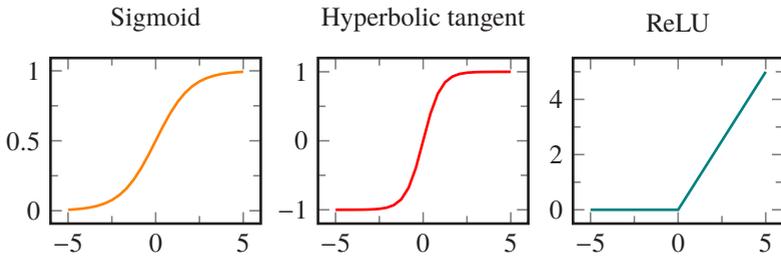


Figure 4.3: Common activation functions [61].

The optimal weights, w_1, w_2, \dots, w_n at the connections between neurons, are found in the training process using an optimization algorithm. Especially in deep learning, the Stochastic Gradient Descent (SGD) algorithm is applied for this. SGD is a version of the well-known optimization algorithm Gradient Descent (GD), at which the whole training set is divided into small sets, called minibatches, to reduce the calculation time while still getting a good optimum estimation.

The GD algorithm optimizes by following the direction of the gradient of a cost function $J(\theta)$. It converges when this gradient tends to zero. The cost function in machine learning is usually the sum over the training examples of a per-example loss function. An illustration of this is the negative conditional log-likelihood function defined by Equation (4.2), where L is the per-example loss $L(x, y, \theta) = -\log(P(y | x, \theta))$, m' is the number of training examples in a minibatch taken randomly from the whole training set, and θ is a vector containing the optimization variables, which in the case of the FFNNs, these are the weights at the connections between neurons.

$$J(\theta) = E_{x,y \sim \hat{p}_{data}}(L(x, y, \theta)) = \frac{1}{m'} \sum_{i=1}^{m'} L(x^{(i)}, y^{(i)}, \theta) \quad (4.2)$$

A numerical calculation of the gradient to execute the SGD algorithm can be computationally expensive. Therefore, its estimation for the cost function, $J(\theta)$, with respect to its parameters, is typically done using the back-propagation algorithm. The procedure consists of applying the chain rule recursively to write the derivative of the loss function with respect to the output vector, \hat{y} , as a multiplication of the derivatives of the outputs of each neuron with respect to its respective inputs.

4.2.2 Recurrent Neural Networks

Recurrent neural networks (RNNs) are a type of neural networks for processing sequential data, $x(1), \dots, x(t)$. RNNs can have different architectures and hidden connections. They either produce an output at each time step or read an entire sequence to produce a single output. They can either have recurrent connections between hidden units or from the output at one time step to the hidden units at the next time step. The selection of the architecture depends on the application.

Figure 4.4 shows as an example the architecture for an RNN with an output at each time step and with recurrent connections between hidden units. The input variable, x , is parametrized by a weight matrix U , the hidden-to-hidden recurrent connections are parametrized by a weight matrix W , and the hidden-to-output connections are parametrized by a weight matrix V . The RNN shares the same weight matrices, U , W , and V , across all time steps.

The inner state, h , at the time step, t , corresponds to the result of an activation function, f , with the linear combination of the last inner state, $h(t-1)$, and the input, $x(t)$, as argument. The output of the example shown in Figure 4.4, $\hat{y}(t)$, corresponds to a scaled and biased version of the inner state. Both mathematical expressions are presented in Equations (4.3) and (4.4).

$$h(t) = f(W h(t-1) + U x(t) + b) \quad (4.3)$$

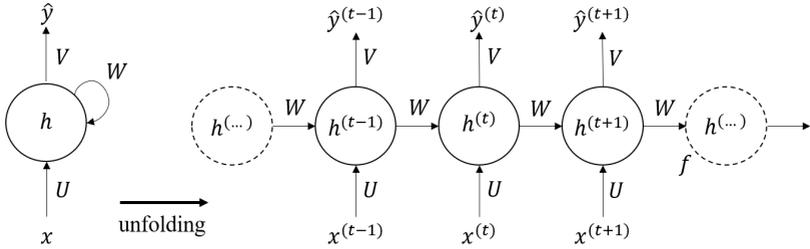


Figure 4.4: Example of the structure of a Recurrent Neural Network [61].

$$\hat{y}(t) = Vh(t) + c \quad (4.4)$$

The input variable, x , is typically operated on minibatches of the original sequential dataset. SGD is, therefore, also valid to train RNNs.

Two RNN architectures were studied during the development of this dissertation due to the good performance the literature declares they have to predict time series. Long short-term memory networks modelled the current-carrying capacity prediction, having as input the weather observations along the overhead line of the case study. Gated Recurrent Units helped in the feature engineering process to determine which input features were the most suitable for the transmission capacity forecasting. The following sections describe them.

Long short-term memory networks

Long short-term memory (LSTM) networks are a sub-group of Recurrent Neural Networks, but unlike general RNNs, LSTM networks have a unique formulation that allows them to avoid the problems of vanishing and exploding gradients (weight changes that quickly become so small as to have no effect or so large to result in overflow) [65]. LSTM networks have also been shown to learn long-term dependencies more easily than the simple recurrent architectures [61].

Instead of having constant weights, U , W , and V , as in the case of the traditional RNN, LSTM networks have gates, which play the role of variable weights. Depending on the input and the last inner state, the gates decide the multiplication factor to use. That creates a mechanism to forget the old state or to turn off the output, depending on the learned rules.

The LSTM gates can be divided in three [65]:

- Forget Gate: Decides what information to discard from the cell.
- Input Gate: Decides which values from the input are used to update the memory state.
- Output Gate: Decides what to output based on the input and the memory of the cell.

They are built up through exposure to the inputs every time step and are used in the calculation of the output.

Chung et al. explain that "unlike the traditional recurrent unit which overwrites its content at each time-step, an LSTM unit is able to decide whether to keep the existing memory via the introduced gates." That means, that LSTM networks are able to detect long-term patterns in a sequence and carry them to the output. [66]

Gated Recurrent Units

Gated Recurrent Units (GRUs) make use of gates to modulate the information flow through the model, as LSTM networks also do. That makes them able to detect patterns in the input sequences and carry them to the model output. However, GRUs do not have the possibility, as LSTM networks do, to control over the output gate the content, which flows to other units in the network. In other words, they do not have an output gate. That makes GRUs simpler to train, by reducing the amount of network parameters. In fact, GRUs show better performance than LSTM networks for small datasets, in terms of convergence in CPU time, of parameter updates and generalization [66].

4.2.3 Convolutional Neural Networks

Note: This section is based on the books of I. Goodfellow et al. "Deep Learning" [61] and of J. Brownlee "Deep Learning with Python" [67]

Convolutional Neural Networks (CNNs), are a kind of neural network for processing data with a grid-like topology. Images are an example of this kind of data since they can be thought of as a 2D grid of pixels. The CNN expects and preserves the spatial relationship between pixels by learning internal feature representations.

The network employs a mathematical operation called convolution, instead of the general matrix multiplication. The convolution occurs between an input matrix for the convolutional layer and a kernel. The latter is drawn across the entire matrix, generating an output matrix called the feature map. Figure 4.5 illustrates how the convolution is calculated. In this example, the kernel is moved a single pixel at the time. However, the distance that it is moved can vary and is referred to as the stride.

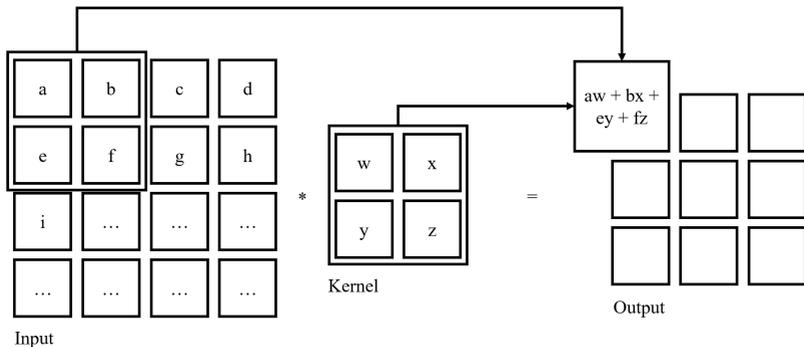


Figure 4.5: Example of a 2D-convolution, with stride equal to one, and no padding [61].

If the size of the input matrix is not divisible by the size of the kernel's receptive field and the size of the stride, then the receptive field can attempt to read of the edge of the input matrix. In this case, techniques like zero padding can be used to invent mock inputs with zero values for the receptive field to read.

These convolutions can also be applied to a sequence of images, as in a video. Also, as Section 7.3.2 explains, it can be used to process the numerical weather predictions arranged in a grid of four dimensions (two dimensions for the horizontal spatial distribution, one for the prediction scope, and one for the weather parameters). Since the input parameter is not a matrix anymore, but a 3D version of it, the variable will be called from now on a tensor.

In the case of having a sequence of 3D-tensors as input, the convolution is called a 3D-convolution. The kernel, shown as a blue box in Figure 4.6, multiplies an extract of the image from the same size and rolls over the whole image. This process generates a new 3D-tensor for each input matrix, i.e., a total of m new tensors, which are added up together to generate a single output tensor. The number and sizes of the output tensors depend on the number and size of the kernels available.

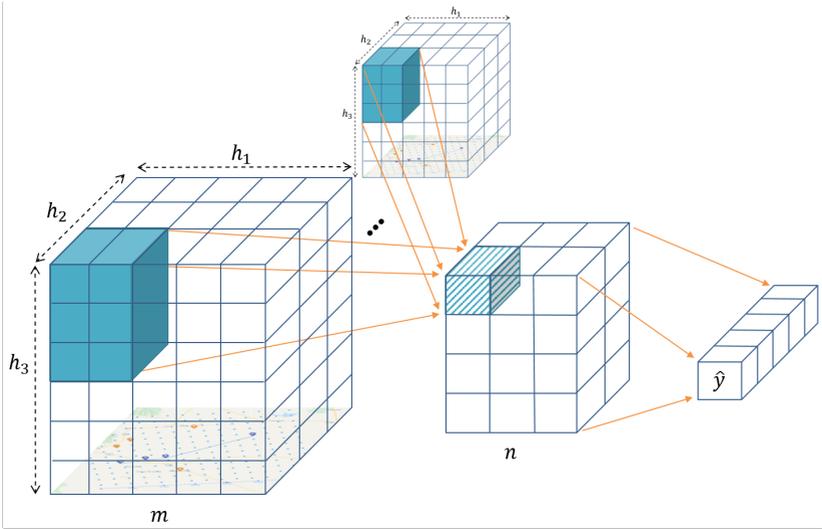


Figure 4.6: Description of 3D convolution. The input corresponds to a sequence of 3D matrices, with dimensions $h_1 \times h_2 \times h_3$ and m the number of elements in the sequence. The output can be an array as shown in this example.

A technique for CNNs to compress or generalize feature representations and generally reduce the overfitting of the training data by the model is called pooling. The pooling layers down-sample its input matrix by taking the average

or the maximum of the input values. They typically follow a sequence of one or more convolutional layers and are intended to consolidate the features learned and expressed in the previous layer's feature map.

Fully connected layers are used at the end of the CNN after the convolutional and pooling layers have performed the feature extraction and consolidation. They are used to create final nonlinear combinations of the features and for making predictions. Fully connected layers are a standard FFNN layer. They may have a nonlinear activation function (sigmoid, hyperbolic tangent, ReLU) or a softmax activation, which output probabilities of class predictions.

4.2.4 Quantile Regression Forests

Quantile Regression Forests (QRFs) correspond to a supervised, probabilistic regression algorithm, which consists of a composition of several trees in the form of a random forest. In graph theory, a tree is defined as a continuous graph with no closed loops, which is built with a set of nodes and edges [68]. A binary tree is, therefore, a kind of tree, which nodes always divide into two other nodes and their edges have a defined direction. The node that has no preceding node is called the root, and the ones with no following nodes after them are called leaves. When these trees are used to classify data, then they are called binary decision trees.

Figure 4.7 shows a simple one-dimensional binary classification problem. The tree has a real number x as input variable. The condition $x \leq t$, with t a threshold, leads to classify the input as y_1 , otherwise as y_2 . The threshold value, t , can be adjusted based on the data, statistics, or heuristics, depending on the application. Decision trees with numeric outputs, $y \in \mathbb{R}$, are used for regression instead of classification problems and are called regression trees.

Binary decision trees can be also used to classify multidimensional variables. For example, in Figure 4.8(a) the graph has as input a two-dimensional variable X . In this case, a plane is spanned and subdivided in partitions based on the threshold values t_i with $i \in 1, 2, 3, \dots$, as shown in Figure 4.8(b) [69].

An ensemble of outputs from several binary trees, trained with N random selected input data elements, is called a random forest [70]. It can be used to solve classification and regression problems, as by simple decision trees.

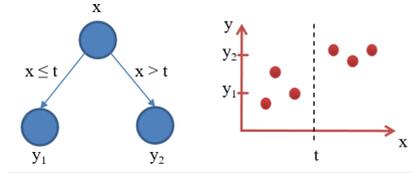


Figure 4.7: Example of a binary tree

A prediction \hat{y} can be calculated as a weighted average of all outputs y_i with Equation (4.5), where the weights w_i are defined by Equation (4.6) [71].

$$\hat{y} = \sum_{i=1}^N w_i y_i \quad (4.5)$$

$$w_i = \frac{1}{N_b} \sum_{k=1}^{N_b} \frac{1_{X_i \in R_{l_{k,p}}}}{\#\{q : \mathbf{X}_q \in R_{l_{k,p}}\}} \quad (4.6)$$

Random forests are robust against overfitting, by creating random subsets of the features and building decision trees using these subsets. Afterward, it combines all subtrees to generate a single prediction result.

QRFs also have the advantage that they work with quantiles. That means that the distribution function of the input data, X , is estimated, as Equation (4.7) shows. For a continuous distribution function, the α -quantile $Q_\alpha(x)$ is defined such that the probability of y being smaller than $Q_\alpha(x)$ is equal to α , given $X = x$ [71].

$$Q_\alpha(x) = \inf\{y : F(y|X = x) \geq \alpha\} \quad (4.7)$$

Quantile regressions can also be used to create prediction intervals, as the 50% interval, defined in between the 0.25 and 0.75-quantiles, given by the Equation (4.8). The higher the prediction interval, the bigger is the probability that, for a $X = x$, the output y is in the interval $I(x)$ [71].

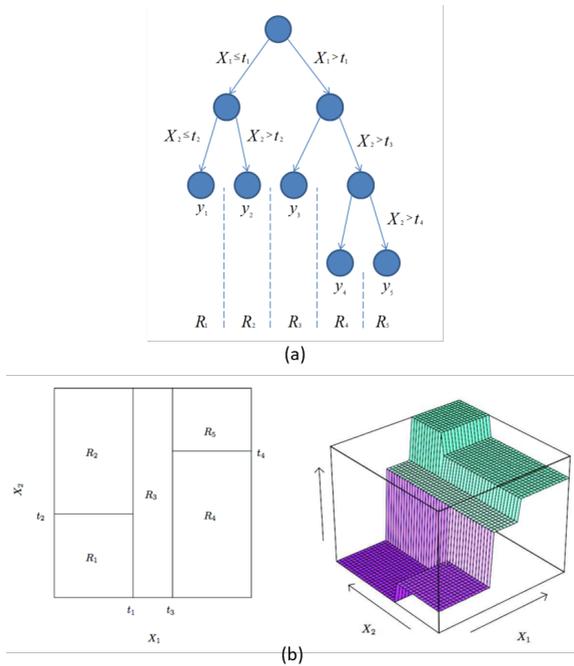


Figure 4.8: Example of a multidimensional binary tree [69].

$$I(x) = [Q_{0.25}(x), Q_{0.75}(x)] \quad (4.8)$$

The QRF models described in this dissertation focus on the prediction of the 0.5 quantile, which corresponds to the median of the output variable distribution. This approach allows the comparison of this model to other solutions.

4.3 Evaluation metrics

The machine learning models can be compared to each other in a benchmark if tested under the same conditions. For this, a set of evaluation metrics has

to be defined. They primarily consist of the difference between the forecast produced by the model and the ground truth, or label.

In PrognoNetz the Mean Absolute Percentage Error (MAPE) plays a main role. It is defined with Equation (4.9), where \hat{y} is the predicted value, y the true value, and N the number of samples. This metric is expressed in percentage when multiplied by 100. It favors risk-averse systems. The MAPE gives importance to errors occurred at low current-carrying capacities, which have a higher probability of being reached in everyday operations.

The research project from the KTH Royal Institute of Technology and the Swedish grid operator Vattenfall established a maximum MAPE of 20% as an evaluation metric for transmission capacity forecasting models [15]. The same limit is used in this dissertation. Moreover, from discussions with system operators at Transnet BW, the most important prediction hours are the first ten hours. Considering that the actual current-carrying capacity of an overhead line is smaller than the predicted, it is necessary to react on time to avoid overloads. The generation adjustment for conventional power plants takes from six to eight hours. Hence, accuracy improvements in the first ten hours are considered valuable in the following analysis.

$$MAPE = \frac{1}{N} \sum_1^N \left| \frac{\hat{y} - y}{y} \right| \quad (4.9)$$

On the other hand, the Mean Absolute Error is an average measure of the error between the prediction and the actual values, given by Equation (4.10). It is often used for error evaluation in the same units as the variable under study.

$$MAE = \frac{1}{N} \sum_1^N |\hat{y} - y| \quad (4.10)$$

The study of the raw error, e_i , is also important when evaluating forecasting models. It is defined as the simple subtraction between the true values and its predictions for every time step, i , as Equation (4.11) describes.

$$e_i = y_i - \hat{y}_i \quad (4.11)$$

5 The PrognoNetz System

As explained in the last sections, TSOs need current-carrying capacity predictions to plan the power generation accordingly and avoid congestion in the electrical grid. The transmission capacity is subject to the weather conditions in the locality of the conductor. The weather changes along the overhead lines depending on the surrounding topography (vegetation, mountains, rivers, etc.). Therefore, the combination of weather measurements along the line and the meteorological models are necessary to obtain an accurate prediction of the current-carrying capacity. This idea is the basis for the research project PrognoNetz, which is the framework of this dissertation.

As illustrated in Figure 5.1, the system consists of weather stations distributed along overhead lines and a central station, where the measurements are collected and the forecast models are running. The TSO has available a graphical user interface to receive information and interact with the system.

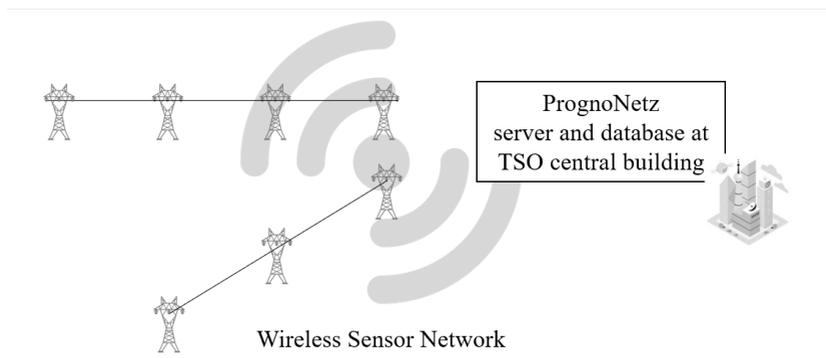


Figure 5.1: PrognoNetz System Diagram.

The project is a cooperation between the Karlsruhe Institute of Technology (KIT) and five industrial partners, financed by the German Federal Ministry

of Economic Affairs and Energy¹ for the period 2019 to 2021. Wilmers Messtechnik GmbH is in charge of developing an affordable, reliable, and energy-efficient weather station. The company Unilab AG implements the information technology infrastructure. UBIMET Deutschland GmbH is a meteorological company in charge of improving the weather interpolation and prediction methods. The TSO, Transnet BW GmbH, and the weather sensor company, GWU Umwelttechnik GmbH, are associate partners. They collaborate on sharing their knowledge, as well as providing the necessary infrastructure for the field tests of the system.

The research team at the KIT has three main tasks in PrognoNetz:

- The development of a tool for weather-dependent grid operation with artificial intelligence capabilities to forecast the current-carrying capacity of the electrical network. It includes a method to determine the number and position of the weather stations along the lines. This dissertation focuses in this task.
- The invention of an installation procedure for the weather stations on the electrical towers using drones. Nowadays, a technician (after required training) has to climb the electrical towers to mount the sensors. A wireless sensor network distributed over the whole electrical grid is only possible if the installation process can be finished as quickly and as cost-effective as possible. This approach should reduce the total time necessary for sensor mounting, allowing the possibility to install more stations per day.
- The design and test of an optical wind sensor for measurements along the line rather than at the sensor location. The correct measurement of the wind is essential for an accurate calculation of the current-carrying capacity. The typical wind sensors measure locally the wind speed and direction. However, the tower structure creates turbulence in the airflow, translating into measurement errors. That can be avoided if the wind is observed far away from the tower.

¹ Bundesministerium für Wirtschaft und Energie, BMWi

This dissertation presents the first results in the machine learning modeling for current-carrying capacity forecasting based on distributed weather measurements and combined with meteorological predictions. The models were developed on a case study, which allowed their proper validation and testing. Moreover, a weather station positioning procedure has been developed and is also described in this dissertation. But first, the following section studies the benefit in terms of congestion avoidance of using a dynamic line rating in the power grid management.

5.1 Investment analysis of dynamic line rating in Germany

The literature on OLM systems [72], [35], [40] agrees on a median potential increment of the current-carrying capacity between 5% and 20%. However, a one-to-one translation from this improvement to the reduction rate on redispatching measures and its corresponding costs requires a comprehensive analysis of the network under review. On one side, the increment in transmission capacity for a selected overhead line does not necessarily mean an improvement in the congestion situation of the electrical network. Since congestion is a localized and temporal phenomenon, it is necessary to study each region as a separate entity with its load and generation capacity.

On the other side, the current-carrying capacity changes over time. This 5% to 20% corresponds to the median of the increment in transmission capacity. The consideration of the DLR curve overtime is necessary to estimate the benefit related to the system.

The current redispatching scenario of the German electrical grid is publicly available and described in detail in the yearly monitoring reports of the German Federal Network Agency. From the report of 2019 [5], Germany had in 2018 a total cost between redispatching (14.9 TWh), network reserve (0.9 TWh) and feed-in management measures (5.4 TWh) of 1438 Million Euros.

In order to calculate how much redispatching and feed-in costs can be saved when using OLM systems in network operations in Germany, it is necessary to have information from each overhead line in the German electrical grid. Data about conductor type, maximum allowed conductor temperature, and

geographical coordinates of each electrical tower, as well as conductor temperature measurements or meteorological observations in the locality of the line routes, are necessary.

Part of this information is not publicly available for critical infrastructures, as the electrical power system. However, our project partner at PrognNetz, the TSO Transnet BW, had provided the data for the overhead line shown in Figure 5.2 as an example.

On the other hand, the German Weather Service (Deutscher Wetterdienst, DWD) has made available a collection of observations since the 1950s. These measurements cover the whole country and offer a first approach for calculations of yearly climate conditions.



Figure 5.2: Overhead Line given by Transnet BW as case study for the benefit analysis of OLM systems.

Given the conductor information and weather data interpolated along the given overhead line, the DLR from 2016 to 2018 was calculated. At each time step, the DLR was compared to the corresponding SLR. Currently, this overhead line is operated under a seasonal thermal rating, i.e., there is a current-carrying capacity limit for summer and another for winter, both fixed over time. The

additional capacity rate can be seen in Figure 5.3 as an increment factor in percentage over time.

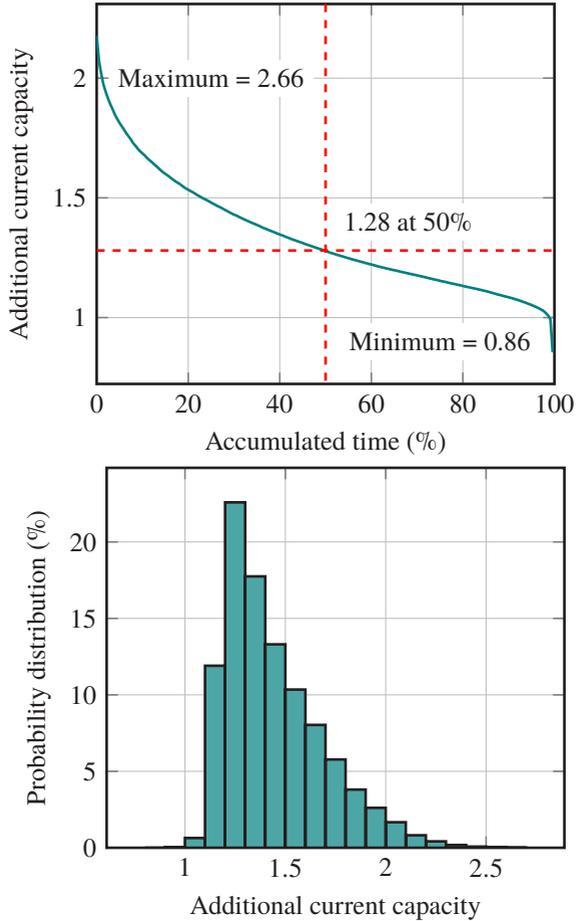


Figure 5.3: Additional current capacity over SLR for the period 2016 to 2018

The following observations can be withdrawn from the capacity rate plot:

- The DLR is 50% of the time 28% bigger than the SLR.

- The best-case-scenario corresponds to an increase in transmission capacity of 165% (2.655 times the SLR). A factor over 2 can be found, however, in less than 1.3% of the time.
- There is an over temperature risk, i.e. a capacity rate smaller than one, in less than 1% of the time.
- The minimum rate factor is 0.85 times the seasonal SLR. Rates smaller than 0.9 are spread from may to july (see Figure 5.4),

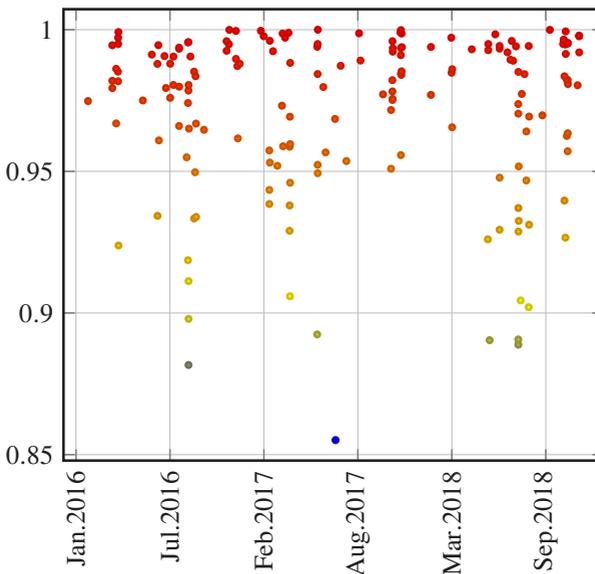


Figure 5.4: DLR cases under SLR for the period 2016 to 2018. The over-temperature risk zone corresponds to about 0.7% of the time.

The severity of over temperature risks depend on the expected demand at those periods. The demand in Germany is from May to July in average lower than in other months of the year, as explained by DBEW in [73]. This explains how overhead lines have worked for decades only based on a seasonal SLR, without damaging the conductor or surpassing the minimum distance to ground.

The prior analysis has resulted in a dynamic factor of capacity increment for DLR over SLR specific for a German region. Considering, as a first approach and under certain error margin, that the meteorological parameters change evenly in the country, it is possible to translate this factor to the redispatching calculations of the whole electrical network.

Redispatching is obtained as a result of the difference in power flow with and without the consideration of the corresponding current-carrying capacities for each line. For this calculation, the electrical system was simulated as a five bidding zones model, based on the work of Staudt [27]. Each zone represents a German TSO (except Tennet, which is divided into a north and south regions). The intra-zonal current-carrying capacities are based on the publicly available static network models of each TSO, status 2017 (Transnet BW [74], Tennet [75], 50 Hertz [76], Amprion [77]).

The following criteria was considered by Staudt for the redispatching calculation and the same are the base for the benefit analysis of OLM systems presented here:

- The transmission capacity is reduced to 50% of the original value to account for the N-1 criterion². Although in reality it is not necessary to reduce the transmission capacity that much, it is considered a conservative but still good approach to simulate safety in the network.
- The reactance was assumed as equal for all lines.
- The power flow was linearized before solving (DC power flow).
- The transmission is considered lossless.
- Staudt distributed the demand into each zone according to the regional gross domestic product and the generation based on the location of the power plants (information from the Bundesnetzagentur).
- With respect to the marginal costs per power plant, Staudt considers that the plant operators "do not act strategically and bid their marginal cost

² The N-1 criterion states that the network has to stay stable, even after failure of any of the system components.

of production" [27]. The marginal cost per power plant was based on Leuthold [78].

For the purposes of the investment analysis presented in this dissertation, the redispatch costs were calculated for the given transmission capacities in the five zones model from Staudt, as well as for DLR as the factor calculated before for the case study. The results for 2018 are shown in the Table 5.1.

Table 5.1: Comparison of the German redispatching situation under SLR and DLR.

	SLR	DLR	Rate of decrease
Redispatching (TWh)	69.26	40.09	42.1%
Redispatching costs (Mio. Euros)	304	135	55.6%

The redispatching results for the SLR case, in TWh, are larger than the actual values reported by the Agency of Energy of Germany, while the costs are smaller. Among the reasons are, first, that this study is not considering imports and exports, which would reduce the amount of redispatching. Second, the costs per TWh considered in the calculations are the minimum possible (marginal costs without any profit). The costs in reality are higher because of the dynamics of the free market. Third, the approximations considered in the redispatching calculation contributed to the deviations as well.

The message that can be taken from this analysis is the relative improvement of the redispatching levels when using dynamic line rating. The possibility to avoid around 40% of line congestion represents a benefit from overhead line monitoring systems, which requires an in-depth study. Extensions of this analysis can include a correct market simulation, running an AC power flow (instead of DC), stability studies, and the calculation of the DLR for each overhead line (this requires the exact coordinates of each line route, tower heights, the corresponding conductor types, and weather observations distributed over the whole electrical grid).

6 Case Study

Meteorology has earned the role of a valuable information source in power systems. Humans act according to the current weather conditions. Therefore, the electrical load has always been influenced by the weather. Since the number of IRES has been increasing, the electrical power generation became more susceptible to weather changes as well. To maintain the system stability, TSOs and DSOs now depend on reliable weather forecasts, which allow them to plan the generation from conventional power plants one to two-days ahead. However, the current-carrying capacity is determined by the hottest overhead line section, which can be in an area below the forecast model resolution.

This dissertation presents a machine learning model solution to improve the forecast accuracy of the current-carrying capacity of the electrical network. It is based on the combination of Numerical Weather Predictions (NWP) and distributed weather observations covering the right of way of a virtual overhead line, which was carefully created from a high-quality weather observation database. This approach allowed the experimentation with different amount of historical data. The knowledge acquired in the development process is now applied in the field, under the framework of the research project PrognNetz.

The following parts describe the weather measurement and NWP datasets used in this case study. Afterward, the criteria to build the virtual overhead line is explained. The correct formatting and pre-processing of the weather data has been the basis for creating the dataset of this dissertation. It is called here the PrognNetz dataset. It is available online as an open-source solution for validation purposes (see Section 6.4).

6.1 NOAA INL Mesonet

The Idaho National Laboratory (INL) is a facility of the United States Department of Energy, which started in 1949 monitoring the weather conditions surrounding the INL. This has been done in order to "describe the meteorology and climatology [of the National Reactor Testing Station] with the focus on protecting the health and safety of site workers and nearby residents" [79].

Weather observations from the INL Mesoscale Meteorological Monitoring Network (INL Mesonet) are available online, published by the National Oceanic and Atmospheric Administration (NOAA) [80]. The data is well maintained and documented. Climatography studies have been published [79], which were of help by designing the experiments of this dissertation and understanding the results obtained.

The database contains high quality, WMO-compliant, weather measurements. The necessary parameters to calculate the current-carrying capacity (ambient temperature, wind, solar radiation) are available, as well as pressure and relative humidity. Weather observations for ambient temperature and wind are available at two different heights, at WMO-Standard¹ and at 15 m above ground. The latter pictures a more realistic current-carrying capacity forecast system scenario. Therefore, they have been considered in this case study. The mean distance between weather stations in the INL Mesonet is around 15 km, which is very dense compared to other weather networks. As an example, the average distance between weather stations from the German Meteorological Service is 25 km according to the description in [81]. Information about the hardware and the position of the weather stations can be found in [79], [82].

Eight weather stations were chosen from a pool of 35 available at the INL Mesonet, as shown in Figure 6.1. The selection had to form plausible geographical arrangements that could model the paths of an overhead line. Furthermore, it had to have a sufficiently large dataset, covering a period of time, which ensures statistical validity of the experiments. This case study is based on measurement data from 2007 until 2017 at a hourly resolution.

¹ The WMO-Standard corresponds to an ambient temperature measurement at 2 m and wind at 10 m above ground.



Figure 6.1: Case study - selected weather stations from the INL Mesonet.

6.2 Numerical Weather Prediction database from NOAA

Numerical Weather Predictions (NWP) are a projection of observational weather data forward in time based on the physical laws of thermodynamics and fluid mechanics (conservation of momentum and continuity). The atmosphere is described by a set of partial differential equations, which are not solved analytically but numerically. Therefore, the resolution and accuracy of the weather predictions depend strongly on the computing power available [83].

The atmosphere is discretized into boxes to solve the NWP models, forming a grid. It has different spatial resolutions depending on the space to cover. For each case, the model can be identified as global, regional or high-resolution local model.

As for every set of partial differential equations, the initial and boundary conditions are of utmost importance to solve them. In meteorology this is achieved

based on the observations from weather sensor networks, which include not only weather stations for the surface information but also satellite, radar and radiosonde. The observations have to be mapped to the corresponding grid resolution using data assimilation techniques. It has to be done carefully, since a simple transfer of the new observations into the forecast field could be physically unrealistic. Shonk [83] gives an example in his book "Introducing Meteorology: a guide to weather" of a wrong data assimilation, at which a surface pressure observation of 1025 hPa was considered into a gridbox that is surrounded by pressures of 1020 hPa in the forecast field. The sharp pressure gradient is unrealistic, which can lead the model to generate unwanted atmospheric waves.

The Continental US Model (CONUS) is a historical weather forecast database from the National Oceanic and Atmospheric Administration (NOAA), which covers the INL Mesonet region. Its spatial resolution is 5 km (CONUS 5) until September 2014. Afterward, the model was updated to CONUS 2.5, offering a higher resolution of 2.5 km. In order to maintain homogeneity over the whole study, while having enough data to obtain statistically meaningful results, the dataset was restricted to CONUS 5, from 2009 to 2014. The time resolution is 3 hours, up to 72 hours. The available weather parameters are wind speed, wind direction and temperature at surface level [84].

NOAA offers Mean Absolute Error (MAE) analyses for each weather element as a verification of the forecast models. The plots in Figures 6.2 and 6.3 were generated based on the data from the NOAA NDFD Monthly Verification web page [85]. They show the monthly (scatter) and the yearly MAE (line) for the year 2018, for a forecast scope of up to seven days.

These curves correspond to the average of the absolute differences between the observations at a weather station and the forecasts at the nearest NWP grid point. The results comprehend measurements of 1320 weather stations distributed over the continental US territory, giving a temporal and spatial overview of the prediction models.

The MAE of the NWP shows a periodicity of 24 hours. It tends to be smaller at night, in comparison to the daylight hours, considering the update time once a day at 00 UTC. The sun is the main figure on the generation of different meteorological conditions in the earth [83]. Wind patterns usually follow a day-night sequence, as temperature does as well. The weather situation

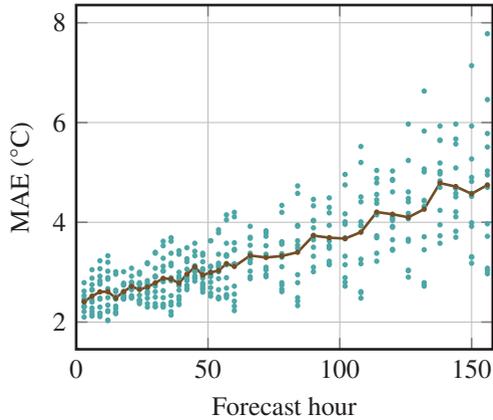


Figure 6.2: MAE for seven days of surface temperature forecast. Monthly errors are shown as a scatter plot. The yearly mean of the forecast error corresponds to the black curve. Data source: NOAA NDFD verification [85].

depends on other factors besides the sun, as cloudiness, global wind flows, and seasons. However, the certainty of having zero solar radiation at night is a hundred percent. This fact assists the atmospheric models, leading to a relative improvement of the forecast error.

This dissertation focuses on a 48 hours prediction scope. The plots in Figure 6.4 show an hourly prediction error analysis specifically for the case study of this dissertation. The verification corresponds to the difference of prediction to observations, from the nearest neighbor NWP grid point to the weather station 690 of the INL Mesonet (see Figure 6.1). The statistics cover the years 2009 to 2014. There is a height difference between the observations from the INL-Mesonet considered in this dissertation and the NWP. The former is at 15 m, and the latter corresponds to 2 m for temperature and 10 m for the wind. Therefore, the error plots in Figure 6.4 comprehend the modeling, the height, and the horizontal distance errors between both datasets.

The 24 hour periodicity seen in the MAE of the NDFD verification of NOAA prevails in the 48 hours forecast MAE for the weather station 690 of the INL-Mesonet (comparison of Figures 6.2 and 6.3 to 6.4). The relative increment of

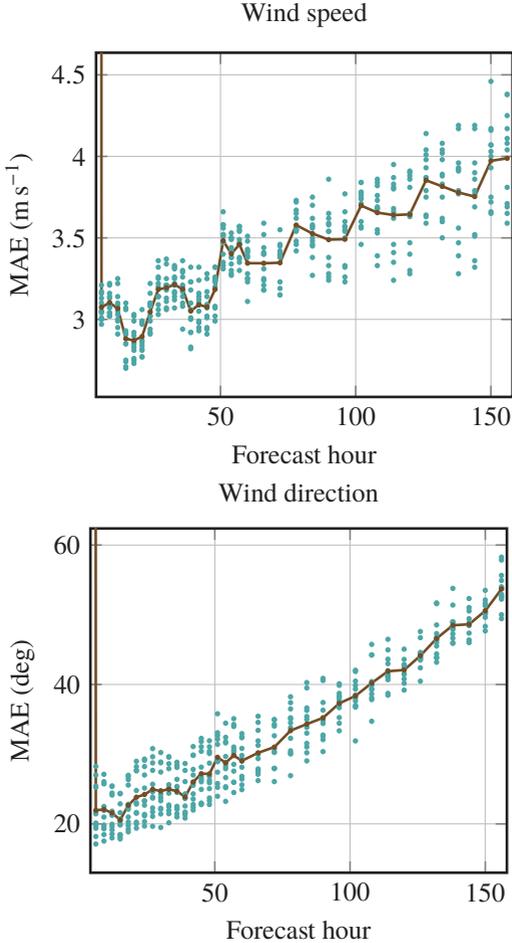


Figure 6.3: MAE for seven days of wind speed and direction forecast. Monthly errors are shown as a scatter plot. The yearly mean of the forecast error corresponds to the black curve. Data source: NOAA NDFD verification [85].

the prediction error with the forecast scope is not considerably as clear in the 48 hours verification, which can be explained by the short analysis window.

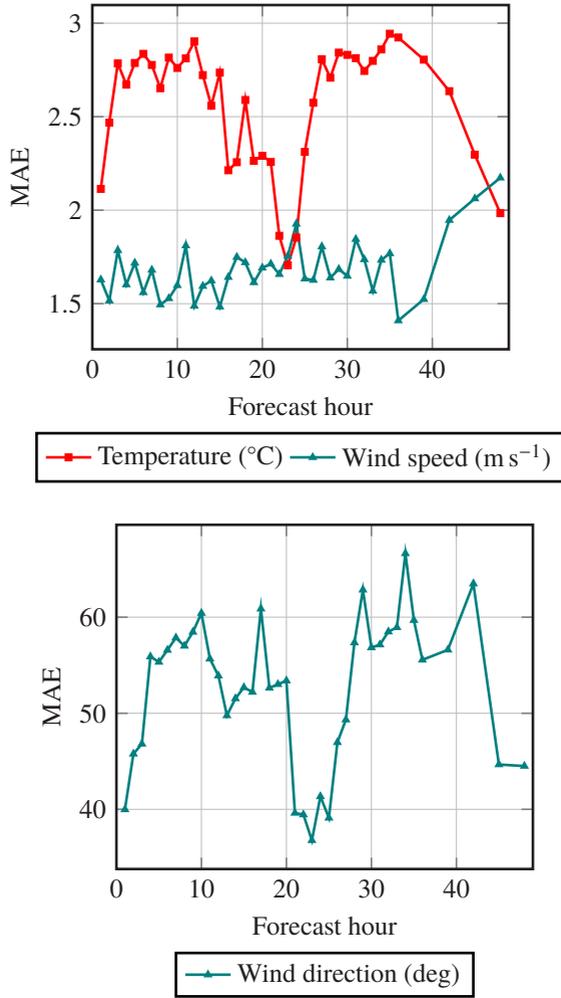


Figure 6.4: MAE for 48 hours NWP in the INL-Mesonet region. Verification done from nearest neighbor NWP grid point to weather station 690.

6.3 Virtual overhead line

The case study of this dissertation recreates historically, over several years, a distributed weather-sensor network along an overhead line as main building block of the PrognoNetz forecasting system. A realistic case study has to meet at least the following three criteria:

- The line has to be shorter than 80 km to remain under the thermal limit. If the line is longer, then the limit for the power transmission is given by the voltage drop limit (from 80 to 320 km) or by stability limits (lines longer than 320 km). These are not related to the weather conditions and have to be controlled using other methods out of the scope of this dissertation [20].
- Changes in the line direction can be up to 90°.
- The number of weather stations going through the overhead line has to be maximized.

Four weather stations from the INL-Mesonet were selected to build a virtual overhead line, which meet these criteria. The result corresponds to a 48.6 km long line, as shown in Figure 6.5, going along the weather stations RWM, 690, CIT, and ROV. The conductor type considered for the current-carrying capacity calculation is a Drake 26/7 ACSR (outer diameter 28.1 mm, steel core diameter 10.4 mm) with emissivity 0.8, solar absorptivity 0.8 and maximum allowable conductor temperature 100 °C. These characteristics allow the direct validation of the current-carrying capacity calculations with the Cigre Standard 601 [9].

6.4 The PrognoNetz dataset

The PrognoNetz dataset² contains historical weather observations and forecasts, the conductor properties, and the current-carrying capacities for the corresponding weather conditions along the virtual overhead line explained in

² Available open-source at github.com/prognonetz/benchmark_idaho



Figure 6.5: NWP grid at the INL Mesonet Region and virtual overhead line. The orange and blue symbols correspond to INL weather stations. The connecting line marks the route of the virtual overhead line. The blue dots represent a portion of the NWP grid, which actually extends to the complete U.S. territory.

Section 6.3. The data has been formatted to train and test machine learning models.

The INL Mesonet Report [79] explains that corrupt or suspect values are often marked using so-called quality flags. That occurs when the sensors have been temporarily disconnected because of maintenance or system faults. These elements are usually indicated as -999 and have to be cleaned. For this study, the corresponding timestamp has been deleted.

In meteorology, the wind is originally arranged in polar coordinates, i.e., wind speed (m s^{-1}) and direction ($^{\circ}$). This arrangement is unfavorable for machine learning models. The main reason is the wind direction discontinuity from 359° to 0° . The solution to this problem is the transformation from polar to Cartesian coordinates. The coordinates system can be oriented north-south, east-west, or rotated towards the first and second principal components

of the wind observations (using the Principal Component Analysis method, PCA). The latter method gives importance to the most common wind patterns. Therefore, it has shown to produce a performance improvement in the machine learning models based on weather observations (see Section 7).

On the other hand, a PCA rotation for the NWP of wind lacks of sense. Each grid point shows their own wind pattern, which means a separate rotation of each grid point is necessary. However, if the same model analyzes several NWP points, the coordinate system has to be consistent throughout all variables. Therefore, the polar coordinate system of the NWP has been converted to a Cartesian system, with components north-south, east-west, from now on called directions U and V, respectively.

The pressure gradient is one of the driving forces, which produces the wind. They were included when developing the machine learning models for the current-carrying capacity forecast based on observational data. They did improve the overall accuracy of the models (see Section 7.2.1) and are considered in the PrognNetz dataset. The pressure gradients were calculated at the position of each weather station along the virtual overhead line. The gradient was obtained from an interpolation of pressure observations from 13 weather stations, at distances ranging from 80 to 950 km from the line.

In the end, the observations dataset is organized as a three-dimensional tensor, with the size (51840, 4, 8) and organized as follows:

- **Timestamp, t_0 :** it represents the real-time in the system. The timestamps were matched to the available NWP dataset and cover the period January 1st, 2009, until December 31st, 2014. There are a total of 51840 timestamps.
- **Weather stations:** the dataset contains measurements for the weather stations located along the virtual overhead line, i.e., RWM, 690, CIT, and ROV (see map in Section 6.3).
- **Weather parameters:** ambient temperature, wind speed, wind direction, solar radiation, pressure, relative humidity, and pressure gradients at each weather station. The data loader transforms the wind from polar coordinates to Cartesian coordinates towards the PCA (see Section 6.5).

For training and a correct evaluation, the NWP dataset has to have the same timestamps (t_0) as the weather observations. The CONUS forecast step corresponds to three hours for the period 2009 to 2014. The INL-Mesonet dataset contains measurements for every hour. The forecast updates always occur at 00:00 UTC and at least once more after 10 to 12 hours. That means that until there is no update to the meteorological model, the NWP are just shifted in time. For example, the third forecast hour at this moment will be the second forecast hour in one hour from now, as the diagram in Figure 6.6 shows.

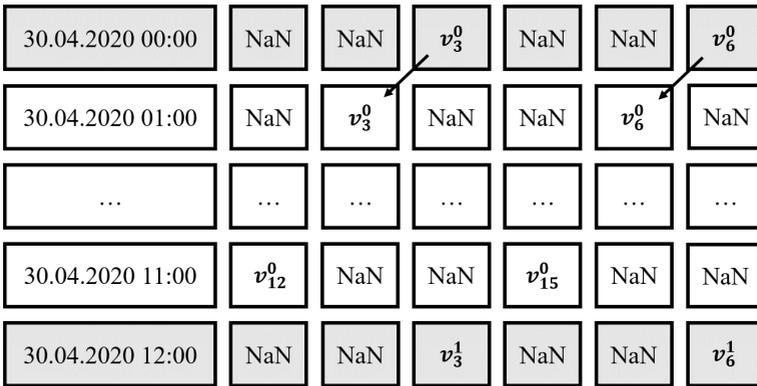


Figure 6.6: Example diagram of temporal alignment of the NWP dataset. The superscripts are related to the forecast update, which occurs every 10 to 12 hours. The subscript corresponds to the original forecast hour. The gray rows mark an update of the NWP model.

The inconsistency in the forecast update rate suggests a careful analysis. A prediction age tensor has been developed to provide additional information to the models. The NWP updates coincide for every weather parameter and every grid point, which reduces the tensor dimension to two. Therefore, it is called from now on, the age matrix, M . It contains the number of timestamps as rows (hourly resolution, from 2009 to 2014, a total of 51840 elements) and the number of forecast elements as columns (equal to the 48 prediction hours).

The age matrix has values from zero to one, with one for a just updated forecast, zero for NaN elements, and the numbers in between depending on the forecast age. The corresponding entries, $m_{i,j}$, are given by Equation (6.1), where n

represents the forecast scope counted from the prediction update time, i the row number, and j the column number. For forecasts older than 48 hours, the value of $m_{i,j}$ is equal to zero.

$$m_{i,j} = 1 - \frac{n - 1}{48} \quad (6.1)$$

The NWP dataset is organized as a five-dimensional tensor, with the final size (51840, 48, 16, 14, 4).

- Timestamp, t_0 : the same parameter as by the weather observations with 51840 elements. The forecasts (t_1, \dots, t_{48}) are referenced to this point.
- Prediction hours, from t_1 to t_{48} : corresponding to a total of 48 elements.
- Horizontal space, (x,y) : two dimensions, describing the spatial position of the NWP grid points as geographic coordinates. The PrognoNetz dataset contains at the moment of writing of this dissertation 16 by 14 NWP grid points in total.
- Weather parameters and age matrix: This dimension contains the wind coordinates U, V, and the temperature. The age matrix is included in this dimension as well, representing the age of each forecast from the last update.

The current-carrying capacity has been calculated for all timestamps at each line section of the virtual overhead line and included in the PrognoNetz dataset. These values are mostly used as labels during the training process of the machine learning models.

The line sections are defined as straight lines between the four weather stations, RWM, 690, CIT, and ROV. Each section has a bearing with respect to the north in degrees, which is included in the calculation of the current-carrying capacity. This value is used to find the strength of the effective wind speed³.

³ Effective wind speed: corresponding wind speed, which produces the same thermal effects on the conductor as the wind with 90° direction with respect to the line. In other words, it is the projection of the wind to the perpendicular of the conductor.

The current-carrying capacity is calculated from the measurements of the weather station at each extreme of the line section. It results in a total of six capacities (three sections, each with two weather stations at the extremes). The actual capacity of the line at a given timestamp is the minimum value between all six. However, for model evaluation purposes, the dataset maintains the capacities per line section as separate features.

The dataset for the current-carrying capacity is organized in a four-dimensional tensor, resulting in a total of (51840, 6, 1, 1) elements.

- Timestamp, t_0 : the same parameter as by the weather observations and forecasts, with 51840 elements.
- Line section: straight lines between the four weather stations, RWM, 690, CIT, and ROV, giving a total of three sections. Each is considered twice since the capacity is calculated at each extreme of the line.
- Line bearing: Angular direction of the line section with respect to the north. For example, zero degrees corresponds to a line north-south.
- Current-carrying capacity: result of the thermal equation in Amperes.

As a summary, the PrognNetz dataset contains three parts. A weather observation dataset with the dimensions (51840, 4, 8), an NWP dataset with (51840, 48, 16, 14, 4) elements, and a current-carrying capacity dataset consisting of (51840, 6, 1, 1) values. All of them covering the same period of timestamps and for the virtual overhead line of the case study. They work as the foundation of the data loader, which is a tool to split the dataset into training, validation, and test sets as needed when modeling (explained in Section 6.5).

6.5 Software development framework and data loader

When developing a benchmark, a single case study and a consistent evaluation metric have to be considered. A development framework helps the machine learning developers, for example, to have a set of standardized tools, to manage the training, validation, and test processes. The framework has access to the

dataset and is able to run pre-processing tasks. Hyperparameter optimization techniques are also available. As a summary, the framework allows a soft execution of the A³ steps from the QUA³CK process [120]: Algorithm selection and training, data Adaption, and hyperparameter Adjustment. Such a framework was implemented during this dissertation specifically for the development of the PrognoNetz models.

The data loader works as a bridge between the PrognoNetz dataset and the software development framework. The dataset is available in the cloud, to make easier the use of the framework from different researchers. The data loader downloads the last version of the dataset, allowing the possibility to keep it up-to-date for modeling. It also presents the option to load a smaller version of the dataset. That is very helpful, especially while prototyping since the running times have to be short and the results do not need to be perfect, while still obtaining an idea of possible model improvements.

The development framework helps explore different machine learning algorithms as well as training features. Therefore, the data loader allows the user to build training and test sets, splitting the timestamps in a 2:1 proportion, with observations, weather predictions, or both.

The PrognoNetz models forecast 48 hours of the current-carrying capacity of a line section, at the location of the weather station. The ground truth, or labels, corresponding to the predictions can be calculated from the observations at that position. For training and test purposes, these datasets are built from historical databases, which makes possible the calculation of the actual values in future timestamps.

The labels dataset is organized as a four-dimensional tensor. It includes the timestamps (t_0), the line sections (considered twice, since the calculation of the current-carrying capacity is done for the two weather stations at the extremes of the line), the bearing of the line, and the 48 corresponding capacities from t_1 until t_{48} . The result is a tensor of size (51840, 6, 1, 48).

The data loader also allows the user to execute pre-processing functions before loading and splitting the dataset. These include the coordinates transformation of the wind (from polar to Cartesian) and the normalization of the data. The latter is done using the Min-Max method, which converts the maximum value in the dataset to one and the minimum to zero.

7 The PrognoNetz Machine Learning Models

The PrognoNetz models have two different information sources available, the weather observations distributed along a virtual overhead line and the corresponding numerical weather predictions in its surroundings. The hypothesis is that machine learning models can extract the information intrinsic into weather measurements and numerical weather predictions, to obtain a reliable current-carrying capacity forecast.

The utilization of appropriate pre-processing methods and machine learning algorithms is one of the key points to achieve appropriate results. That is a problem that requires a structured framework, a well-defined case study, and a proper division of the problem into simpler ones.

The weather observations and predictions correspond to different datasets, both available along the studied virtual overhead line. They have different temporal coverage and data formats. Therefore, they were considered separately in the feature engineering and modeling process, to extract the most of each dataset. The diagram in Figure 7.1 shows the sequence of steps followed in this stage of the project. First, one of the two available datasets is selected for the process. Then, the feature engineering step gives information about the feature set, which improves the performance of the current-carrying capacity prediction. Different combinations of weather parameters, weather stations or NWP grid points, and different data formatting methods were considered.

The following sections present only machine learning models, which directly forecast the current-carrying capacity, instead of those producing a weather prediction to calculate indirectly the ampacity. The former achieved the best performance, as shown in the studies [130], [134], [135]. A reason for this accuracy discrepancy relies on the non-linear transformation of the weather forecast errors due to the thermal equations of overhead lines (described in Section 3.1). Therefore, the models presented here optimize directly the

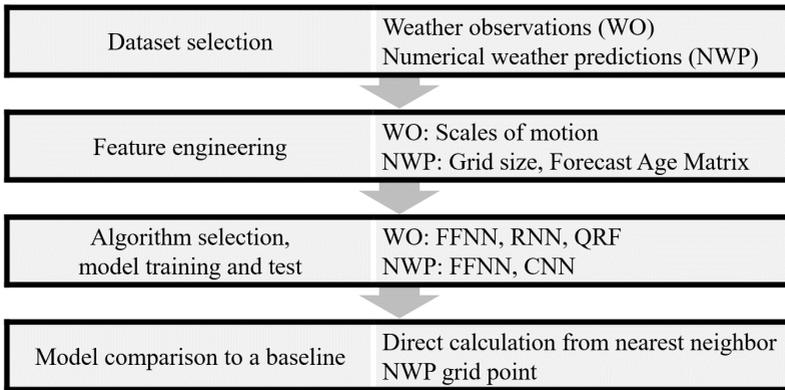


Figure 7.1: Process description for performance improvement of a machine learning forecast model. The dataset is either the weather observations or the numerical weather predictions. They were considered separately to extract the most of each dataset.

prediction accuracy of the current carrying capacity, i.e., 48 hours capacity forecast as output parameter.

After the identification of a suitable feature set, machine learning algorithms are tested and, after benchmarking, one is selected. The procedure corresponds to a mixture of domain knowledge and experimentation. As well as for the feature engineering step, this study is executed for each dataset (weather observations and NWP) separately. The reason is the reduction of the degrees of freedom and easier identification of the approaches, which lead to performance improvements.

Finally, once suitable solutions for each dataset have been selected, the fusion of the knowledge acquired in the last steps takes place. Section 7.4 explains how the final model for the PrognoNetz forecasting system has been created.

The following sections consider the line section between the stations 690 and CIT of the case study (see virtual overhead line in Figure 6.5). The ground truth for the current-carrying capacity is calculated from the weather observations at the station 690. The predictions are all considered in a scope of 48 hours.

The training and test processes were performed on the supercomputer ForHLR funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research of Germany.

7.1 Baseline model

The baseline model provides a reference point to compare the results from the machine learning modeling process. It represents the realistic case of a direct calculation of the current-carrying capacity from the numerical weather predictions. The baseline regards the meteorological predictions at the nearest NWP grid point to the weather station under study, i.e., the sensor node 690. The distance between the nearest neighbor NWP to the weather station is around 2.5 km.

The NWP model used in this case study (CONUS 5) provides temperature and wind predictions. However, the calculation of the current-carrying capacity also needs solar radiation. A realistic and simple way to solve this is applying a persistence forecast of solar radiation. The American Meteorological Society defines persistence as a simple method, at which "the future weather condition will be the same as the present condition" [86]. In this study, the last 48 hours of solar radiation observations are extended as the prediction of the next two days.

Figure 7.2 shows the MAPE for the current-carrying capacity forecast calculated from the nearest neighbor NWP grid point. The accuracy is better during the night hours, compared to the daytime, considering the forecast update once a day at 00 UTC. The variations are as high as 12% mean error. This can be explained by the fact that the sun, the most influencing factor of different weather conditions, is not present during the night with a probability of one.

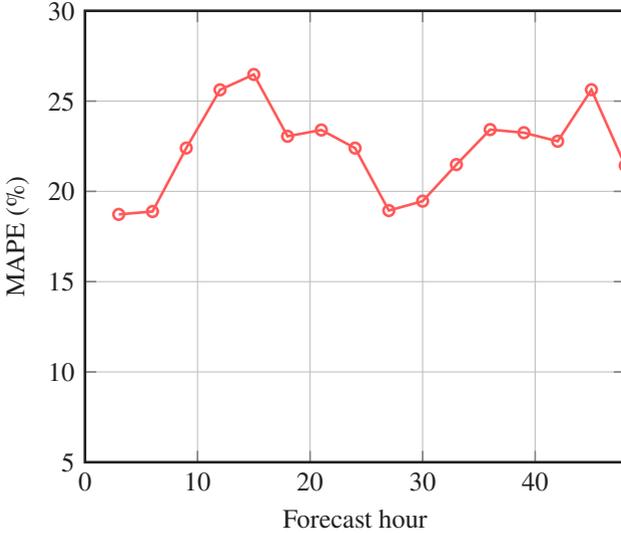


Figure 7.2: Accuracy plot of NWP-nearest neighbor model. The current-carrying capacity is calculated directly from the weather predictions.

7.2 Models based on weather observations

The experiments performed in this phase exhibit the possibility to forecast the current-carrying capacity of an overhead line based only on weather observations while improving the prediction accuracy over the baseline model. The first exploration step has as main goal the determination of a suitable feature set (Section 7.2.1). The experiments were carried out, maintaining the same machine learning algorithm throughout the process. Based on theoretical knowledge, GRUs from the family of the RNNs were selected to execute the feature engineering phase. The reason is based on the capacity of this algorithm to describe sequences of data using recurrent connections in the neural network (see Section 4.2.2).

Once the feature set of weather observations was determined, machine learning algorithms with known forecasting capabilities were explored: Feedforward Neural Networks, Long-Short Term Memory Networks (LSTM), and Quan-

tile Regression Forests (QRF). As a result, LSTMs and QRFs were the best performing models. Both showed a yearly median forecast MAPE of less than 18% for the 48th forecast hour, which is an improvement over the baseline model.

QRFs had the advantage of being trained with one year of historical data, compared to the three years needed for LSTMs [142]. Therefore, QRFs are considered from the practical perspective more suitable for the PrognONetz system than the LSTM approach. Section 7.2.2 describes the former in detail.

7.2.1 Feature engineering

The reduction of the set of input features to its most essential components, maximizing their information density, is crucial to improve the performance of the machine learning model, given a limited amount of training data. This process has been done based on a set of experiments using Recurrent Neural Networks (RNN), specifically Gated-Recurrent-Units (GRU). RNN can learn long-term dependencies of the input variables. The network ends up having memory, which facilitates the modeling of time series (see Section 4.2.2 for the theoretical explanation).

The experiments were divided to test different meteorological scales of motion, i.e., the micro-scale covering a radius of 10 m to 100 m, the mesoscale from 100 m to 1000 m, and the synoptic-scale over the 1000 m radius [87]. For each one of these feature sets, the GRU is optimized for a 48 hours forecast of the current-carrying capacity. The diagram in Figure 7.3 illustrates the standard model used for feature engineering.

Table 7.1 shows a summary of the change in MAPE at each experiment compared to a reference model. The reference consists on a GRU model trained on the normalized ambient temperature, wind vector (rotated using the PCA), solar radiation, barometric pressure, and relative humidity at the station 690. In the following experiments, new features were added to this set. The reference model resulted in a MAPE of 19.22%.

- **Experiments in the microscale:** This case considers the vertical inversion of the temperature gradient, one of the main factors in the change of wind direction at night. This factor consists of the change in ambient

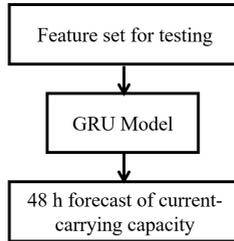


Figure 7.3: Feature engineering standard model. The input dataset changes depending on the feature set to be tested. The GRU hyperparameters are optimized at each experiment, considering in all cases the 48 hours forecast of the current-carrying capacity as output of the model.

Table 7.1: Comparative table for feature engineering

Experiment	Change in MAPE
Microscale	+0.05%
Mesoscale	-0.12%
Synoptic scale	-0.29%
Combination best models	-0.63%

temperature with respect to height, and it is called the adiabatic lapse rate [79]. This parameter was added to the input set used in the reference model. The experiment shows a worsening of 0.05% of the MAPE.

- Experiments in the mesoscale:** Observations from the weather parameters from all four weather stations along the overhead line were added to the feature set. However, this experiment resulted in an inferior performance compared to the reference (+0.29% MAPE). Therefore, a different approach for the mesoscale was developed. Wind and pressure data from the auxiliary sensor nodes along the virtual overhead line, RWM, CIT, and ROV, were added to the weather observation of the central weather station, the node 690. Wind data was included to stimulate wind pattern recognition. Pressure data was added due to the direct correlation between pressure changes and wind conditions. The addition of pressure measurements from auxiliary weather stations did

not perform well on the MAPE. The addition of only wind measurements from the auxiliary stations, on the other hand, brought MAE and MAPE improvements on the current-carrying capacity forecast.

- **Experiments in the synoptic-scale:** These experiments were focused on using remote weather stations for feature extraction. Observing processes at the synoptic scale is a key part of understanding weather conditions [87]. A total of 13 weather stations, at distances ranging from 80 to 950 km from the primary sensor node, 690, were used to reconstruct the distribution of air pressure surrounding the local weather stations using interpolation. This information was employed to derive the geostrophic wind approximation¹, which can be calculated using the pressure gradient (see Figure 7.4). The addition of this variable to the input set resulted in a mean improvement of 0.29% over the baseline MAPE.
- **Combination of mesoscale and synoptic-scale:** Adding wind measurements from the auxiliary weather stations along the virtual overhead line, together with the geostrophic wind approximation, has been the base for the last experiment for feature engineering. Both produced improvements in the error levels of the current-carrying capacity forecast. Their combination resulted, as expected, in an even greater accuracy improvement of 0.63% over the baseline MAPE.

As a summary, the best feature set for current-carrying capacity forecasting consists of weather measurements from the weather station under study (station 690), plus wind observations (rotated towards their PCA) of the auxiliary weather stations along the overhead line (RWM, CIT, and ROV), plus the geostrophic wind approximation at the location of the main weather station. The MAPE of the GRU with the combination input set was 18.59%. Since the accuracy improvement over the reference model is small, reproducibility checks were executed with a new random initialization of the training process. Afterwards, the reference and combination input datasets were tested with the QRF algorithm, which consistently proved, that the latter dataset actually helps

¹ The geostrophic wind approximation is a theoretical wind vector, which magnitude is in m/s and is directed parallel to the isobars (lines of constant pressure). This vector is generated by calculating the Coriolis force and the horizontal pressure gradient force for a given location [88], [87].

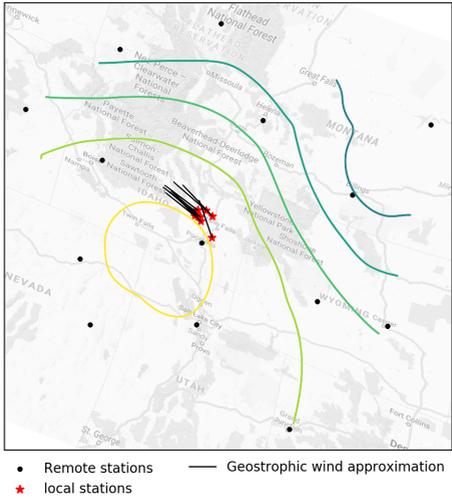


Figure 7.4: Geostrophic Wind Approximation at the INL-Mesonet area [130].

the machine to forecast the current-carrying capacity. Therefore, the following section only considers the combination input dataset.

7.2.2 Quantile Regression Forests

Within different machine learning approaches, QRFs showed together with LSTMs the best median accuracy. They have the additional practical advantage of being trained with a year of historical data, compared to the three years needed for LSTMs [142].

QRFs are, in general, very expensive in training time and memory. This algorithm constructs a forest of randomly and independently generated trees. Each of them considers a subset of all samples. In the end, the best performing tree is used as the final model (see Section 4.2.4 for the theoretical details). Therefore, the trained model is relatively lightweight, but the modeling process

is resource-expensive. That was the original reason to consider seasonal forecasts. Increasing the amount of training data experimentally from 3 to 6 months meant an increment of the training time from 5 hours to more than three days. Therefore, the weather measurements were divided into four seasons, training four models to forecast a year.

The QRF input-output structure is shown in Figure 7.5. Past weather observations enter the model. The QRF then forecasts the current-carrying capacity directly, obtaining 48 models, one for each forecast hour. The actual capacity values are calculated from the actual weather measurements along the overhead line, which are already integrated into the PrognNetz dataset (as explained in Section 6.4).

In the training and test processes, the past and the future are relative. In both cases, a sliding window is used. It has as center the present time, t_0 , and from there, the past and future are defined. In reality, t_0 corresponds to real-time.

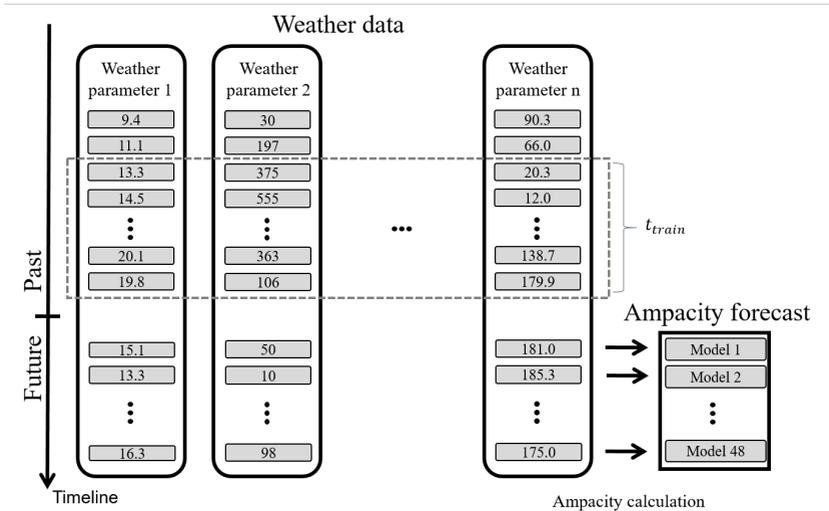


Figure 7.5: QRF input-output structure.

The number of past time steps, t_{train} , to consider for each input weather parameter was optimized. This value can range from one, i.e., only the current observation; to the entire measurement history. As Figure 7.6 shows, the QRF

Mean Absolute Error (MAE) for a 24-hour current-carrying capacity forecast decreases for greater t_{train} values. It reaches a plateau at 20 hours of past input observations. Therefore, the following QRF models were trained based on a t_{train} equal to 20 hours.

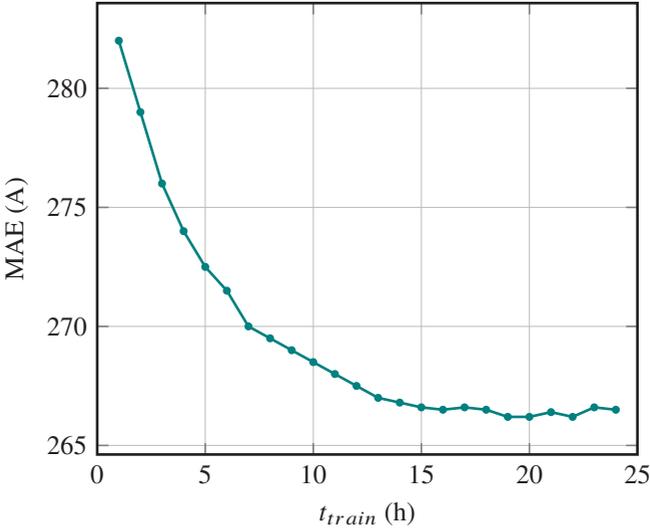


Figure 7.6: Accuracy plot from experiments at different amounts of past weather observations

The input weather parameters were, as discussed in the feature engineering Section (7.2.1), observations at the weather station 690 (the temperature at 15 m height, solar radiation and wind), wind observations from the auxiliary weather stations covering the virtual overhead line (RWM, CIT, and ROV), and the pressure gradients at the station 690. In all cases, the wind has been rotated towards its principal components. The training set contains three months of weather observations, and the model is optimized to forecast within the same period.

Figure 7.7 shows the MAPE for the seasonal QRFs using 0.5-quantile, for comparative reasons with the other machine learning algorithms. The training set covers the year 2007, while the test was carried on three years, from 2014 to 2016. The figure also shows a yearly MAPE, which is the result of calculating the error of the concatenated models. For comparison, the MAPE

of the baseline model was added to the graph. The accuracy improvement is stronger in the first eight hours, which is the most meaningful period for TSOs to guarantee the safety of the grid.

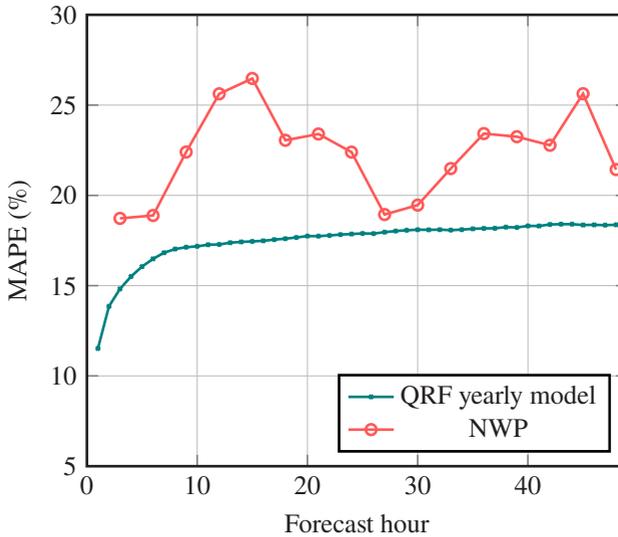


Figure 7.7: MAPE for Quantile Regression Forests, based on weather measurements. Results compared to the baseline model.

The models were programmed using the R programming language and the functions from the *quantreg* package. The training process of a 48 hours forecast model took around 5 hours, which means 20 training hours for the four seasonal models. The test process over three months of data needed around 25 minutes and a single 48 hours prediction lasts for around 500 ms, which are acceptable running times for TSOs. The programs ran over an Intel Xeon Processor E5-2660 v3². A summary of the final QRF model is presented in the Table 7.2.

² Parallel High Performance Computing Cluster ForHLR II, access through the Steinbuch Centre for Computing at the KIT.

Table 7.2: Summary of key features of final QRF model based on weather observations.

	QRF
Input time span	Measurements from the past 20 hours
Input features from main weather station	Wind rotated towards PCA Ambient temperature Solar radiation Pressure gradient
Input features from/at auxiliary weather stations	Wind rotated towards PCA
Output	48 hours prediction of current-carrying capacity
Programming language	R
Training time per model	About 5 hours
Testing execution time	25 minutes for three years of data
Hardware	Intel Xeon Processor E5-2660 v3

7.3 Models based on numerical weather predictions

In this section, the performance of current-carrying capacity forecast models relying only on NWP data is evaluated. Theoretically, the number of grid points to be considered in the calculation of the current-carrying capacity forecast can be extended to the whole earth, including in this way, global information about the weather patterns. However, the number of observations influences the calculation time (training and evaluation) directly and, of course, the memory space required by the models. This section analyzes the number of NWP grid points, which are necessary to reduce the forecasting error of the current-carrying capacity of an overhead line while maintaining the complexity of the problem manageable.

The goal of the study described in this section is to determine if the integration of several NWP grid points in the current-carrying capacity forecast is

beneficial for the prediction accuracy. For this, an algorithm, which can work with distributed information over many dimensions, is necessary. Therefore, Convolutional Neural Networks (CNN) play a central role in this chapter. They are well-known for analyzing structures in n-dimensional datasets [61].

Two different grid sizes have been considered to analyze the influence of the spatial distribution of NWP with CNN, i.e., a small mesh of 4 by 4 points and other of 16 by 14 grid points (equivalent to an area near to 75 km by 65 km), as shown in Figure 7.8.



Figure 7.8: Map of NWP grid sizes for evaluation with CNN. Small mesh grid of 4 by 4 grid points and the original mesh grid of 16 by 14. Both including the weather station 690 as ground truth.

The following subsections describe the models, which were developed to study the influence of the NWP on the current-carrying capacity forecast. The analysis utilizes the NWP dataset, as explained in Section 6.4. The models are organized by the number of grid points they consider and the machine learning technique. First, a single grid point, the nearest neighbor to the station 690, is examined using a Feedforward Neural Network. It models the NWP at the

location of the weather station to minimize the error of the current-carrying capacity forecast. Second, Convolutional Neural Networks (CNNs) take as input a mesh of 4 by 4, or 16 by 14. Both approaches are compared. Moreover, the analysis also considers a CNN, which takes as input a prediction age matrix. The models are compared to the direct calculation from the nearest NWP grid point without any spatial adjustment (baseline experiment).

For training the machine learning models, the NWP dataset has to have the same resolution as the weather observations. The NWP forecast step corresponds to three hours for the period 2009 to 2014. The INL-Mesonet dataset contains measurements for every hour. Three possible approaches, zero filling, backfilling, and linear interpolation, were evaluated to match both resolutions. The results show that there is just a small difference in the influence of each method on the capacity forecasting error. Therefore, the dataset was built on the simplest approach, i.e., the zero-filling.

7.3.1 Feedforward Neural Networks

Feedforward neural networks (FFNN) are well-known for finding non-linear relationships between variables (see Section 4 for the theoretical explanation of neural networks). This experiment uses FFNN with the nearest neighbor NWP point to the position of the weather station 690 as the input set of the model. It is trained by minimizing the current-carrying capacity forecast error. Thus, obtaining an indirect non-linear spatial interpolation of the weather predictions.

The input features to the FFNN are the wind forecast, transformed to Cartesian coordinates (U, V), and the temperature prediction, both for the next 48 hours (a total of 3x48 features). The output corresponds to 48 forecast hours of the current-carrying capacity.

The hyperparameters were optimized using random search³. As a result, the best model ends up having three layers, with 637, 431, and 824 nodes, respectively, with a batch size of 17 and a learning rate of 0.026. The training epochs were fixed because of consistency to 75. The training loss function

³ Random search usually achieves similar results as sophisticated techniques, as the Bayesian optimization algorithm, by lower code complexity. For more information, please refer to [89].

was the MAE. The training time was around ten minutes per model, the code was written in Python, using the Pytorch functions for machine learning, and executed on a NVIDIA GeForce GTX980 Ti graphic card⁴.

Figure 7.9 shows the comparison of MAPE for the FFNN model against the nearest neighbor approach (baseline). There is an improvement of about 10% during the day, while during the night the improvement is around 2%, considering the forecast update hour of the baseline of 00 UTC.

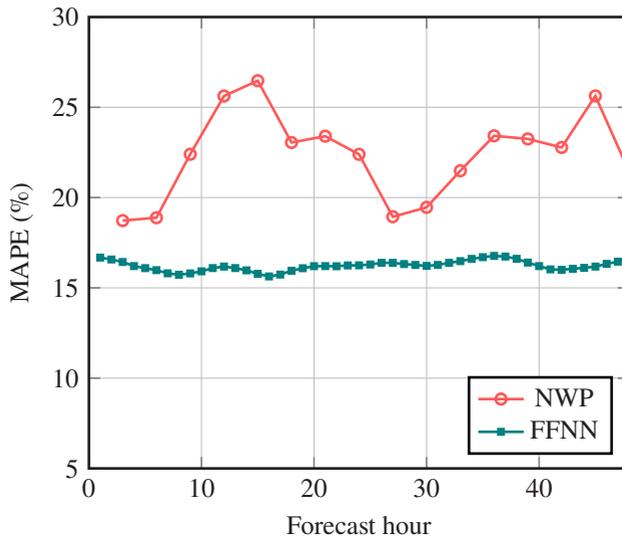


Figure 7.9: Accuracy comparison of FFNN model to the NWP-nearest neighbor approach.

7.3.2 Convolutional Neural Networks

The NWP dataset can also be seen as a set of three-dimensional images. Consider again Figure 4.6 explained in Section 4.2.3. The dimensions (h_1 , h_2) can be seen as the horizontal space (x , y). The values for h_3 , the depth of the image,

⁴ Parallel High Performance Computing Cluster ForHLR II, from the Steinbuch Centre for Computing at the KIT.

and m , the number of 3D-images, can be determined experimentally. There is a difference between these two parameters from the convolution perspective. Depending on the parameter used as image depth, the network extracts different patterns. This assignment is determined in the first experiments described in this section.

Two models have been evaluated, i.e., one with the convolution rolling over the forecast hours and other over the weather prediction parameters. For these experiments, the simplest possible configuration for the CNN was chosen, with two identical convolutional layers, followed by a fully connected layer. Due to the different input variables, different kernel, pooling sizes, and number of channels had to be considered. A single output was chosen for the experimental model. That means that a separate model was trained for each forecast hour. A drop-out of 30% was used to ensure good generalization capabilities, as well as weight-decay. The batch size was set to 64. The learning rate, the parameter of weight-decay, and the training epochs were left constant. They were initially chosen according to experience, to ensure comparability between experiments.

The models were trained and tested in a small dataset to ensure a quick experimenting process. As a result, an improvement is achieved if the image depth is the forecast hours, and each 3D-image corresponds to each weather parameter (NWP wind coordinates (U, V) and temperature). The performance improvement makes sense since the convolutions are extracting patterns of each weather parameter over their structural relationships over time, which is a known phenomenon in meteorology. In the further analysis only the best approach is considered, i.e. $h_3 = 48$ and $m = 3$.

The following experiments lead to the identification of a suitable amount of NWP grid points as input features to the CNN. Two sizes for the horizontal space, the dimensions (h_1, h_2), have been considered, i.e., 4 by 4 and 16 by 14 elements (see Figure 7.8). Both cases centered at the weather station 690. The selection of these sizes are based on two criteria: first, the minimum amount of grid points to achieve a symmetrical arrangement, which also is the smallest grid that can be reasonably processed with a CNN, and second, the coverage of the whole virtual overhead line of this case study.

Figure 7.10 shows a comparison of both models. It seems that the wider the spatial coverage of NWP mesh, the more accurate the current-carrying

capacity forecast gets. That agrees with the original hypothesis and the reason to include CNN in the analysis of this case study.

As explained before, the NWP updates are not always consistent in time. Therefore, a prediction age matrix has been considered as an additional input feature for the next experiment with CNNs. It has values from zero to one, with one for a just updated forecast, zero for NAN elements, and the numbers in between depending on the forecast age. The calculation of each entry follows the Equation (6.1) (Section 6.4).

The reason to include an age matrix into the input features of the CNN model is the information it contains about the prediction importance. In fact, as Figure 7.10 shows, the system is taking advantage of this for the first 24 hours. The forecasts from around 30 to 48 hours do not depend anymore on the prediction age.

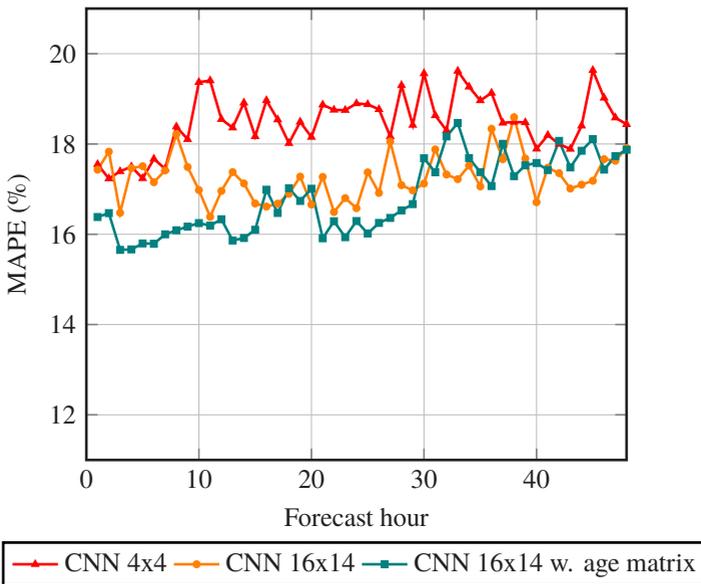


Figure 7.10: MAPE comparison of two CNN models trained for a small coverage (case 4 by 4 grid points) and a bigger spatial area (16 by 14 NWP grid points). A prediction age matrix was included to the best performing CNN, which also improved the MAPE of the model.

A hyperparameter comparison of the three CNN models is shown in Table 7.3. The maximum training epochs were fixed to 75 and the drop-out to 0.1. The training loss function was the MAE. The training time for the CNN models was around one hour per model, the code was written in Python, using the Pytorch functions for machine learning, and executed on a NVIDIA GeForce GTX980 Ti graphic card⁵. The CNN with the age matrix as input has the best prediction accuracy and is considered in further analysis.

Table 7.3: Hyperparameters of CNN models

Hyperparameters	CNN 16x14	CNN 4x4	CNN 16x14 with age matrix
Batch size	128	16	512
Learning rate	0.000125	0.0005	0.000125
Convolutional layers	10	9	5
Channels	2	9	3
Channels last layer	2	4	3
Weight-decay	0.003	0.0003	0.1

7.3.3 Comparison and discussion

Figure 7.11 compares the accuracy of the models based on NWP data, i.e., the baseline, which uses the nearest neighbor NWP grid point directly to calculate the current-carrying capacity forecast; the FFNN model, which interpolates the nearest neighbor NWP to the weather station position; and the best performing CNN model (16 by 14 NWP grid size including a prediction age matrix).

Two main conclusions can be derived from this section. First, machine learning has advantages over the baseline. The accuracy of the current-carrying capacity forecasting could be improved in the day hours in about 10%, compared to the nearest neighbor NWP grid point approach. FFNN outperformed CNN, especially in the last prediction hours (from 28 to 48 hours).

⁵ Parallel High Performance Computing Cluster ForHLR II, from the Steinbuch Centre for Computing at the KIT.

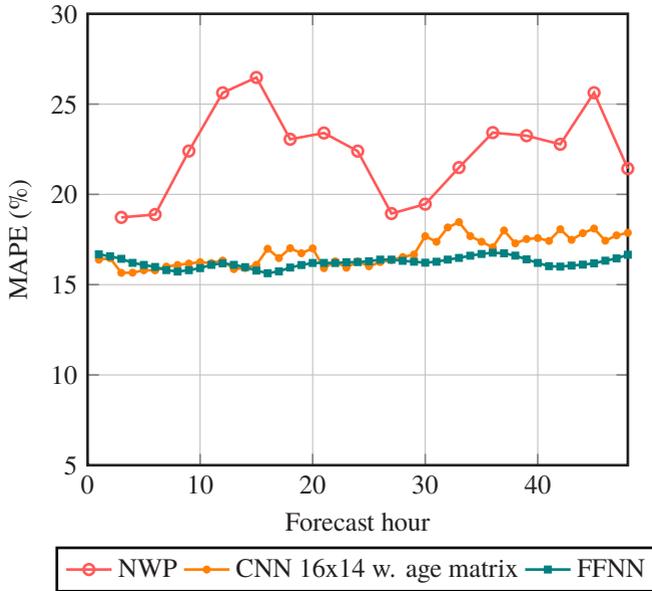


Figure 7.11: Comparison of the prediction models based on NWP.

The second conclusion corresponds to the importance of increasing the grid size when extracting spatial patterns. There was an accuracy improvement when the CNN model transitioned from a grid of 4 by 4 NWP elements to 16 by 14. That means that a further analysis of wider NWP grid sizes could be valuable. However, there is a trade-off between the number of NWP grid points, the run-time, and network complexity, which has to be optimized. These experiments should be considered as outlook of this dissertation.

The following section uses the resulting FFNN model to produce a final current-carrying capacity forecast solution. Table 7.4 shows the summary of the key-features of the FFNN model.

Table 7.4: Summary of key features of final FNN model based on NWP.

	FNN
Input features	48 hours NWP at nearest neighbor
Preprocessing	Wind prediction in cartesian coordinates
Output	48 hours current-carrying capacity prediction
Training loss function	Mean Absolute Error (MAE)
Training time per model	about 10 minutes
Number of models trained for hyperparameter optimization	250 models
Programming language	Python, machine learning functions from Pytorch
Hardware	Four NVIDIA GeForce GTX980 Ti graphic cards

7.4 Ensemble and final solution

Experimenting with models based on historical weather observations separately from those based on NWP data helped to gain knowledge about these different datasets. Weather measurements provide more information if they are considered together with the observations distributed along the overhead line, as well as the geostrophic wind approximation calculated at the synoptic scale. On the other hand, the nearest neighbor NWP grid point offers more information when adjusted to the location of the weather station using FNNs, compared to the consideration of several of them as a mesh.

The QRF model, based on observational data, has a better accuracy for the short-term forecasts, while the FFNN model based on the nearest neighbor NWP performs better for long-term predictions (short and long-term are here referred to 0-10 hours and 11-48 hours, respectively). These strengths can be combined using ensemble methods, creating a new model, which takes advantage of the best prediction of each of the original models.

The combination can be done after the models are trained, a method called stacking. The ensemble model is trained as a supervised learning algorithm, with the QRF and FFNN outputs as input features, and the actual current-carrying capacities as labels.

Stacking can be done using any supervised machine learning algorithm, from a simple average to neural networks. The most simple but still well-performing approach in this case is a multilinear regression [139], [145]. This algorithm generates a weighted average of the inputs. The goal of the training process is to find the proper weights for the combination. The diagram in Figure 7.12 shows the structure of the implemented ensemble approach. Each forecast hour is combined separately to ensure proper model weighting. That results in 48 multilinear regression models. The calculation of the weights lasts around a minute running on the training dataset and the prediction time is negligible.

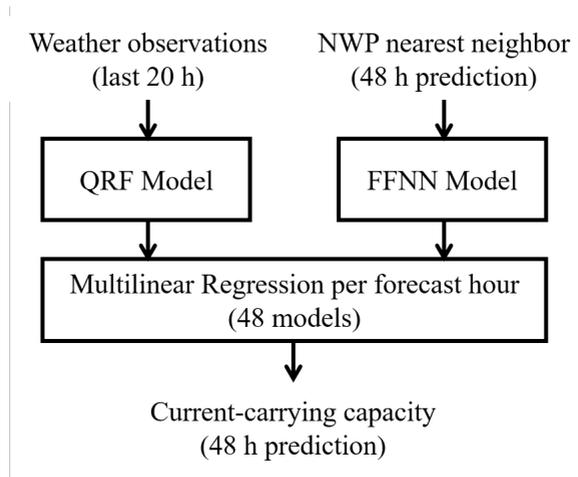


Figure 7.12: Block diagram of the ensemble of the best performing models: QRF from the weather observations and the FFNN from the NWP nearest neighbor.

Figure 7.13 shows the MAPE of the ensemble compared to the original models. It is calculated as the mean absolute percentage error between the predicted ampacity in 48 hours by the combined model and the actual corresponding capacities. The true values were calculated and saved as labels for testing in the PrognoNetz dataset. Based on this metric the combination achieves a

percentage improvement from the 4th to the 10th forecast hour, from 17th to the 21st and at the 36th hour. The error reduction in the first ten hours is beneficial for the system and aligns to the requirements of the TSOs.

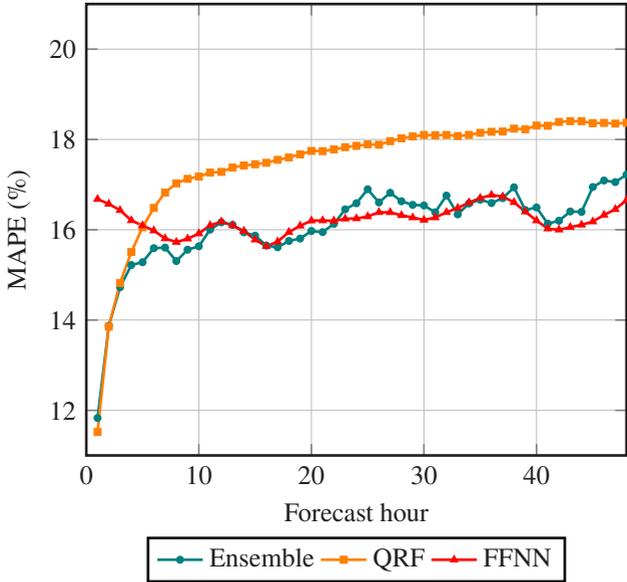


Figure 7.13: MAPE of model fusion compared to the accuracy of the original models.

When analyzing the raw error distribution, the ensemble model also achieves an improvement in the errors dispersion. Given the raw error, calculated as the difference between the predicted and the actual capacities, negative errors correspond to underestimations, and vice versa. Figures 7.14 and 7.15 show the comparison of the box plots for the raw error distributions of the ensemble model, the QRF and the FNN models, at the 1st and 48th forecast hour, respectively. The filled box in the figure represents the region, where the 50% of the errors are located. The whiskers show the maximum and minimum errors obtained in the test set. In both figures, the ensemble presents a higher concentration of the errors around the zero. This corresponds to a greater accuracy and precision of the solution.

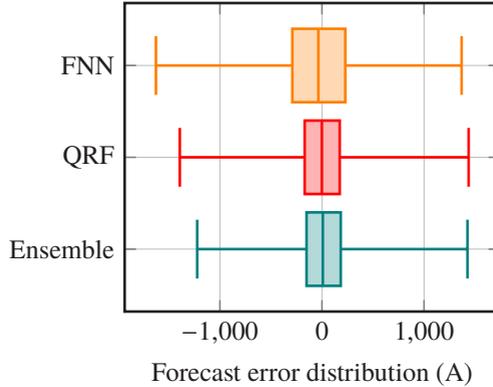


Figure 7.14: Statistics of final model compared to QRF and FNN at the 1st forecast hour. The filled box represents the 50% of the errors and the whiskers the maximum and minimum errors in the test dataset.

7.4.1 Comparison to the baseline and discussion

The final solution for the current-carrying capacity prediction is compared to the baseline model, i.e., the direct capacity calculation from the nearest neighbor numerical weather prediction. The comparison of the mean absolute percentage error (MAPE) is presented in Figure 7.16. The ensemble model shows an average accuracy increment with respect to the baseline (called in the figure as NWP) of 16.19%. The accuracy improvement over the direct calculation from the weather predictions is clearer for the first prediction hours, which fulfills the requirement of the TSOs for a reliable redispatching. The error was reduced in around 10% for the daily hours and 3% to 4% in the night hours, considering the forecast updates at 00 UTC for the baseline.

Figure 7.17 presents the box plot for the raw error distributions (predicted minus actual transmission capacities) of the ensemble model compared to the baseline (NWP) at the 48th forecast hour. The filled box in Figure 7.17 represents the 50% of the errors and the whiskers the maximum and minimum errors in the test dataset. The ensemble model shows a higher concentration of the errors around the zero, i.e., a greater accuracy and precision compared to the baseline.

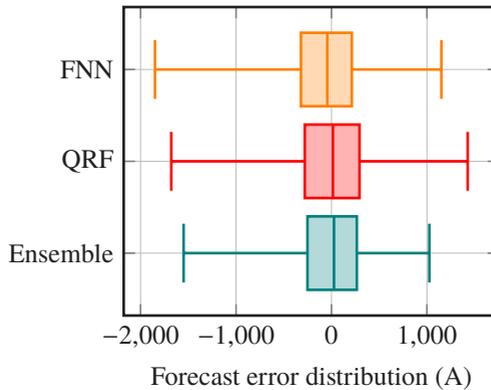


Figure 7.15: Statistics of final model compared to QRF and FNN at the 48th forecast hour. The filled box represents the 50% of the errors and the whiskers the maximum and minimum errors in the test dataset.

The raw error distribution of the ensemble model is symmetric around zero. This behaviour is expected, since all models were trained for high accuracy, which allows benchmarking in the development process. It also means, that the probabilities of over- and underestimation are both around 50%.

7.4.2 Prediction safety factor for system operations

The safety and the life span of the electrical network have the highest priority. System operations need a cooperation of an online monitoring together with the prediction of the current capacity for efficient planning. Real-time measurements of the conductor temperature (or its derivation from other variables, as line length) allow TSOs to correct an overloading of the grid in short-term. Thus, avoiding infringements of the minimum distance of the overhead line to ground or a reduction of the life span of the material because of annealing.

The nearer the current capacity prediction to real-time, the more important overestimations are. Allowing 2% of overrate cases, TSOs still have time to react without producing long-term damages of the line. Since prediction systems always have deviations with respect to the actual values, a safety factor has to be considered.

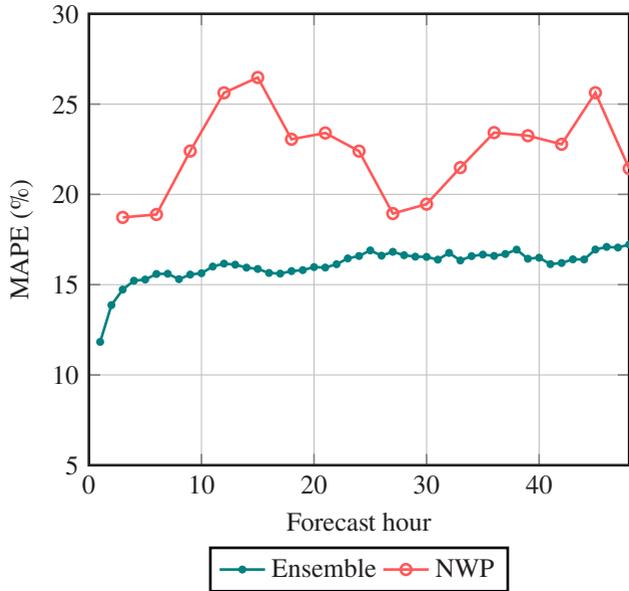


Figure 7.16: MAPE comparison of the ensemble model to the NWP-nearest neighbor approach.

The safety factor for system operations at a maximum of 2% overestimations rate is calculated from the 0.02 quantile of the errors. The final solution of this dissertation has for the third prediction hour a 0.02 quantile of 660 A, compared to the 964 A of the state-of-the-art (direct calculation using numerical weather predictions, without any machine learning). Considering now the 0.02 quantile of the first forecast hour, which is the nearest to real-time that the machine learning solution can deliver, the safety factor is reduced to 550 A.

The impact of the machine learning solution of this dissertation can be analyzed considering a dispatching plan based on a corrected model with 98% reliability. The following calculation is based on the simplified, 5-zones model of the German electrical network from Staudt [27] (see Section 5.1 for a detailed explanation of this model).

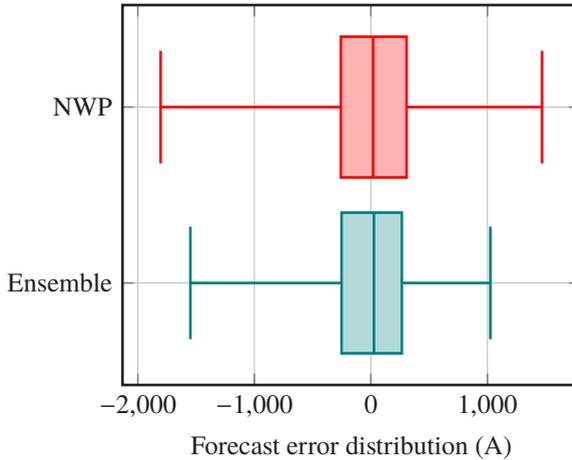


Figure 7.17: Statistics of final model compared to the Numerical Weather Predictions (NWP) at the 48th forecast hour. The error is calculated as (predicted - true). The filled box represents the 50% of the errors and the whiskers the maximum and minimum errors in the test dataset.

The correction of the one-hour prediction consists on subtracting the safety factor from each first hour forecast value. If the new prediction results smaller than the Static-Line-Rating (SLR), it is limited to the corresponding SLR as a minimum. In this scenario, the number of cases with higher current capacity than the SLR is almost 15% of a year.

Table 7.5: Comparison of the German redispatching situation under SLR and DLR with a prediction safety factor.

	SLR	DLR w. prediction safety factor	Improvement
Redispatching (TWh)	69.26	66.33	2.93
Redispatching costs (Mio. Euros)	304	278	26

Table 7.5 results from calculating the redispatching costs for this new setting and comparing them to the SLR case for the year 2018 (compare with Section

5.1). The redispatching improvement of around 3 TWh in a year is the based for the calculation of return-of-investment found in the summary and conclusions of this dissertation.

8 Integration of PrognoNetz into System Operations

This dissertation centers in the development of the machine learning models for the weather-based prediction of the current-carrying capacity of overhead lines. However, these models have to be embedded into a system, which follows the requirements of system operations, for their appropriate application.

8.1 Implementation of a distributed sensor network

The integration of the PrognoNetz system into system operations include the correct selection of the hardware and measuring devices, the improvement of the installation process of the sensor nodes, to achieve it in less time, thus reducing the costs; a reliable and safe wireless communication between the sensor nodes and the base station, and the correct implementation of a server and a database to manage the models accordingly. These topics are outside the scope of this dissertation.

Besides these points, the correct application of PrognoNetz also includes the careful positioning of the weather stations along the overhead lines. The limited budget of many TSOs does not allow to consider PrognoNetz as a system with complete coverage of the electrical network. Instead, the available sensor nodes have to be meticulously located along the overhead lines to cover the spots with the most probable highest conductor temperatures, the so-called *hotspots*. This section presents an algorithm to locate these hotspots as a previous-to-deployment phase of the PrognoNetz system, independent of the region where it is applied.

The current-carrying capacity for each section of an overhead line and for each point in time is calculated based on an interpolation of the weather observations to the electrical towers. The capacities are calculated, and the line section where the minimum occurs is annotated. The number of occurrences of minimum capacities at each line section is then accumulated into a probability distribution function. The histogram in Figure 8.1 shows an example case (for more information about the case study, please refer to [119]). The plot shows regions, which have a high probability of acting as a bottleneck, where the conductor temperature gets the hottest more often. There are also regions where the capacity finds its minimum just on rare occasions. These can be filtered out, considering a threshold that the TSO is prepared to take. In other words, all points along the line, with a probability value lower than a threshold, t , can be neglected from the list of hotspots.

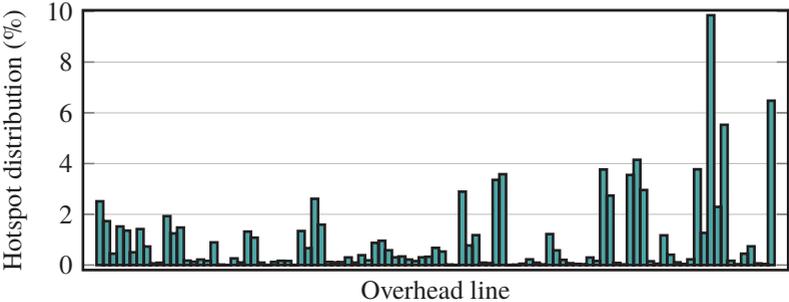


Figure 8.1: Distribution of current-carrying capacity minima along the overhead line. The greater the number of occurrences, the higher the probability that the corresponding line section acts as a capacity bottleneck.

The identification of the filtering threshold, t , is done based on a percentile. The probability distribution along the line is organized from the smallest to highest probability, and the percentiles are extracted, as the example shown in Figure 8.2. The probability of occurrence of the hotspots to the left of the selected threshold (10% percentile in this example) is considered negligible. Therefore, all these elements are rounded to zero, thus leaving only the most probable hotspots.

The non-filtered sections of the line are reorganized back and then clustered. In the example considered before 18 clusters can be identified, as shown in

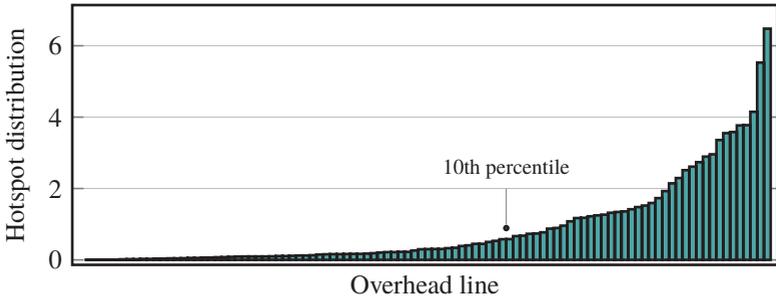


Figure 8.2: Example for the determination of the filtering threshold for the hotspot analysis. In this example, the 10th percentile is shown. All elements to the left of this threshold are rounded to zero. Thus considering only higher probability spots.

Figure 8.3. The procedure consists of assigning a one for all line sections with hotspot probability different to zero. A cluster corresponds to a group of ones bounded by a zero on each side.

The variation of the filtering threshold changes the amount of identified clusters. While t increases, the clusters are divided into parts, thus augmenting the total number of hotspot-areas. At some point, whole clusters begin to disappear. A further increase of t decreases the amount of clusters until there are no recognizable hotspots along the line anymore. This behavior can be seen in Figure 8.4 for the case example. The selection of a reasonable threshold depends on the risk level and the budget the TSO is ready to carry.

Once the TSO has determined a budget, it is transformed into an equivalent amount of weather stations, k . The procedure of cluster identification is executed. Each contains one or more towers per group (possible installation positions), and the k sensor nodes are distributed among them.

The electrical towers selection process gives a higher priority to the positions with greater probability to be a bottleneck, i.e., to be the place of the minimum current-carrying capacity of the line at a point in time. The positions are selected from the highest to the smallest probability without repeating a cluster unless k is higher than the total amount of clusters available. The pseudocode

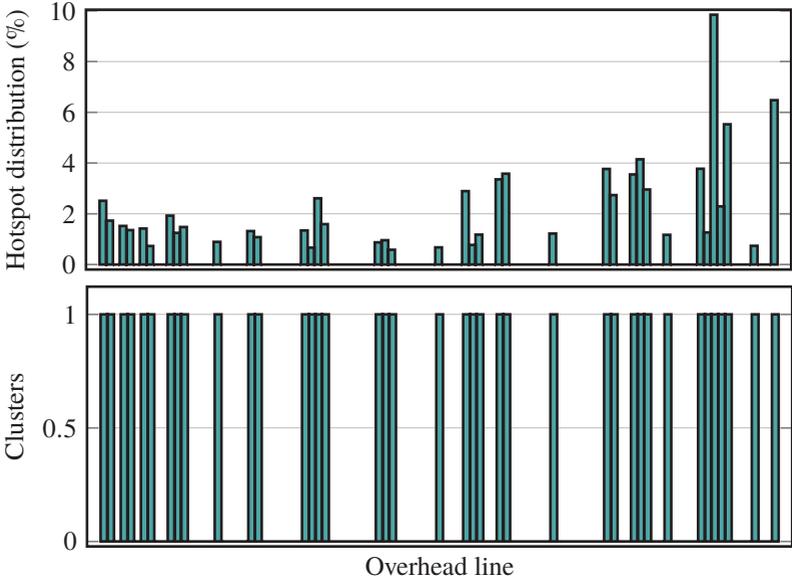


Figure 8.3: Clustering example of hotspots along the overhead line.

for this algorithm can be found below. The process assumes that the number of clusters and available weather stations are known.

Algorithm 1: Weather station positioning

```

Result: Placement of available weather stations at hotspots
Given a number of weather stations to position;
Initialize list of available clusters;
for each cluster do
    Sort available installation positions from highest to lowest
    hotspot probabilities;
end
while still weather stations to position do
    Assign weather station to first installation position from next
    cluster in the list;
    Delete installation position from cluster;
end
    
```

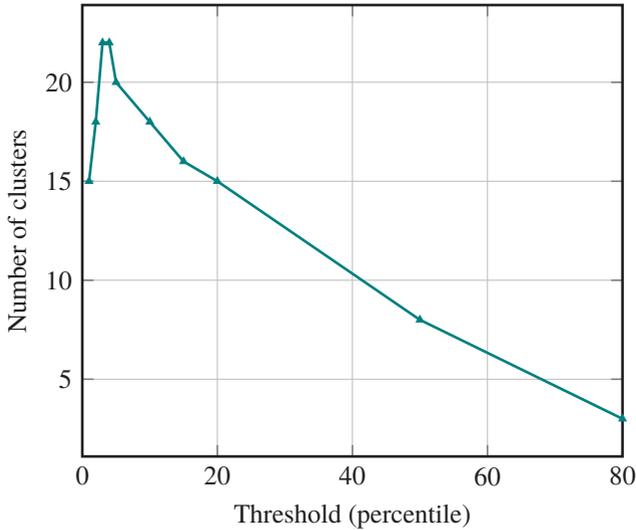


Figure 8.4: Influence of threshold value on the number of hotspots clusters.

The analysis of the current-carrying capacity error determines the minimum number of weather stations, k , an overhead line monitoring system should have to cover the most critical hotspots. This evaluation includes the analysis of the mean capacity error for different k .

It is essential to notice that the current-carrying capacity cannot be underestimated by leaving hotspots out of the monitoring system coverage. In the best-case scenario, the minimum capacity is always measured. Otherwise, the measurement error is always positive (observation greater than the actual value). TSOs have to consider this risk when executing this hotspot analysis.

The results for a clustering threshold of 10% are shown in Figure 8.5 as an example. For a k equal or greater than 20, out of 18 clusters and 40 possible installation positions, the error is near to zero. For a smaller k , the error increases exponentially. That makes sense, because of the amount of information that is lost when fewer clusters are considered. For a line of 45 km length, an amount of 20 weather stations covering the most important hotspots of the line implies an average sensor node density of 2.2 km.

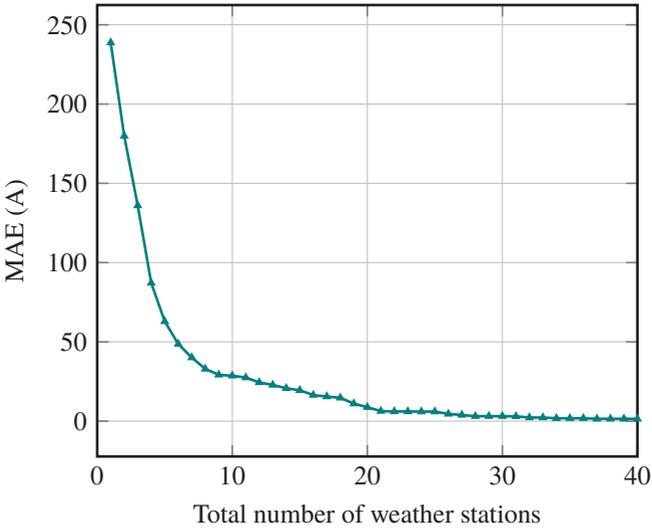


Figure 8.5: Example for mean current-carrying capacity error for different number of weather stations covering the overhead line.

There is an alternative approach to the procedure explained above. Instead of considering only a single line section per timestamp, at which the minimum current-carrying capacity appears, there is also the possibility to include all sections with capacities *near* the bottleneck. The hotspot analysis can be done at different tolerance levels, and by comparing the results, the most convenient solution can be determined.

Moreover, the number of hotspots can be reduced by evaluating those, which are spatially correlated and hold that correlation over time. In other words, the areas along the line, where the maximum temperatures (under a tolerance level) are present at the same time. That would reduce the amount of installation positions since measuring at only one of the spatially correlated hotspots corresponds to measuring several of them.

8.2 Software tool for system operators

The PrognoNetz system additionally offers a tool for active interaction between the user and the information available about the grid, going beyond a simple passive information source. For demonstration, a web-based representation has been developed. Figure 8.6 presents the first view of the website. On the left side, the overhead lines, which are currently under test in PrognoNetz, is shown in a map. In the final version, the whole electrical grid would be represented. The capacity of the overhead lines, are presented on the right side of the site. It is calculated separately for each overhead line, as the minimum between the capacities of each line section. The plot shows the past 48 hours (from actual weather measurements) and the corresponding two days prediction.

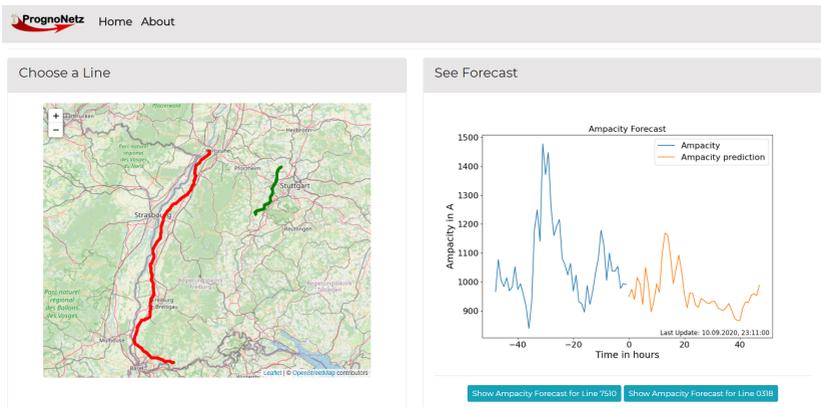


Figure 8.6: PrognoNetz software tool: view of the overhead line and its ampacity.

The second view of the website (Figure 8.7) offers a transient analysis of the system. Given the current weather conditions and the power flowing through the line as a function in time, the transient changes in the conductor temperature can be calculated (see Section 3.1.7). Text boxes have been added for a manual input of the current flowing through the line and the analysis time. The system generates a warning if the calculated conductor temperature exceeds the maximum allowed temperature. It also informs when it could occur.

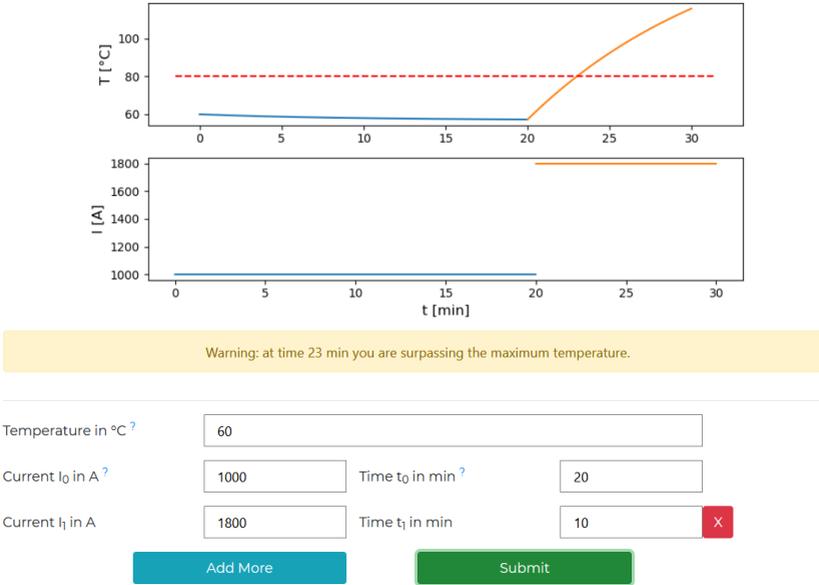


Figure 8.7: PrognoNetz software tool: transient analysis of the conductor temperature.

The possibility to analyze the changes in conductor temperature in short-term, gives the system operator the flexibility to free momentary congestion cases. They can make use of the slow thermal time constant of the conductor (between 1 and 15 min, depending on the conductor type and the weather conditions) and allow the transmission of more power than the actual current-carrying capacity, without overheating the line.

For example, considering there is an ACSR Drake overhead line at a sunny, windy and fresh day (980 W m^{-2} solar radiation, 2 m s^{-1} wind speed perpendicular to the line, $22 \text{ }^\circ\text{C}$ ambient temperature). For simplicity, the weather conditions will be considered constant in the analysis. The line has been driven at its half current-carrying capacity (668.87 A), thus the initial conductor temperature is $38 \text{ }^\circ\text{C}$, as shown in Figure 8.8 in the negative side of the temporal axis.

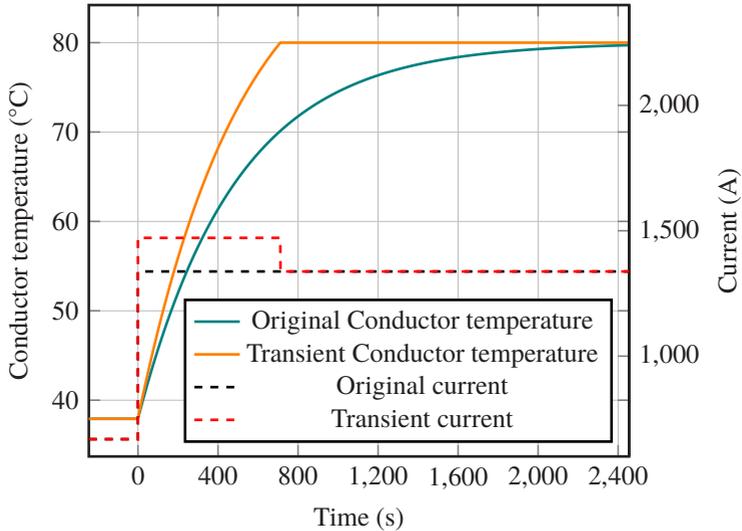


Figure 8.8: Example of a transient management measure.

At some point there is the need to transport more power to cope with congestion. If the current is increased to the maximum transmission capacity, then the maximum conductor temperature of 80 °C will be reached after 40 min (2400 s in Figure 8.8). However, the TSO determines, that the congestion can be solved if the line is driven at 1.1 times the current-carrying capacity for at least 5 min. A transient analysis confirms that this decision is safe, giving more than 11 min (712.54 s) to reach the maximum conductor temperature. In case the TSO decides to manage the line at 80 °C conductor temperature, as shown in Figure 8.8, the energy transmission increases in 2.9%.

The PrognNetz software tool developed in this dissertation also provides a calculator for the time a particular overhead line can transport that extra-power to cope with congestion. The TSO inputs the electrical current that has to be transmitted during a short period and the system tells how long this setting can be operated under current weather conditions. Figure 8.9 shows the corresponding user-interface.

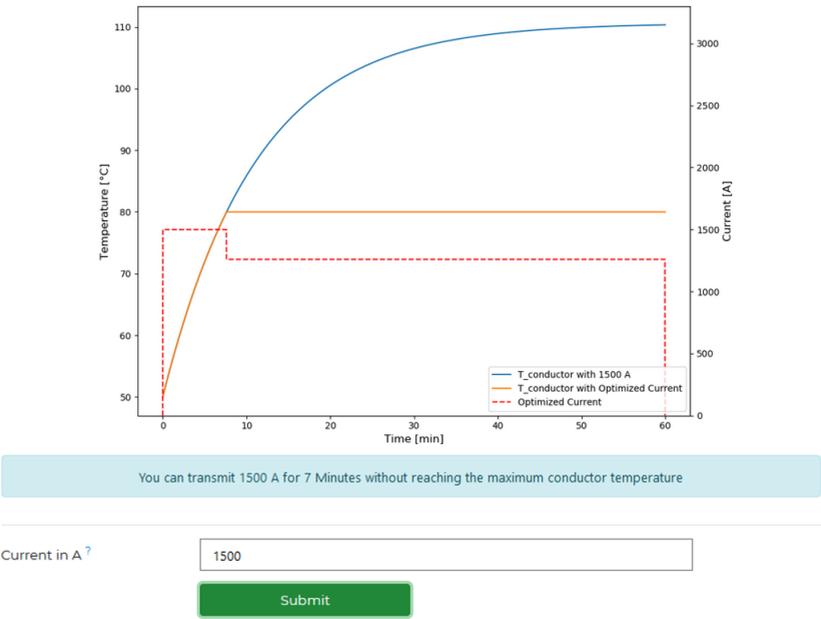


Figure 8.9: PrognNetz software tool: transient management, which calculates how long the TSO can transmit the given electrical current through the line under the current weather conditions.

These short-term reactions are possible if a monitoring system and the corresponding analysis tool are available. Moreover, it is highly recommended to have redundant measurement systems (combination of weather stations and conductor temperature sensors, for example) to guarantee the lifespan of the overhead lines and safety of the system.

9 Conclusion and Outlook

This dissertation presented the development of a tool for predicting the dynamic line rating of the electrical grid based on the local weather conditions along the overhead line routes. The system helps transmission system operators (TSOs) to avoid congestion scenarios, which is a key-point towards the flexible grid necessary to achieve the Energy Transition Plan.

The economic impact of dynamic line rating has been analyzed on a sample overhead line in Germany. The study is based on abstractions of the German electrical network and a generalization of the additional current-carrying capacities of the overhead lines. When these optimization mechanisms are used, the transmission capacity has the potential to be 50% of the time 28% bigger than the static line rating approach. This extra-capacity is translated into the amount of power that could have been transmitted, instead of redispatched or curtailed. Therefore, the use of dynamic line rating can reduce the redispatching measures in 42%, which means a reduction in the congestion management costs of around 55%. Considering a cost of 1438 Million Euros in 2018, the use of these optimization mechanisms have clear benefits.

The exploitation of dynamic line rating systems is possible if the TSOs can use this information to plan one to two-days-ahead the schedules of traditional power plants. The examination of the state-of-the-art shows that most of overhead line monitoring systems focus on real time measurements, which are hard to exploit since the decisions in systems operations are taken some days in advance. The existing systems for the forecasting of the current-carrying capacity are a direct translation from weather predictions, which do not consider the actual local surrounding conditions along the line routes. Therefore, this dissertation focused on the study of weather-based 48 hours forecasting of the current-carrying capacity using machine learning models. It presents a solution, which adjusts the meteorological predictions to the weather conditions along the line, based on weather observations collected in the surroundings of the conductor.

The partitioning of the problem into models based on weather observations and those based on numerical weather predictions showed that the former had a better prediction accuracy in the first hours, while the latter was more appropriate for the last prediction period. An ensemble model resulted in a combination of the strengths of each approach. It had an average accuracy increment with respect to the baseline (direct calculation of the current-carrying capacity from numerical weather predictions) of 6.13%.

The TSOs consider the safety of the electrical network as the highest priority when taking decisions. This means that overloading the grid has to be avoided. The machine learning models presented here were trained to predict with a confidence rate of 50%, i.e., centering the mean error on zero. That leads to having statistically as many overestimations as underestimations of the transmission capacity. In the development phase, especially when comparing different algorithms, this behavior is useful and wanted. For the application of this system in operations, a safety factor is considered to reduce the overestimation probability to 2%. Comparing the machine learning result presented in this dissertation to the state-of-the-art, this safety factor could be improved in about 300 A for the third forecast hour.

The impact of dispatch based on the result of this dissertation was calculated by running a redispatching analysis using this system and compared to the static-line-rating case. The redispatch power was reduced in 3 TWh in a year, when adding the DLR predictions under their safety factor. In the studied simplified scenario, that corresponds to a total of 26 million Euros saved costs.

By defining the efficiency of the power transmission plan as the percentage of the actual current-carrying capacity that is utilized if the plan is carried out, the 48 hours generation plan had an efficiency of 64.7% by planning based on the final model of this dissertation, 53.3% for the corrected baseline and 47.9% for the static line rating. Meaning more than 10% gained efficiency with the forecast model presented in this dissertation.

Moreover, the number of necessary weather stations and their positioning along the overhead line routes are two important key-points for the success of the application of this system. A procedure to locate the weather stations at the hotspots (places along the line route, where the conductor temperature tends to be higher than the rest of the line) was developed and tested on a sample overhead line in Germany. The number of sensor nodes is also adjusted to the

available budget of the TSO. For the sample line of 45 km length, an amount of 20 weather stations covering the most important hotspots of the line implied an average sensor node density of 2.2 km. Considering a cost per weather station of about 4000 Euros, plus the installation costs of 1000 Euros per station, the total cost for this line adds up to 100 thousand Euros. For the German transmission grid with around 35000 km overhead lines, a coarse scaling up results in a total installation cost of around 70 Million Euros. Considering the redispatching costs saving in case DLR with a prediction safety factor is used of around 26 million Euros for 2018 (see Section 7.4.2), the return-of-investment for the installation of a distributed weather measurement system is almost three years. This is a rough calculation, not considering that in the north of Germany less weather stations per kilometer may be needed because of the flat topography. Maintenance costs are also not considered in this result. However, it gives the order of magnitude of the investment return for the system.

The next steps in the development of the final current-carrying capacity forecasting model presented in this thesis corresponds to the reduction of the time to put the system into operations. The results of this dissertation require a year of weather measurements after installation of the hardware. Using transfer learning methods, pre-trained models can have just small adjustments in a short period of time. Moreover, continuous learning can be implemented to have a constant update of the models based on the current measurements from the system.

In general, research is required on forecasting low wind speeds, which is the most influencing factor in the current-carrying capacity estimation. Meteorologists are working on the use of downscaling techniques and the integration of more sophisticated wind turbulence descriptions into mesoscale meteorological models [10]. These models could take advantage of the distributed weather measurements from systems like PrognoNetz.

Finally, other applications, in the area of smart grids or smart cities, could find valuable a network of weather observations distributed over vast areas coupled with predicting models. For example, the prediction of air quality curves is today important to organize the traffic flow in the cities and reduce the localized levels of urban air pollution. That could be managed by a network of sensors collecting weather information (primarily wind measurements) and a machine learning model, which correlates the traffic peak hours, the location of the cars, and the weather conditions.

This dissertation is part of the first steps towards a flexible electrical grid. Optimization mechanisms, as dynamic line rating forecasting systems, offer a short-term solution to extend the need of the construction of new overhead lines. Machine learning algorithms showed the possibility to adjust the current-carrying capacity prediction to the surrounding conditions of the conductor. The dataset, carefully created for training and evaluation, was published open-source to motivate a standardized analysis of the models, allowing other researchers to compare their results. The hope is to see a constant growth of systems supporting the energy transition plan in the years to come.

A Appendix

A.1 The Nusselt number in convective cooling of overhead lines

Note: if not specified otherwise, this section is based on the Cigré Standard 601 [9]

The calculation of the Nusselt number depends on the convection scenario, natural or forced convection, which are directly related to the wind speed. The goal of the following sections is to provide a summary of the most important equations for the calculation of the convective cooling of stranded conductors, which are the most common used in overhead lines in Germany.

A.1.1 Forced convection

This case applies when the wind speed exceeds 0.5 m s^{-1} . Equation (A.1) shows the general mathematical description of the Nusselt number for the forced convective cooling of stranded conductors. The constants B and n depend on the value of the Reynolds number, Re . On the other side, k_1 and k_2 depend on the angle between the line and the wind direction, δ .

$$Nu = BRe^n(0.42 + k_1 \sin(\delta)^{k_2}) \quad (\text{A.1})$$

The Reynolds number is defined by Equation (A.2), with V representing the wind speed (m s^{-1}), D the conductor diameter (m), and ν_f the kinematic viscosity ($\text{m}^2 \text{ s}^{-1}$) of the air.

$$Re = V \frac{D}{\nu_f} \quad (\text{A.2})$$

The kinematic viscosity of the air ($\text{kg m}^{-1} \text{s}^{-1}$), ν_f , can be calculated as the quotient of the dynamic viscosity of the air, μ_f , at the temperature of the film of air in contact with the conductor surface, T_f , divided by the density of the air (kg m^{-3}), γ , at the elevation of the conductor above sea level, y . The expressions (A.3) and (A.4) are valid for a maximum film temperature of 300°C .

$$\mu_f = (17.239 + 4.635 \cdot 10^{-2} \cdot T_f - 2.03 \cdot 10^{-5} \cdot T_f^2) 10^{-6} \quad (\text{A.3})$$

$$\gamma = \frac{1.293 - 1.525 \cdot 10^{-4} \cdot y + 6.379 \cdot 10^{-9} \cdot y^2}{1 + 0.00367 \cdot T_f} \quad (\text{A.4})$$

The coefficients B and n in Equation (A.1) depend on the values of the Reynolds number and the conductor roughness, R_s , which is calculated as $R_s = d/[2(D-d)]$, where d is the wire diameter in the outermost layer and D is the overall conductor diameter. The values for B and n are summarized in Table A.1.

Table A.1: Values of the coefficients B and n for the calculation of the Nusselt number. The selection depends on the Reynolds number, Re , and the conductor roughness, R_s .

$R_s \leq 0.05$			$R_s > 0.05$		
Re	B	n	Re	B	n
100 - 2,650	0.641	0.471	100 - 2,650	0.641	0.471
2,650 - 50,000	0.178	0.633	2,650 - 50,000	0.048	0.800

The coefficients k_1 and k_2 depend on the angle between wind direction and the line. Table A.2 summarizes the possible values of both constants.

Table A.2: Values of the coefficients k_1 and k_2 for the calculation of the Nusselt number.

	k_1	k_2
$0^\circ \leq \delta < 24^\circ$	0.68	1.08
$24^\circ \leq \delta \leq 90^\circ$	0.58	0.90

A.1.2 Natural convection

For the case of no wind, or a wind speed equal to 0 m s^{-1} , the Nusselt number is calculated as shown in Equation (A.5). The coefficients A and m depend on the product of the dimensionless Grashof, Gr , and Prandtl, Pr , numbers.

$$Nu = A(Gr \cdot Pr)^m \quad (\text{A.5})$$

The Grashof number, Gr , depends on the overall diameter of the conductor, D , the average conductor temperature, T , the ambient temperature T_a , the gravity (m s^{-2}), g , the thin film temperature, T_f , and the kinematic viscosity of air, ν_f , as described in Equation (A.6).

$$Gr = \frac{D^3(T - T_a)g}{(T_f + 273)\nu_f^2} \quad (\text{A.6})$$

The Prandtl number, Pr , is a coefficient calculated from the specific heat capacity of air at constant pressure ($\text{J kg}^{-1} \text{K}^{-1}$), c , the dynamic viscosity, μ , and the thermal conductivity of air, λ , as seen in Equation (A.7).

$$Pr = \frac{c\mu}{\lambda} \quad (\text{A.7})$$

Finally, Table A.3 shows the values for A and m depending on the product $Gr \cdot Pr$.

Table A.3: Values of the coefficients A and m for the calculation of the Nusselt number.

	A	m
$10^{-1} \leq GrPr \leq 10^2$	1.020	0.148
$10^2 < GrPr \leq 10^4$	0.850	0.188
$10^4 < GrPr \leq 10^7$	0.480	0.250
$10^7 < GrPr \leq 10^{12}$	0.125	0.333

A.2 Relationship between sag, tension and conductor temperature

Note: if not specified otherwise, this section is based on the Cigré Standard 324 [90].

Depending on the height of the electrical towers the limiting factor for the current-carrying capacity can be the minimum distance of the overhead line to ground instead of the material maximum temperature, given by the manufacturer. Therefore, a short review of the relationship between the conductor temperature, its length and its distance to ground is summarized in this section.

Consider the inclined span shown in Figure A.1, where the supports of the conductor are at different heights.

An overhead line can be mathematically described with the catenary equation (A.8), where H is the horizontal component of the tension (N), w is the conductor weight per unit length (N m^{-1}), x the horizontal distance from the lowest point of the conductor, and $y(x)$ the elevation relative to the lowest point.

$$y(x) = \frac{H}{W} \left[\cosh \left(\frac{wx}{H} \right) - 1 \right] \quad (\text{A.8})$$

If $\frac{x^2 w^2}{12H^2} \ll 1$ then Equation (A.8) can be approximated to a parabolic, as seen in Equation (A.9).

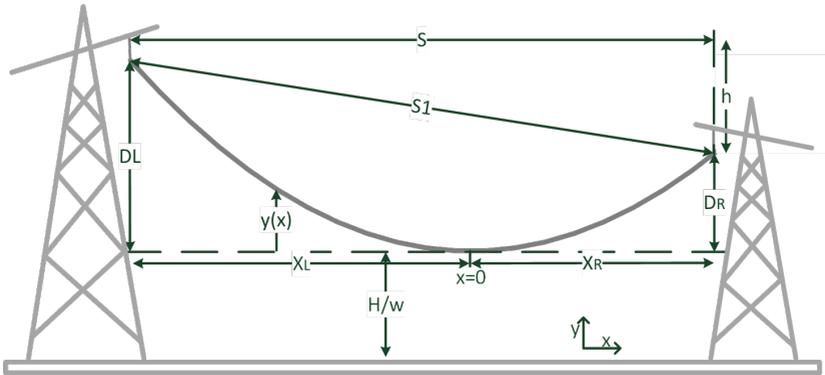


Figure A.1: Inclined span diagram for calculation of line length

$$y(x) \approx \frac{wx^2}{2H} \quad (\text{A.9})$$

In case of a level span, the lowest point of the line is located in the middle of the span. This means that $x_L = x_R = S/2$, with S the span length (m), and $D_L = D_R = D$. Then the conductor sag (m), D , can be calculated as the elevation at the supports, $D = y(S/2)$. Introducing this into Equation (A.9), we obtain an approximated expression for D (Equation (A.10)).

$$D = \frac{wS^2}{8H} \quad (\text{A.10})$$

The term $\frac{H}{w}$ is called the catenary constant. It depends on conductor temperature, wind and ice loading and typically takes values between 500 and 2000 m.

In general, the conductor length can be calculated as the integral of infinitesimal elements of the catenary curve, giving the Equation (A.11) as a result. This expression can be approximated considering again $\frac{x^2w^2}{12H^2} \ll 1$.

$$L(x) = \frac{H}{w} \sinh\left(\frac{wx}{H}\right) \approx x \left(1 + \frac{x^2 w^2}{6H^2}\right) \quad (\text{A.11})$$

In case the line is a level span, the conductor length can be calculated as twice the length at the support, i.e., $L = L\left(\frac{S}{2}\right)$. From the parabolic approximation, Equation (A.12) is obtained.

$$L \approx S \left(1 + \frac{S^2 w^2}{24H^2}\right) \quad (\text{A.12})$$

Or in terms of the conductor sag, D , the conductor length of a level span can be described as Equation (A.13).

$$L \approx S + \frac{8D^2}{3S} \quad (\text{A.13})$$

The relationship between sag and conductor temperature is normally approximated as a linear function, as shown in Equation (A.14).

$$L = L_{ref} \left[1 + \alpha(T_c - T_{c,ref})\right] \quad (\text{A.14})$$

Where L_{ref} is the reference conductor length measured at a conductor temperature considered as reference, $T_{c,ref}$. The coefficient of linear thermal elongation, α , is a constant, which depends on the conductor materials.

The total tension at the supports of an overhead line corresponds to the vector addition of its horizontal and vertical components. Therefore, the magnitude of the total tension can be described as Equation (A.15).

$$T^2 = V^2 + H^2 \quad (\text{A.15})$$

For a level span, the vertical component of the tension at the supports, V , is equal to half the weight of the conductor.

$$V = w \frac{L}{2} = H \sinh \left(\frac{wS}{2H} \right) \quad (\text{A.16})$$

The horizontal component of the tension, H , can be derived from the sag, as in Equation (A.10), or it can be written in terms of the length of the conductor, by introducing Equation (A.11) into (A.10), obtaining the following.

$$H = \frac{wS}{2} \sqrt{\frac{S}{6(L-S)}} \quad (\text{A.17})$$

For inclined spans, the lowest point is not located at $x = S/2$ but at

$$x_{min} = \frac{S}{2} - \frac{H}{w} \sinh^{-1} \left(\frac{wh}{2H \sinh \left(\frac{wS}{2H} \right)} \right) \quad (\text{A.18})$$

From the calculations above, the angle of the line with respect to the horizontal at any given point, x , can be calculated as

$$\delta = \tan^{-1} \left(\frac{dy}{dx} \right) \quad (\text{A.19})$$

Where the derivative of $y(x)$ is the derivative of the catenary equation (A.8), which is

$$\frac{dy}{dx} = \sinh \left(\frac{wx}{H} \right) \quad (\text{A.20})$$

In other words, given the conductor angle at any given x , the value of the catenary constant, H/w , can be obtained. From Equation (A.12), in case of a level span and given the span length, S , the conductor length can be then derived.

B Acronyms

AC	Alternating Current
ACSR	Aluminium Conductor Steel-Reinforced
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
CONUS	Continental US Model
CPU	Central Processing Unit
DC	Direct Current
DLR	Dynamic Line Rating
DSO	Distribution System Operator
DWD	Deutscher Wetterdienst
FFNN	Feedforward Neural Network
GD	Gradient Descent
GRU	Gated Recurrent Unit
INL	Idaho National Laboratory
IRENA	International Renewable Energy Agency
IRES	Intermittent Renewable Energy Sources
LSTM	Long Short-Term Memory

MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MESONET	MESOScale meteorological monitoring NETwork
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OLM	Overhead Line Monitoring
OLMS	Overhead Line Monitoring System
PCA	Principal Component Analysis
QRF	Quantile Regression Forest
ReLU	Rectified Linear Unit
RES	Renewable Energy Sources
RNN	Recurrent Neural Network
SGD	Stochastic Gradient Descent
SLR	Static Line Rating
TSO	Transmission System Operator
WMO	World Meteorological Organization
WWF	World Wide Fund for nature

References

- [1] German Federal Ministry of Research, “German Energy Transition - BMBF,” 2019. [Online]. Available: www.bmbf.de/en/german-energy-transition-2319.html
- [2] WWF, “Germany’s Electric Future II: Regionalization of renewable power generation,” Berlin, Germany, Tech. Rep., 2019. [Online]. Available: www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF-Germanys-Electric-Future-2-RES-Regionalization.PDF
- [3] Bundesnetzagentur, “Quartalsbericht zu Netz- und Systemsicherheitsmaßnahmen: Gesamtjahr und Viertes Quartal 2018,” Bonn, Deutschland, Tech. Rep., 2019.
- [4] BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., “Fakten: Redispatch in Deutschland,” BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., Berlin, Germany, Tech. Rep., 2019.
- [5] Bundesnetzagentur, “Monitoringbericht 2019,” 2020. [Online]. Available: www.bundesnetzagentur.de
- [6] Bundesministerium für Wirtschaft und Energie (BMWi), “Rahmenbedingungen für den Netzausbau.” [Online]. Available: www.bmwi.de/Redaktion/DE/Artikel/Energie/stromnetze-und-netzausbau-regulierung-rahmenbedingungen.html
- [7] “Nova-Prinzip - Netzentwicklungsplan,” 2019. [Online]. Available: www.netzentwicklungsplan.de/de/node/489
- [8] S. Ropenus, F. Peter, and P. Graichen, “Optimierung der Stromnetze,” *Agora Energiewende*, 2017. [Online]. Available: www.agoraenergiewende.de/veroeffentlichungen/optimierung-der-stromnetze/
- [9] Cigre, “Guide for Thermal Rating Calculations of Overhead Lines,” Technical Report, 2014.
- [10] A. Michiorri, H. M. Nguyen, S. Alessandrini, J. B. Bremnes, S. Dierer, E. Ferrero, B. E. Nygaard, P. Pinson, N. Thomaidis, and S. Uski, “Fore-

- casting for dynamic line rating,” *Renewable and Sustainable Energy Reviews*, no. 52, pp. 1713–1730, 2015.
- [11] T. Montmerle, “Statement of Guidance for High-Resolution Numerical Weather Prediction,” World Meteorological Organization, Tech. Rep., 2018. [Online]. Available: www.wmo.int/pages/prog/www/OSY/SOG/SoG-HighRes-NWP.pdf
- [12] Deutscher Wetterdienst (DWD), “Wettermodelle,” Deutscher Wetterdienst, Tech. Rep., 2017.
- [13] R. Steinacker, M. Dorninger, S. Schneider, and S. Tschannett, “High Resolution Analysis and Nowcasting over Complex Terrain by Using Physical a Prior Knowledge,” in *Proc., International Symposium on Nowcasting and Very Short Range Forecasting WSN05*, Toulouse, France, 2005.
- [14] Bundesnetzagentur, “Monitoringbericht 2018,” 2019. [Online]. Available: www.bundesnetzagentur.de
- [15] N. Doban, “Building predictive models for dynamic line rating using data science techniques,” mathesis, KTH, Stockholm, 2016.
- [16] A. J. Conejo and L. Baringo, *Power System Operations*. Springer International Publishing, 2018. [Online]. Available: link.springer.com/book/10.1007%2F978-3-319-69407-8
- [17] H. Huang, D. Kumar, V. Rumaswami, and D. Retzmann, “UHV 1200 kV AC Transmission,” in *GridTech*, New Delhi, 2007. [Online]. Available: www.ptd.siemens.de/070201_AC1000kV_GRIDTECH.pdf
- [18] Museum of Energy of the Urals, “Unique Power Transmission Line - 1150,” 2012. [Online]. Available: musen.ru/chronicle/1987_2
- [19] K. Padiyar, *FACTS: Controllers in Power Transmission and Distribution*. Tunbridge Wells: Anshan LTD, 2009.
- [20] P. Kundur, *Power system stability and control*. McGraw-Hill Education, 1994.
- [21] K. Van Den Bergh, E. Delarue, and W. D’haeseleer, “DC power flow in unit commitment models,” KU Leuven Energy Institute, Tech. Rep. May, 2014. [Online]. Available: www.mech.kuleuven.be/en/tme/research/energy_environment/Pdf/wpen2014-12.pdf
- [22] IPCC, *Summary for Policymakers*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2014. [Online]. Available: www.ipcc.ch/working-group/wg3

- [23] E. Hawkins, P. Ortega, E. Suckling, A. Schurer, G. Hegerl, P. Jones, M. Joshi, T. J. Osborn, V. Masson-Delmotte, J. Mignot, P. Thorne, and G. J. van Oldenborgh, "Estimating changes in global temperature since the preindustrial period," *Bulletin of the American Meteorological Society*, vol. 98, no. 9, pp. 1841–1856, 2017. [Online]. Available: doi.org/10.1175/BAMS-D-16-0007.1
- [24] International Renewable Energy Agency (IRENA), "Energy Transition," 2019. [Online]. Available: www.irena.org/energytransition
- [25] Energiewende Web, "Energiewende: Geschichte, aktuelle Situation, Zukunft - energiewende.de," 2019. [Online]. Available: www.energiewende.de/
- [26] A. Zablocki, "Fact Sheet: Energy Storage (2019)," 2019. [Online]. Available: www.eesi.org/papers/view/energy-storage-2019
- [27] P. Staudt, "Transmission congestion management in electricity grids: Designing markets and mechanisms," PhD Thesis, Karlsruhe Institute of Technology, 2019.
- [28] Cigre, "CIGRE Technical Brochure 299 - Guide for the selection of weather parameters for bare overhead conductor ratings," no. August, 2006.
- [29] "DIN EN 50341-2 VDE 0210-2:2002-11 - Freileitungen über AC 45 kV," 2002.
- [30] Amprion, Tennet, Transnet BW, and 50 Hertz, "Grundsätze für die Ausbauplanung der deutschen Übertragungsnetzes," Tech. Rep., 2018.
- [31] W. Black and R. Rehberg, "Simplified Model for Steady State and Real-Time Ampacity of Overhead Conductors," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 10, pp. 2942–2953, oct 1985. [Online]. Available: ieeexplore.ieee.org/document/4112966
- [32] H. A. Enos, "Current-carrying capacity of overhead conductors," *Electrical World*, no. May 15, pp. 60–62, 1943.
- [33] B. V. Viscomi, "Steady state thermal ratings for 220 kV overhead conductors," Lafayette College, Tech. Rep., 1972. [Online]. Available: archive.org/details/ECL-197
- [34] H. E. House and P. D. Tuttle, "Current-Carrying Capacity of ACSR," *Electrical Engineering*, vol. 77, no. 8, pp. 1169–1173, 1959.

- [35] S. Uski-Joutsenvuo and R. Pasonen, "Maximising power line transmission capability by employing dynamic line ratings - technical survey and applicability in Finland," Technical Report, 2013.
- [36] G. Biedenbach, "Kapazitäten in Stromnetzen Optimal Ausschöpfen. Monitoring macht den Betrieb von Freileitungen Sicherer." *EW Magazin für die Energiewirtschaft*, pp. 74–83, 2009.
- [37] Micca, "Monitoring on overhead transmission lines. EMO Easy Monitoring Overhead Transmission," Technical Information.
- [38] "OTLM sensor - OTLM." [Online]. Available: www.otlm.eu/energy/otlm-device/
- [39] E. Cloet, J. L. Lilien, and P. Ferrieres, "Experiences of the Belgian and French TSOs using the "Ampacimon" real-time dynamic rating system," *Cigre session C2-106-2010*, 2010.
- [40] Ampacimon, "Innovative solution for Dynamic Line Rating," 2015. [Online]. Available: www.ampacimon.com/wp-content/uploads/2015/09/Ampacimon-Brochure.pdf
- [41] S. Kurth, "Astrose - Autarkes Sensornetzwerk zum Monitoring in der Energietechnik. Abschlussbericht." 2012.
- [42] S. Voigt, "Autarkes Monitoringsystem für Hochspannungsleitungen," Article, 2013.
- [43] NOAA, "What is LIDAR?" 2018. [Online]. Available: oceanservice.noaa.gov/facts/lidar.html
- [44] J. McCall, P. Spillane, and K. Lindsey, "Determining crossing conductor clearance using line-mounted lidar," *Cigre*, 2015.
- [45] LineVision, "LineVision Technology," 2019. [Online]. Available: www.linevision.co/technology
- [46] J. Marmillo, N. Pinney, B. Mehraban, and S. Murphy, "A Non-Contact Sensing Approach for the Measurement of Overhead Conductor Parameters and Dynamic Line Ratings," in *Cigre 2017 Grid of the Future Symposium*. CIGRE, 2017.
- [47] Erbeuerbare Energien, "Leistung der Stromnetze erhöhen statt auf Netzausbau zu warten." [Online]. Available: www.erneuerbareenergien.de/archiv/leistung-der-stromnetze-erhoehen-statt-auf-netzausbau-zu-warten-150-406-105759.html

-
- [48] S. D. Foss and R. A. Maraio, "Evaluation of an overhead line forecast rating algorithm," *IEEE Transactions on Power Delivery*, no. 7, pp. 1618–27, 1992. [Online]. Available: ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=141882
- [49] J. F. Hall and A. K. Deb, "Prediction of overhead transmission line ampacity by stochastic and deterministic models," *IEEE*, vol. 3, no. 2, 1988.
- [50] D. A. Douglass, "Weather-Dependent versus static thermal line ratings," *IEEE Transactions on Power Delivery*, vol. 3, no. 2, pp. 742–753, 1988.
- [51] EU-Project, "Twenties project. Final report," 2013.
- [52] H. M. Nguyen, J. J. Lambin, F. Vassort, and J. L. Lilien, "Operational experience with Dynamic Line Rating forecast-based solutions to increase usable network transfer capacity," *CIGRE Session 45 - 45th International Conference on Large High Voltage Electric Systems 2014*, vol. 2014-August, 2014.
- [53] H. E. Hoekstra, C. P. J. Jansen, J. Hagen, and J. W. V. Schuylenburg, "Weather forecasted thermal line rating model for the Netherlands," *Cigre*, no. B2-105, 2012.
- [54] T. Ringelband, "Unisicherheit witterungsabhängiger Übertragungskapazitäten bei der Netzbetriebsplanung," Dissertation, RWTH Aachen, Aachen, Deutschland, 2011.
- [55] J. Aznarte and N. Siebert, "Dynamic Line Rating Using Numerical Weather Predictions and Machine Learning: A Case Study," *IEEE transactions on power delivery*, vol. 32, no. 1, pp. 335–343, 2017.
- [56] Ampacimon, "ADR Horizon." [Online]. Available: www.ampacimon.com/adr-horizon
- [57] Lindsey, "Smartline." [Online]. Available: lindsey-usa.com/dynamic-line-rating
- [58] Energy&Meteo Systems, "Grid Load Forecasts." [Online]. Available: www.energymeteo.com/products/power_forecasts/grid_load_forecasts.php
- [59] S. S. Smith, "Blockchain, artificial intelligence and financial services: Implications and applications for finance and accounting professionals," Cham, 2020. [Online]. Available: doi.org/10.1007/978-3-030-29761-9

- [60] B. Wollenberg and T. Sakaguchi, "Artificial Intelligence in Power System Optimization," in *Proceedings of the IEEE*, vol. 75, no. 12, 1987, pp. 1678–1685. [Online]. Available: ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1458202&tag=1
- [61] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Massachusetts: MIT Press, 2016.
- [62] R. T. Kreutzer, "Understanding artificial intelligence : Fundamentals, use cases and methods for a corporate ai journey," Cham, 2020. [Online]. Available: doi.org/10.1007/978-3-030-25271-7
- [63] T. Mitchell, *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997.
- [64] E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," 2010.
- [65] J. Brownlee, *Long Short-Term Memory Networks With Python: Develop Sequence Prediction Models With Deep Learning*, 2017.
- [66] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014. [Online]. Available: arxiv.org/abs/1412.3555
- [67] J. Brownlee, *Deep Learning with Python*, 2018.
- [68] J. Becker and H. M. Lipp, *Grundlagen der Digitaltechnik [Introduction to Digital Techniques]*, 2009.
- [69] D. Rosenberg, "Classification and regression trees," New York University, Tech. Rep., 2015.
- [70] G. Biau, "Analysis of a Random Forests Model," *Journal of Machine Learning Research*, vol. 13, 2012.
- [71] N. Meinshausen, "Quantile Regression Forests," *Journal of Machine Learning Research*, vol. 7, 2006.
- [72] D. Douglass, W. Chisholm, G. Davidson, I. Grant, K. Lindsey, M. Lancaster, D. Lawry, T. McCarthy, C. Nascimento, M. Pasha, J. Reding, T. Seppa, J. Toth, and P. Waltz, "Real-Time Overhead Transmission-Line Monitoring for Dynamic Rating," *IEEE Transactions on Power Delivery*, vol. 31, no. 3, pp. 921–927, 2016.

- [73] BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., “Monatlicher Stromverbrauch in Deutschland,” 2019. [Online]. Available: www.bdew.de
- [74] “TransnetBW - Engpass.” [Online]. Available: www.transnetbw.de/de/strommarkt/engpassmanagement/engpass
- [75] “TenneT - Statisches Netzmodell.” [Online]. Available: www.tennet.eu/de/strommarkt/transparenz/transparenz-deutschland/
- [76] “50Hertz - Statisches Netzmodell.” [Online]. Available: www.50hertz.com/de/Transparenz/Kennzahlen/Netzdaten/StatischesNetzmodell/
- [77] “Amprion - Statisches Netzmodell.” [Online]. Available: www.amprion.net/Strommarkt/Engpassmanagement/Statisches-Netzmodell/
- [78] F. Leuthold, H. Weigt, and C. von Hirschhausen, “Efficient pricing for European electricity networks - The theory of nodal pricing applied to feeding-in wind in Germany,” *Utilities Policy*, vol. 16, no. 4, pp. 284–291, 2008. [Online]. Available: [dx.doi.org/10.1016/j.jup.2007.12.003](https://doi.org/10.1016/j.jup.2007.12.003)
- [79] K. L. Clawson, R. M. Eckman, N. F. Hukari, J. D. Rich, and N. R. Ricks, “NOAA Technical Memorandum OAR ARL-259: Climatology of the Idaho National Laboratory, 3rd Edition,” National Oceanic and Atmospheric Administration, Tech. Rep. December, 2007.
- [80] NOAA, “INL Mesonet Data Extractor.” [Online]. Available: www.noaa.inel.gov/extract
- [81] Deutscher Wetterdienst (DWD), “Climatological Maps of Germany.” [Online]. Available: www.dwd.de/EN/ourservices/klimakartendeutschland/detailbeschreibung_en.html
- [82] NOAA, “NOAA/INL MESONET Instrumentation,” NOAA, Tech. Rep.
- [83] J. Shonk, *Introducing Meteorology: A Guide to the Weather*. Edinburgh: Dunedin Academic Press Ltd., 2013.
- [84] National Oceanic & Atmospheric Administration, “National Digital Forecast Database and Local Database: Description and Specifications,” Tech. Rep., 2019. [Online]. Available: www.nws.noaa.gov/directives
- [85] NOAA, “NDFD Monthly Verification Web Page,” 2020. [Online]. Available: sats.nws.noaa.gov/~verification/ndfd
- [86] American Meteorological Society, “Persistence Forecast,” 2012. [Online]. Available: http://glossary.ametsoc.org/wiki/Persistence_forecast

- [87] F. Lutgens and E. Tarbuck, *The Atmosphere: An Introduction to Meteorology*. Prentice Hall, 1995.
- [88] J. R. Holton and G. J. Hakim, *An Introduction to Dynamic Meteorology*. Elsevier, 2013.
- [89] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *Journal of Machine Learning Research*, vol. 13, pp. 281–305, 2012.
- [90] Cigre, “Sag-Tension Calculation Methods for Overhead Lines,” Technical Report, apr 2016.
- [91] Bundesministerium für Wirtschaft und Energie (BMWi) and Bundesministerium für Umwelt (BMU), “Energy Concept for an Environmentally Sound, Reliable and Affordable Energy Supply,” Tech. Rep., 2010.
- [92] EUMeTrain, “Skill Scores,” 2019. [Online]. Available: www.eumetrain.org
- [93] Fraunhofer ISE, “Stromproduktion: Energy Charts,” 2019. [Online]. Available: www.energy-charts.de
- [94] Agora Energiewende und Energynautics, “Toolbox für die Stromnetze - Für die künftige Integration von Erneuerbaren Energien und für das Engpassmanagement,” Tech. Rep., 2018. [Online]. Available: www.agora-energiewende.de
- [95] C. W. G. 22.12, “207 - Thermal behaviour of overhead conductors,” Cigre, Tech. Rep. August, 2002.
- [96] T. Krontiris, A. Wasserrab, and G. Balzer, “Weather-based Loading of Overhead Lines - Consideration of Conductor’s Heat Capacity,” *Proceedings of the Modern Electric Power Systems 2010*, pp. 1–8, 2010.
- [97] The Valley Group, “CAT-1 Transmission Line Monitoring System.” [Online]. Available: www.nexans.be
- [98] USi, “CAT-1 DYNAMIC LINE RATING SYSTEM - USi,” 2016. [Online]. Available: www.usi-power.com/cat-1-dynamic-line-rating-system/
- [99] C. Singh, A. Singh, P. Pandey, and H. Singh, “Power Donuts in Overhead Lines for Dynamic Thermal Rating Measurement, Prediction and Electric Power Line Monitoring,” *International Journal of Advanced*

- Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 5, pp. 9394–9400, 2014.
- [100] Lindsey, “Real Time Transmission Line Conductor Monitor,” Azusa, USA, p. 2, 2016. [Online]. Available: www.lindsey-usa.com
- [101] Bundesnetzagentur, “Jahresbericht 2018,” Bundesnetzagentur, Tech. Rep., 2018. [Online]. Available: www.bundesnetzagentur.de
- [102] G. Biedenbach, “Echtzeitmonitoring von Freileitungen: Leiterseiltemperatur zuverlässig ermitteln.” *EW Magazin für die Energiewirtschaft*, vol. 7104, no. 15-16, pp. 70–73, 2012.
- [103] National Oceanic and Atmospheric Administration, “National Digital Forecast Database (NDFD) - NOAA.” [Online]. Available: www.ncdc.noaa.gov/data-access/model-data/model-datasets/national-digital-forecast-database-ndfd
- [104] M. Giuntoli, P. Pelacchi, D. Poli, F. Bassi, and G. Giannuzzi, “Thermomechanical dynamic rating of OHTL: applications to Italian lines,” *Cigre*, no. C2-112, 2014.
- [105] E. Fernandez, I. Albizu, G. Buigues, V. Valverde, A. Etxegarai, and J. G. Olazarri, “Dynamic line rating based on numerical weather prediction,” *2015 IEEE Eindhoven PowerTech, PowerTech 2015*, pp. 1–6, 2015.
- [106] R. Martínez, A. González, A. Madrazo, M. Mañana, A. Arroyo, M. A. Cavia, R. Domingo, A. Sierra, and A. Laso, “Ampacity forecasting using neural networks,” *Renewable Energy and Power Quality Journal*, vol. 1, no. 12, pp. 120–123, 2014.
- [107] A. Babs, “Weather-based and conductor state measurement methods applied for dynamic line rating forecasting,” *APAP 2011 - Proceedings: 2011 International Conference on Advanced Power System Automation and Protection*, vol. 1, pp. 762–765, 2011.
- [108] Deutscher Wetterdienst (DWD), “Wetter und Klima - Deutscher Wetterdienst - NWV-Änderungen.” [Online]. Available: www.dwd.de
- [109] Institute for Essential Services Reform, “Energy Transition in a Nutshell: 8 Q & A on the German Energy Transition and Its Relevance for Indonesia,” *Agora Energiewende*, 2018. [Online]. Available: www.agora-energiewende.de

- [110] NASA, "Causes. Facts - Climate Change: Vital Signs of the Planet," 2019. [Online]. Available: climate.nasa.gov/causes
- [111] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, 05 2012.
- [112] Energy&Meteo, "Current Carrying Capacity Forecast for Overhead Power Lines," 2019. [Online]. Available: www.energymeteo.com/products/power_forecasts/grid_load_forecasts.php
- [113] A. Lanciani and M. Salvati, "Spatial interpolation of surface weather observations in alpine meteorological services," *Environmental Protection*, p. 44, 2008.
- [114] European Commission, "2020 climate & energy package." [Online]. Available: ec.europa.eu/clima/policies/strategies/2020_en

Personal publications

Journal articles

- [115] G. Molinar and W. Stork, “Ampacity forecasting using machine learning: an approach based on distributed weather measurements,” *High Performance Computing in Science and Engineering 2020*, 2020, submitted.

Conference proceedings

- [116] G. Molinar and W. Stork, “Infrared spectral imaging for damage detection and prevention of overhead power lines,” in *OCM 2017 - Optical Characterization of Materials - conference proceedings*. Hrsg.: J. Beyerer, F. Puente León, T. Längle. KIT Scientific Publishing, Karlsruhe, 2017, pp. 159–168.
- [117] N. Popovic, G. Molinar, and W. Stork, “From data points to ampacity forecasting: Gated Recurrent Unit networks,” in *IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService)*. Bamberg, Germany: IEEE, March 2018, pp. 200–207.
- [118] G. Molinar, L. T. Fan, and W. Stork, “Ampacity forecasting: an approach using Quantile Regression Forests,” in *IEEE International Conference for Smart Grid Technologies - North America*. Washington DC: IEEE, 2019.
- [119] G. Molinar, M. Li, and W. Stork, “Positioning of distributed weather stations for overhead line monitoring systems based on historical weather data,” in *IEEE AFRICON*. IEEE, 2019.
- [120] J. Becker, D. Grimm, T. Hotfilter, C. Meier, G. Molinar, M. Stang, S. Stock, and W. Stork, “The QUA³CK Machine Learning Development

Process and the Laboratory for Applied Machine Learning Approaches (LAMA),” in *Symposium Artificial Intelligence for Science, Industry and Society (AISIS 2019)*. Mexico: Symposium Artificial Intelligence for Science, Industry and Society (AISIS 2019), 2019. [Online]. Available: indico.cern.ch/event/781223/book-of-abstracts.pdf

- [121] G. Molinar, J. Bassler, N. Popovic, and W. Stork, “Ampacity forecasting from Numerical Weather Predictions: a fusion of the traditional and machine learning methods,” in *IEEE International Conference for Smart Grid Technologies - Europe*. IEEE, 2020, accepted.

Invited talks

- [122] G. Molinar, “Optimizing the grid using machine learning: towards a sustainable electrical system,” in *Sino-EU Doctoral School for Sustainability Engineering 2018*, Shanghai, China, September 2018.
- [123] —, “Ampacity forecasting using machine learning: an approach based on distributed weather measurements,” in *Cigre TAG 4, Working Group B2.59*, Hakodate, Japan, April 2019.
- [124] —, “PrognoNetz - Strombelastbarkeitsprognose mittels künstlicher Intelligenz,” in *ZVEI - Zentralverband Elektrotechnik- und Elektronikindustrie e.V.*, Hamburg, Germany, October 2019.
- [125] G. Molinar and N. Shen, “Ampacity forecasting using machine learning: an approach based on distributed weather measurements,” in *7th International Women4Energy Conference*, Stuttgart, Germany, December 2019.

Supervised Student Work

- [126] A. Haque, “Master Thesis: Development of an optimal synchronization system for remotely located sensor units of overhead line monitoring systems,” March 2017.
- [127] A. Silini, “Bachelor Thesis: Langzeitige Belastbarkeitsprognose der Freileitungen mittels Machine Learning Algorithmen,” November 2017.
- [128] M. Leiberich, “Master Thesis: Auswahl, Implementierung und Validierung maschineller Lernalgorithmen für lokale, kurzzeitige Wetterprognosen,” November 2017.
- [129] Y. Yang, “Bachelor Thesis: Entwicklung eines Windgeschwindigkeits- und Windrichtungssensor mittels der Charakterisierung von Schallwellen,” November 2017.
- [130] N. Popovic, “Bachelor Thesis: Spatial Correlation of Weather Data and its Impact on Weather Forecasting by Means of Machine Learning,” December 2017.
- [131] L. Öfele, “Bachelor Thesis: Umsetzung und Auswertung von Gauss-Prozessen zur Verbesserung der Strombelastbarkeitsprognose von Hochspannungsleitungen,” April 2018.
- [132] X. Zhu, “Bachelor Thesis: Implementierung eines energieeffizienten Kommunikationskanals für weitreichende Sensornetzwerke,” April 2018.
- [133] B. Chirumamilla, “Master Thesis: Development of an optical prediction tool for ramp-down events on photovoltaic systems,” May 2018.
- [134] L. T. Fan, “Master Thesis: Wetterbedingte Belastbarkeitsprognose von Freileitungen mittels Quantile Regression Forests,” June 2018.
- [135] F. Völz, “Master Thesis: Implementierung und Bewertung rekurrenter neuronaler Netze für die Strombelastbarkeitsprognose einer Freileitung,” October 2018.

- [136] F. Schweikardt and J. Herrmann, "Master Thesis: Objekterkennung mittels Transfer Learning durch Convolutional Neural Networks (CNN): Eine Fallstudie zur automatisierten Mahlzeitenerkennung," October 2018, [Supervised together with Simon Stock].
- [137] P. Hofmann, "Bachelor Thesis: Implementierung eines drahtlosen Sensornetzwerks," November 2018.
- [138] B. Haas, "Master Thesis: Reinforcement-Learning für die Strombelastbarkeitsprognose einer Freileitung über 48 Stunden mittels Daten verteilter Wetterstationen," December 2018.
- [139] M. Mandhouj, "Bachelor Thesis: Implementation of Ensemble Methods to Increase the Accuracy of the Ampacity Forecasting of High Voltage Overhead Lines," January 2019.
- [140] H. Raza, "Master Thesis: Development and Evaluation of a Few Shot Segmentation Algorithm for Computer Vision Applications," March 2019.
- [141] M. Zhang, "Bachelor Thesis: Two-day ahead ampacity forecasting of overhead lines based on parallelized Quantile Regression Forests," May 2019.
- [142] F. Marquardt, "Bachelor Thesis: Untersuchung des Einflusses unterschiedlicher Datenaufbereitungsmethoden beim Training von LSTM- und QRF-Modellen," June 2019.
- [143] K. Shi, "Master Thesis: Data Processing of Numerical Weather Predictions and Evaluation of their Suitability for Ampacity Forecasting of Overhead Lines," June 2019.
- [144] M. Li, "Master Thesis: Process development for spatial distribution of sensor nodes for overhead line monitoring systems based on meteorological and topographical features," July 2019.
- [145] X. Chen, "Bachelor Thesis: Ampacity forecasting accuracy improvement based on a dynamic and automatic selection over time of the best performing ensemble method," July 2019.
- [146] P. Jahnke, "Bachelor Thesis: Vergleich der räumlichen Interpolationsverfahren für Wetterdaten und ihre Integration in selbstlernende Modelle zur Strombelastbarkeitsprognose des Hochspannungsnetzes," September 2019.

- [147] R. Schmitt, "Bachelor Thesis: Development of an Energy Saving Wake Up Pattern Detection and Transferring of the race|result Active Transponder Protocol on an 32-bit Microcontroller," October 2019.
- [148] L. H. Fiedler, "Master Thesis: Entwicklung und Test einer adaptiven Funkverbindung von Smart-Grid Sensorknoten," December 2019, [Supervised together with Lintao Toni Fan].
- [149] J. Bassler, "Bachelor Thesis: Prognoseverfahren zur Strombelastbarkeit mithilfe von Convolutional Neural Networks und räumlich verteilten, numerischen Wetterprognosen," Februar 2020.
- [150] J. Denig, "Bachelor Thesis: Feature-Engineering verschiedener CNN Modelle zur Strombelastbarkeitsprognose von Freileitungen in Deutschland," August 2020.