

**Forecasting of Dynamic Thermal Line Rating Under the Conditions of
Temporal Discretization and Correlation**

by

Tomas Barton

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering
University of Alberta

© Tomas Barton, 2021

Abstract

Dynamic Thermal Line Rating (DTLR) is a technology that optimizes the utility of overhead power transmission lines by dynamically adjusting the rating according to current ambient conditions. To be truly useful, forecasting of DTLR must be applied within processes governing the function of the electrical system. This thesis focuses on medium term DTLR forecasts on the timescale of days and hours.

Two DTLR forecasting systems are developed within this thesis. Both systems are categorized as indirect probabilistic rating systems, where Numerical Weather Predictions (NWP) are processed to estimate the rating. The systems differ in the method used to quantify uncertainty. In the first system, the uncertainty is described by a custom statistical model that is fitted onto the historical forecasts of the model in a scheme called Model Output Statistic. In the other system a Random Forest machine learning model is employed to produce probabilistic output associated with the NWP output. This system uses the regression-via-classification approach, where most processing occurs in the discrete domain even though the inputs and outputs are continuous. Both developed systems provide DTLR predictions that perform better than Static Line Rating and a reference method.

The DTLR forecasting systems in this thesis have been specifically designed to produce forecasts of temporally discretized rating. Temporal discretization is a term defined in this thesis as the process of taking the continuous DTLR and turning it into a single value valid over a period. Temporal discretization decreases DTLR benefit and this relationship has been studied and simulated across a vast dataset spanning all of Canada. A figure with the relative benefit of DTLR for different lengths of

discretization periods is provided.

The effects of temporal discretization on DTLR forecasts are also addressed. A simulation of DTLR forecasts with different discretization period lengths was performed to analyze the sensitivity of DTLR forecasts to this variable. It was revealed that the overall benefit of DTLR forecasts is relatively insensitive to the discretization length.

It was argued that temporal correlations within DTLR time series impact the discretized value. It was argued that ignoring these correlations will negatively affect the probabilistic prediction by skewing the distribution toward the extreme value. This was demonstrated on the simulated forecast. A Monte Carlo system with a novel sampling method was proposed to mitigate this problem by generating realistic time series that are used to estimate the rating. This method yields well calibrated results.

Preface

This thesis is an original work by Tomas Barton performed under the supervision of Dr. Petr Musilek. Parts of the thesis have been published or submitted to journals or conferences, as indicated below.

Chapter 3 has been published as Bartoň, T., Musilek, P., Derivative based prediction with look ahead. Proceedings of the International Joint Conference on Neural Networks (IJCNN 2016), 24-29 July 2016, Vancouver, Canada, art. no. 7727461, pp. 2118-2123.

Chapter 4 has been published as Barton, T., Musilek, P., Day-Ahead Dynamic Thermal Line Rating Using Numerical Weather Prediction. 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE 2019), Edmonton, Canada, art. no. 8861883.

Chapter 5 has been published as Barton, T., Musilek M., Musilek, P., The Effect of Temporal Discretization on Dynamic Thermal Line Rating, 21st International Scientific Conference on Electric Power Engineering (EPE 2020), Prague, Czech Republic, October 19-21, 2020

Chapter 6 has been submitted to International Journal of Electrical Power and Energy Systems as Barton, T., Musilek P.: Probabilistic Forecasting of Dynamic Thermal Line Rating with Temporal Correlations.

Acknowledgements

I would like to thank my supervisor Dr. Petr Musilek, who deserves my gratitude for his support, guidance, and the opportunity he offered me at the University of Alberta that allowed me to start a new chapter of life here in Canada.

I would also like to thank my wife Amanda Barton for keeping me on track and staying with me up through endless nights helping me edit my papers.

In particular, I would like to thank Rodrigo Castro Martins for all the fruitful discussions.

This work has been supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Alberta Electric System Operator (AESO), ATCO Electric, Altalink LP, and Canada and Alberta Ingenuity Technology Futures (AITF).

Table of Contents

1	Introduction	1
1.1	Motivation	2
1.2	Thesis Objectives	4
1.3	Research Originality	5
1.4	Thesis Outline	6
2	Background and Related Work	9
2.1	Background	9
2.1.1	Benefits of DTLR	16
2.1.2	Thermal Model of Transmission Line	17
2.1.3	Static Thermal Line Rating	19
2.1.4	Time Series	19
2.1.5	Extreme Values of Time Series	20
2.2	Related Work	21
3	Derivative Based Prediction with Look Ahead	23
3.1	Methods	25
3.1.1	Artificial Neural Networks	25
3.1.2	Genetic Algorithms	27
3.2	Improved DBP Algorithms	28
3.2.1	Original DBP	28
3.2.2	Delayed DBP	30

3.2.3	DBP with look ahead	31
3.3	Evaluation	35
3.4	Conclusions	37
4	Day-Ahead Dynamic Thermal Line Rating using Numerical Weather Prediction	40
4.1	Methods	44
4.1.1	Probabilistic Prediction	44
4.1.2	Naive prediction	48
4.1.3	Perfect prediction	48
4.1.4	NWP with derating	50
4.2	Results	50
4.3	Conclusion	57
5	The Effect of Temporal Discretization on Dynamic Thermal Line Rating	59
5.1	Temporal Discretization	61
5.2	Wind speed model	62
5.2.1	Static Line Rating	64
5.2.2	Dynamic Thermal Line Rating	66
5.2.3	Dataset	66
5.3	Discussion	67
5.3.1	Temporal Discretization of Wind Speed	67
5.3.2	Temporal Discretization of Rating	69
5.3.3	Limitations	72
5.4	Conclusion	72
6	Probabilistic Forecasting of Dynamic Line Rating with Temporal Correlations	75

6.1	Methods and Procedures	78
6.1.1	Machine Learning Model	80
6.1.2	Conductor Temperature Estimation	81
6.1.3	Time Series Sampling	84
6.1.4	Over-Temperature Cost Function	87
6.1.5	Static Line Rating Overtemperature	89
6.1.6	Rating Estimation Methods	89
6.2	Results	90
6.2.1	Dataset	90
6.2.2	Simulation details	91
6.2.3	Classification model	91
6.2.4	Percentile accuracy	93
6.2.5	Distribution of sampled minimum	94
6.2.6	Comparison of rating strategies	96
6.3	Conclusion	101
7	Conclusion and Final Remarks	106
7.1	Future Research Direction	108
	Bibliography	110

List of Tables

3.1	The performance of DBP and its variants on two temperature series.	37
4.1	Simulation results	56
5.1	Discretized DTLR benefit	71

List of Figures

1.1	The timeline of the electrical market	3
2.1	Diagram of the electrical system	9
2.2	Length of the transmission line and its rating	12
3.1	A schematic diagram of the NARX network	26
3.2	An example of DBP algorithm processing a time series	29
3.3	DBP and Delayed DBP comparison	31
3.4	Diagram of DBP, Delayed DBP and DBP with look-ahead	32
3.5	Performance of proposed DBP algorithms	36
4.1	Histogram of the real-time ampacity	43
4.2	Example of a probabilistic forecasts	45
4.3	Example of the time series of ampacity	51
4.4	Relationship between the overestimated and transmitted energy	52
5.1	Weibull fit to discretized wind speed	63
5.2	Shape and scale parameters of the Weibull distribution	64
5.3	Experimental results	65
5.4	Histogram of seasonal SLR	65
5.5	Extreme value distribution fitted to the rating	70
5.6	The power law approximation	71
6.1	Temperature of an overloaded conductor	82

6.2	Diagram of the rating prediction system	85
6.3	Histogram of median forecast residuals	93
6.4	PIT Histogram of the probabilistic forecast	94
6.5	Q-Q diagram of probabilistic forecast	95
6.6	PIT of minimal values	96
6.7	Rating forecast for different discretization periods	97
6.8	Rating forecast for different over-temperatures	98

Chapter 1

Introduction

Electricity has become an essential pillar in society and the advancements made in this field have revolutionized everyday life. It was not too long ago when candles were used to illuminate homes, cooking was done on wood burning stoves, and nobody owned cell phones. Now it seems as though electricity is integrated into all facets of daily living from turning on lights, to driving electric vehicles, all the way to cleaning with robot vacuums. It truly feels like everything is or soon will be just a flick of a switch away.

In 1873, the first electrical light was switched on in Winnipeg, Canada. Then the first Canadian movie theater opened in 1906. Soon after, in 1964, Canada was the third country to launch a satellite into space [1]. It is evident, that there has been tremendous growth in the electrical system over the years. In fact, 100% of the Canadian population has access to electricity [2]. So while at one time it may have only been possible to power a single light bulb, as a society we have become capable of the unimaginable.

At the heart of all these achievements lies the transmission and distribution system. In order to get to its target location, electricity must first be produced and then delivered through this complex system. To produce electricity, power plants use either fuel or harvest renewable natural resources. Then the electrical grid provides the infrastructure to transport the electricity from the producers to the consumers

who utilize the energy for personal, business, or industrial use.

Power can be injected into the electrical network at any location and then extracted in another completely distant location. Thus, the electrical grid works behind the scenes and over vast distances only to be noticed if something goes wrong. As we increase our electrical demands, power lines are being pushed to operate at their maximum capacity. Consequences of insufficient transmission capacity are price elevations due to a lack of access to cheaper sources, curtailment of renewable power sources, and in the most unfortunate case load shedding. Congestion is therefore, a serious emerging problem that demands innovation and investment into the system.

Dynamic thermal line rating (DTLR) is one such innovation that aims to best utilize the available capacity of the current transmission system. It does this through precisely estimating the true capacity. This is in contrast to the more conservative but crude approach known as static line rating (SLR). While simple and easy to implement, it wastes a lot of potential capacity as it sets a rating that assumes the worst possible conditions.

1.1 Motivation

DTLR, by definition allows for the optimal utilization of conductor capacity. Assuming the sensors perfectly monitor both the conductor and environment without error or uncertainty, the operator can use the thermal model to calculate the true rating. This process is termed Real-Time DLR (RT-DTLR) [3].

This strategy, while theoretically optimal, comes with practical limitations that reduce its utility in many scenarios. Firstly, highly accurate sensors with substantial spatial and temporal resolution are required to be installed directly on the transmission line conductor. The sensors are distributed in sufficient quantities and are capable of providing real-time communication with the center of operations. This process implies costs associated in installation, ongoing maintenance, and data collection.

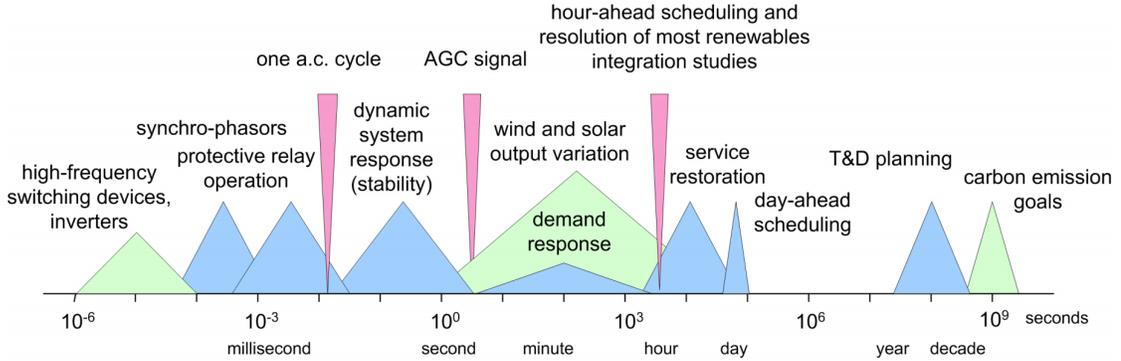


Figure 1.1: The timeline of the electrical market. Figure from [4]

The second and arguably more important issue is the impracticality of rating in real-time with the current electrical system. For many processes, the available transmission capacity needs to be preemptively known. Figure 1.1 shows a timeline of a typical electrical market spanning across 15 orders of magnitude. The RT-DTLR can be assumed to be valid for up to several minutes due to the thermal inertia of the conductor and can be extrapolated to a few hours using statistical methods. However, this only covers processes that facilitate sub-second operations, and respond to demand spikes and emergency situations. Any process that operates on a longer time scale, such as weekly and day-ahead scheduling, therefore cannot benefit from RT-DTLR.

For DTLR to be more valuable within the current infrastructure and system, it needs to be incorporated into longer scale processes that operate in horizons of hours or days. To achieve this, DTLR must be forecasted with sufficient accuracy, precision and reliability.

The focus of the research presented in this thesis is on the innovation and development of medium-term DTLR forecasting to be used with daily to weekly scale processes within electricity markets.

1.2 Thesis Objectives

The objectives for the research presented in this thesis can be arranged into 3 groups as listed below.

1. Design a novel DTLR forecasting system:
 - Review the available literature to find the most promising combination of methods for DTLR prediction.
 - Develop a technique to provide a probabilistic prediction of DTLR.
2. Examine the effects of temporal discretization on DTLR:
 - Develop a process for calculating the temporal discretization of DTLR.
 - Establish a baseline for DTLR when temporal discretization is applied.
3. Improve the performance of discretized DTLR forecasts:
 - Analyze the impact of temporal correlations on transmission line rating predictions.
 - Propose a technique to integrate temporal correlations into the DTLR forecasting system.
 - Evaluate the proposed methods on experimental data.

1.3 Research Originality

The work presented in this thesis is rooted in the present state of the art, however, a substantial portion of the research is novel and original. This section highlights the contributions to the current state of knowledge. These original contributions are significant not only from the academic perspective, but they also promise important improvements to the electric power transmission industry.

Two novel extensions to an established algorithm are proposed in Chapter 3, which describes an algorithm for predicting time series data collected by wireless sensors. The first extension, Delayed-DBP, allows the user to make a compromise between the latency of the algorithm and the compression ratio. An additional configuration option was added to the algorithm and the modified pseudo-code is provided. The second extension aims to decrease the latency, while providing a portion of the compression ratio benefit by performing local prediction using an artificial neural network.

A probabilistic rating prediction system is presented in Chapter 4. The system differs from other DTLR forecasting systems as it concentrates solely on daily forecasts. The majority of available research in DTLR focuses on predicting real-time rating. In this chapter, the limitations of daily rating are clearly outlined and daily rating is compared to realtime rating to highlight the impact of temporal discretization. Moreover, this research uses Model Output Statistics (MOS) with a custom statistical model. While MOS is usually used with ensemble systems, this work applies it to a single NWP with the understanding that the forecast uncertainty is inversely correlated with wind speed.

The concept of temporal discretization is formally introduced in Chapter 5. To the best of the author's knowledge, this topic had not previously been addressed. In this chapter, the implications of temporal discretization on extreme values of winds speed and DTLR are discussed. Additionally, the effects are evaluated on a large scale

simulation and arguments are presented to support the claim that DTLR benefit decreases according to the power law. Finally, the connection between DTLR and reproductivity through formation of the minimum, a not commonly known property of Weibull distribution, is established.

The work in Chapter 6 applies the idea of temporal discretization to DTLR prediction, which prior to this had not been done before. The sensitivity of DTLR prediction to the discretization period length is evaluated. An unexpected result is obtained when the DLR prediction is not very sensitive to the change in discretization period. Furthermore, the DTLR benefit is lower for a 1-hour discretization period than for a 2-hour one. The regression-via-classification framework was applied to DTLR prediction for the first time. To estimate temporally discretized forecasts, a novel method based on Monte Carlo simulation is proposed.

1.4 Thesis Outline

The thesis is structured into 7 chapters.

Chapter 2 - Background and Related Work

In this chapter the related work from the field of DTLR is summarized. This chapter contains background information that is relevant for the subsequent chapters.

Chapter 3 - Derivative Based Prediction with Look Ahead

RT-DTLR requires sensors to be installed on multiple locations on the transmission line. The data must be transmitted nearly in real-time. This chapter presents Derivative Based Prediction with Look Ahead, which is a compression algorithm that reduces the number of messages transmitted through wireless sensor networks. It is an extension to the derivative Based Prediction (DBP) algorithm. The algorithm computes a linear fit over the time series and sends only updates of the linear model to the base station. Two extensions of the original algorithm that further decrease

the number of data packets sent are introduced. The first variant, Delayed DBP, computes the slope of the linear fit using data points in front of and after the model reference point. This is in contrast to basic DBP that uses only the reference point itself. The second extension, DBP with look-ahead, is based on Delayed DBP, but uses a recurrent neural network (RNN) to predict several points from the future for slope computation. All three algorithms have been implemented and simulated on a temperature time series. DBP reduced the number of required transmissions to transfer the entire dataset to 4.4 %, Delayed DBP to 2.6 % while introducing 7.5 minutes of delay and, DBP with look ahead to 4 %.

Chapter 4 - Day-Ahead Dynamic Thermal Line Rating using Numerical Weather Prediction

In this chapter, a DTLR medium-term forecasting system is developed and tested. The rating is set daily based on a Numerical Weather Prediction (NWP) generated the previous day. This method is similar to a seasonal rating, also known as Quasi or Semi-Dynamic line rating, but differs in that it uses a very short time frame of one day instead of a month or quarter year. A statistical model is used to quantify the uncertainty in prediction to assure reliability of the predicted ampacity. The method is tested on measured data from a transmission line located in southern Alberta, Canada.

Chapter 5 - The Effect of Temporal Discretization on Dynamic Thermal Line Rating

DTLR is often applied in discrete time periods during which the rating is held constant. As the applied rating must always be safe under the worst case conditions, it is selected as the minimal rating over the time period. This conservative approach decreases the potential benefit of DTLR because the capacity that is over the minimum of the period is unused. In this chapter the effect of different discretization period lengths is investigated. The stability under the minimization property of Weibull

distribution is used to derive the relation between the discretization length and the expected value of discretized wind speed. Moreover, data from 490 weather stations located across Canada are used to support the analytical result and to quantify the benefit of DTLR over a large geographical area.

Chapter 6 - Probabilistic Forecasting of Dynamic Line Rating with Temporal Correlations

In this chapter, the effects of temporal discretization on rating prediction are examined, the importance of temporal correlations are discussed, and an algorithm that includes these correlations in a rating prediction system is proposed. The sampling algorithm is based on iterative sampling from a chain graphical model, where the values of already sampled variables influence the distribution of their neighbours to emulate the correlations seen in real data. The system is able to produce samples that are usable in a Monte Carlo optimization of an objective function and yield overall good calibration.

Chapter 7 - Conclusion and Final Remarks

The thesis is recapitulated and concluded in this final chapter.

Chapter 2

Background and Related Work

2.1 Background

The purpose of the electrical system is to provide electricity for personal, business, and industrial use. The three main components of the system include generation, transportation, and distribution. The purpose of generation is to convert different forms of energy to electricity. Conventional generation uses fuel to spin generators that provide electrical power, whereas renewable generation utilizes energy that is freely available in the environment like wind, solar, or hydro [5, 6].

After generation, the energy has to be transported to the users. The transmission network achieves this by converting the electricity to higher voltages and sending it through power transmission lines over great distances. Once it reaches its destination, the electricity is converted back to a lower voltage and is delivered to users through

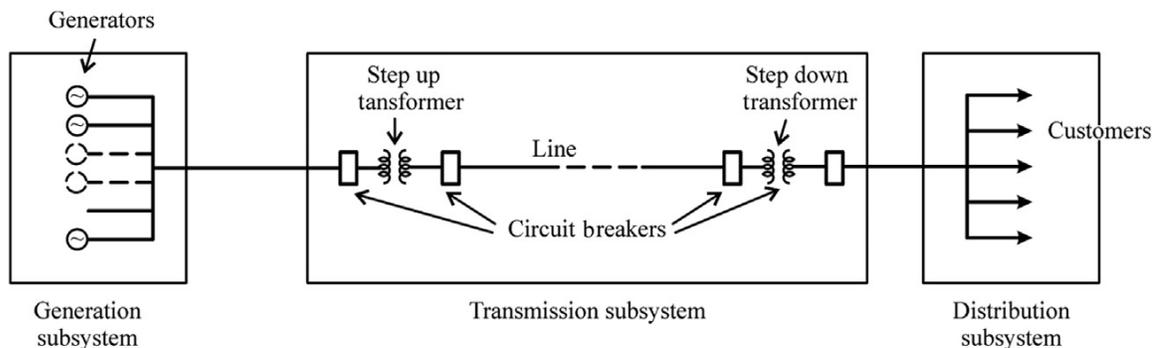


Figure 2.1: Diagram of the electrical system. Image from [6]

a distribution network [5, 6].

Historically, the electrical system was operated by one entity or a small number of parties controlled by the state or local government. In recent years, however, many regions have moved to a free market electrical system, not too different from regular commodity markets. In this free market electrical system, electricity is traded between the parties that produce it (the generation) and those that consume it (the consumers) [7].

Electricity has several properties that make the market unique. Unlike a conventional commodity, electricity cannot be easily stored in large quantities [8]. The production has to be equal to the consumption at all times. This poses a problem, as cheap forms of electricity cannot be produced in advance and used when needed. The cost of generation follows the economies of scale, where larger quantities of energy can be produced at lower unit costs. Furthermore, as larger generators are more difficult to control the cost of electricity is typically proportional to the reaction speed at which the output of the generator can be changed. In other words, generation that can be altered at moments notice is more expensive than generation that takes hours or even days to control [7].

To overcome this limitation, a hierarchy of power generation has been established. There is bulk power generation that covers the baseline energy needs at low costs. These generators produce power at a steady rate and rarely have a change in output. Then there are smaller peaking generators that engage as needed to satisfy any power demand peaks that exceed the limits of the bulk generators.

Another issue with electricity is that the amount of consumption cannot be controlled. Physically, a customer can draw as much or as little electricity as they want regardless of any previous arrangements or contracts [8].

A different distinctive property of the electricity market is that many consumers are generally insensitive to the cost of electricity [7]. In other markets, when the cost of a commodity increases, a certain portion of buyers will decide not to buy, thereby

decreasing demand and subsequently price. Hence, this process provides a degree of self regulation. In the electricity market, however, many customers utilize electricity regardless of the price.

A typical electricity market works on the principle of an auction. The producers and consumers submit their bids and then the operator clears the market to find the price. The bids of the producers are sorted by cost and the lowest cost that satisfies the demand is found. This process guarantees the lower cost energy sources will be used first followed by the more costly generators [7].

After the auction is closed, the price and the total amount of generation is known. It is at this time that the operator of the electrical network can notify the generators of how much energy is needed to be produced at specific times.

The locations of electrical generation and consumption are typically in different geographical areas. Consumption is concentrated around cities, whereas large producers of electricity are often located in remote areas. It is postulated that proximity to fuel or energy sources, as well as concern for pollution from a large power plant may be the rationale for this. Thus, electricity is often transferred over extensive distances.

Additionally, the capacity of transmission lines are not unlimited and several factors are known to restrict this capacity. To achieve sufficient quality and stability over medium distances (80-240 km), good control of electrical flow is needed, and the parasitic parameters of inductance and capacitance should be taken into account. In even longer lines (over 240 km), these parameters must to be treated as distributed, and even greater accuracy of control is required [6]. Figure 2.2 illustrates this relationship.

In short lines (less than 80 km), the limiting factors are simply the thermal properties of the conductor [6]. When electricity flows through the conductor, heat is generated secondary to the effects of electrical and magnetic losses, and the temperature of the conductor rises. The temperature of the conductor has to be capped for three reasons. Firstly, if the temperature of the metal were to rise to its annealing

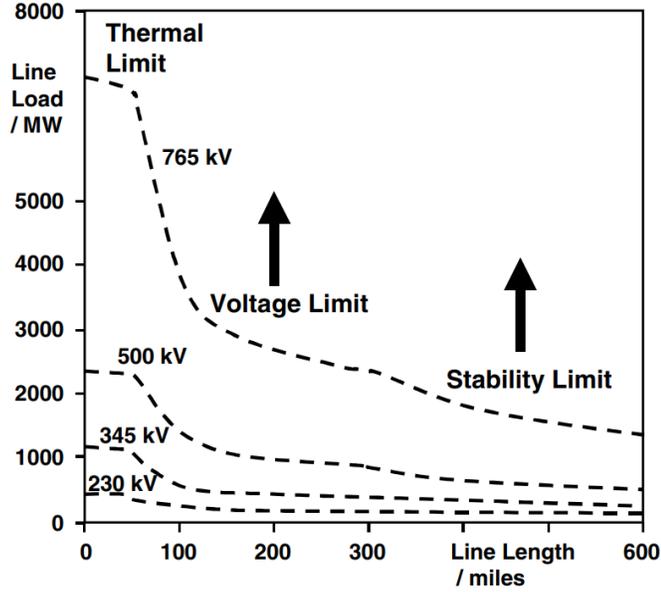


Figure 2.2: The relationship between the length of the transmission line and its rating. Figure from [8].

point, the conductor would abruptly lose tensile strength and break. Secondly, an increased temperature accelerates the ageing of the conductor. The lifespan of a conductor is measured in decades, however, if it is operated over its design temperature the conductor might fail before its projected service time [9]. Finally, an increased temperature causes metal expansion and subsequent loss of tension and increased sag over a segment of the power transmission line. When a transmission line is built, a specific amount of clearance between the conductor and the other object is assumed to accommodate for sag [10]. If, however, the conductor were to sag below the allotted clearance, it may come into contact with foreign objects, which can cause deleterious effects to the line itself or nearby property. In the worst case scenario this may lead to human injury or death.

For the reasons specified above, it is therefore critical to carefully control the amount of power being transmitted to ensure it does not surpass the defined limits. Methods of controlling power transmission are limited. Traditionally, it was not a straightforward task to decrease the power flow in a transmission line, because power

flows through a network following Kirchhoff's circuit laws. Often, the only methods for control are changing the output of the power plants, or shedding load. Recent developments in power electronics, however, has enabled the application of devices that offer a more directed but still limited control over power flow [8].

Ideally, transmission line capacity far exceeds what is required and the market operator needs not direct any special attention to the transmission network besides accounting for transmission losses. On the other hand, if the transmission line is utilized to full capacity, the operator has to schedule the generators in such a way that the flows are lesser than the transmission limits. Typically, this can be achieved by scheduling more generation at other locations to enable the required power to be provided through alternative transmission lines. This ultimately decreases the load on the congested line [7].

Unfortunately though, this solution deviates from the optimal schedule and leads to what is commonly known as local pricing. This is when electricity is sold at an inflated price due to both the inability to access cheaper electricity, and the necessity to use locally available but more expensive generators [7].

The demand for electricity is ever increasing from a 20% share of total energy consumption in 2020 to a projected 24-31% in 2040 [11]. This constant uplift of electricity demand, together with wide spread adaptation of renewable generation, requires adequate investments in the electrical system including the electrical grid. In the next decade, an additional 2 million kilometers of transmission lines and 14 million kilometers of distribution lines are planned to be added [11]. Together with the upgrades for the current grid, the investment is projected to reach \$460 billion in 2030 [11].

Construction of new lines or upgrading existing lines is an extremely costly process that necessitates not only large investments, but also causes interruptions in service for the consumers, and opportunity costs for the producers. To minimize cost, the investment strategy relies on modernization and digitization of current equipment to

increase the capacity of the network by utilizing existing equipment in a more efficient way. There are numerous new technologies in development that aim to optimize the way transmission networks function.

The improvements in development related to demand include utility scale battery electricity storage. Utility scale battery storage is essentially a large battery that is able to provide power levels comparable to the power coming from a utility operator. This storage can be charged when traffic on the electrical grid is low and used later to supply energy when the network is congested. This process of smoothing large surges of demand is called peak shaving and decreases costs by removing the need to use expensive peaking generators [12]. It is projected that there will be 220 GW of battery storage installed by 2030 compared to 11.5 GW in 2019 [11].

The greater utilization of decentralized production mainly from renewable sources can be viewed as an improvement in electricity production. It is estimated, that solar photo-voltaic additions will reach 280 GW by 2030, and wind power generator additions will reach 145 GW [11]. The more consumers that can be satisfied with renewable production of closer proximity, the greater the capacity of the transmission system left free for bulk transfers. This does however bring its own problems, as the transmission and distribution system has been designed for mainly one way flow from large power plants to consumers.

Even the transmission system itself has been a focus of research in recent years. To name a few examples, there are new discoveries in network topology [13]; new optimization techniques for generation and transmission planning [14]; improvements in DC converter technology that allow construction of high voltage DC lines which are more efficient and economically beneficial over long distances [15]; and the introduction of Flexible AC Transmission Systems (FACTS) devices that provide fast and reliable way to modify flows in the network by controlling the voltage, current, or impedance [8].

Dynamic Thermal Line Rating (DLR) is another technology that aims to optimize

the utility of the electrical grid [16]. As mentioned previously, the transmission lines are subjected the operational limits. This limit is termed the rating of the transmission and dictates the maximal current that can flow through the conductor at any given moment. It is a value decided by the operator to ensure equipment safety. For shorter lines (<80km) the rating is governed by the thermal properties of the conductor [8]. The rating is set in such a way, as to avoid the temperature of the conductor rising beyond the safe limit of the material.

Historically, rating has been derived by using estimates of the worst case ambient conditions to calculate the maximal possible current that does not result in overheating. This strategy is called Static Line Rating (SLR) [16]. It is apparent that SLR is not optimal, as worst case conditions rarely occur and therefore, the amount of energy carried by the line for the majority of the time will be well below its physical limit. One of the conditions used in the selection of SLR, as defined by standard CIGRE-299 [17], is that SLR must be lower than the real rating with 99% probability. This means that by definition, a transmission line controlled by SLR will be operating sub-optimally 99% of the time.

It has been well known in the industry that SLR is a largely conservative method that under-utilizes transmission line capacity. The need to overcome the inherent limitations of SLR lead to the proposal of a new innovative strategy in the 1980s known as Dynamic Thermal Line Rating (DTLR). Unlike SLR, this strategy takes information from the current ambient conditions to set the rating of the line precisely to the value that maximizes the capacity all while honoring the thermal limit. In its simplest form, DLR is implemented by installing sensors on the transmission line and using the measurements to estimate the constantly changing rating in real time. Early deployments by the end of 1980s showed that DLR can increase the average capacity by 20-70% over that of SLR [18].

2.1.1 Benefits of DTLR

DTLR provides an increase of capacity to overhead power lines. This additional capacity can reduce congestion and provide additional options to guarantee N-1 security, support wind power integration, and increase economic benefits in general. The following paragraphs mention several figures from other research papers regarding the benefits of DTLR.

Dabbaghjamanesh [19] experiments with scheduling of a reconfigurable microgrid. They show, that employing DTLR instead of SLR provides the necessary capacity to allow the network to function safely even when the grid operates near full capacity. Without DTLR, the scheduling algorithm was unable to find a solution to the flow due to violation of limits in certain cases. Schneider [20] shows that employing DTLR in the European grid could reduce the redispatch volume necessary to satisfy the N-1 condition by 60%.

Wind power integration and DTLR have natural synergy. Wind power plants generate electricity when wind speed is high and, at the same time, the transmission lines receive a lot of cooling. Kazerooni [21] investigated the application of DTLR to support the installation of off-shore wind power plants in the United Kingdom. They concluded DTLR application can decrease the constraint cost by 53% or £1.2 million. Michiorri [22] estimated that deployment of DTLR in a transmission network on Orkney Isles in Scotland could allow installation of an additional 4MW of generation while reducing the Curtailment from 38.5 to 9.7%. Ringelband [23] discovered that by simulating a part of the German power grid with thermal and wind power plants, DTLR could reduce the redispatched energy by 85%. Viafora [24] presents a figure showing a positive correlation between wind power and DTLR. They also conclude that DTLR decreases wind curtailment thereby reducing the cost by 17.6% in their test case.

An analysis of the potential economic benefit gained by application of DTLR to

the power corridor that connects Siberian and European zones in Russia was conducted by Bubenchikov [25]. It was concluded that up to 8% of costs could be saved by transporting more of the cheaper Siberian energy to the European zone by employing DTLR. Bhattarai [26] presented a study where it was concluded that DTLR application on a line in southern Alberta, Canada could increase the average rating by 22%. Additionally, with DTLR this line could also allow for almost doubling of the connected wind power capacity, which would not be possible with SLR. Teng [27] comments on the synergy between DTLR and FACTS and they estimate that the benefit can be increased by additional 30% when these two technologies are employed together.

2.1.2 Thermal Model of Transmission Line

In this section the thermal model that is used to calculate the rating is described. The IEEE-738 model is used throughout this thesis. There are other commonly used models like CIGRE 601 [28], however, these models are very similar [29].

The model from IEEE-738 balances the elements that heat up the conductor (electrical current and solar irradiation) with the elements that typically provide cooling (environmental conditions).

$$q_s + q_j + q_m = q_c + q_r, \quad (2.1)$$

where q_s is heat gained from solar irradiation, q_j is heat generated by the electrical current due to the non-zero resistance of the conductor, q_m is the energy of magnetic losses, and q_c and q_r are the amount of heat lost due to convection and radiation to the environment.

The convective loss, q_c , is divided into two types: natural convection and forced convection. Natural convection is caused by the warm air rising from the conductor and being replaced by cool air. This process occurs even in still-air condition. Forced convection occurs when wind is moving the air around the conductor forcing cool air

into contact with the conductor. Natural convection provides less cooling effect than forced convection, to the point where natural convection is equivalent to around a 0.2 m/s wind speed.

For lower wind speeds, the recommended procedure of calculating convective loss is to evaluate both the natural convection and the forced convection and use the greater of these two values. The forced convection depends on the wind direction and other variables. Wind that is perpendicular provides more cooling than wind that is parallel to the conductor. During the calculation, the wind vector is converted into a so called effective wind speed, which is a perpendicular wind speed that has the same cooling power as the original wind speed vector.

The radiated heat loss, q_r , occurs when the conductor is warmer than its surroundings and radiates heat to the environment. The rate depends primarily on the difference between the conductor temperature and the environment. It is also effected by the emissivity, which is a material property of the conductor. Emissivity is a number between 0 and 1 and represents the emission *efficiency* of the material, with 0 and 1 being no emission and full emission, respectively.

The heat gained from solar irradiation, q_s , can be calculated from the amount of solar energy reaching the conductor, the conductor shape and orientation towards the sun, and the absorptivity of the conductor material. The solar heat is calculated from the Sun's position in the sky, the solar constant, and the portion of the light that is passed through the atmosphere.

The heat, q_j , generated by the electrical current is calculated from Ohm's law. The conductor resistance depends on the material properties of the conductor. Typically, it depends on the conductor temperature through a temperature coefficient. In the case of conductors with a steel core, the resistance is a mixture of properties from both materials. The conductor heat capacity determines the temperature increase for a given gain in heat. A conductor with higher heat capacity needs more heat to increase its temperature. The heat capacity also effects the dynamic properties of the

conductor. For a higher heat capacity, more heat transfer is necessary to change the temperature and therefore, the conductor will have more thermal momentum.

The exact formulas to calculate the quantities in the thermal model can be found in the standard IEEE-738 [16].

The thermal rating can be calculated from the thermal model by separating the current I from q_j and setting the conductor temperature T to the maximal temperature of the conductor.

$$I = \sqrt{\frac{q_c + q_r - q_s}{RT}} \quad (2.2)$$

The conductor temperature cannot be directly calculated, but can be estimated through numerical integration of the following equation:

$$q_c + q_r + m \cdot C_p \cdot \frac{dT}{dt} = q_s + I^2 \cdot RT \quad (2.3)$$

2.1.3 Static Thermal Line Rating

Static Thermal Line Rating (SLR) is a conservative rating strategy that ensures safe thermal operation of a transmission line by assuming the worst case ambient conditions for rating calculation. It is based on the assumption that the worst case scenario is rare and therefore, the line will be operating safely most of the time.

CIGRE-299 [17] provides the requirements for rating as a) the average temperature of a line segment does not exceed the operating temperature by more than 10 °C, b) at least 99% confidence that the line temperature is within range, and c) the spot temperature does not exceed the design temperature by more than 20 °C.

2.1.4 Time Series

Time series can be described using a number of different models. One can differentiate between models based on knowledge of the underlying system and non-parametric models that are generated based on the time series itself [30, 31]. As the former

approach requires a deep understanding of the measured signal and the system, it is not universal and reusable for different variables. Non-parametric linear models for time series prediction include autoregressive, moving-average or autoregressive moving-average (ARMA) models. As these models are linear, they do not perform well when the nature of the underlying system is non-linear.

Time series prediction is the task of finding a function $y_t = \phi(y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots)$ that will take the history of the output signal y along with some additional variables x , and generate the next value of y [32]. Variables x , called exogenous inputs, are used if they can provide any correlation with the output signal. If x is omitted, then the model (or network) is purely autoregressive. If the values of y are not included, then the function transforms one time series to another.

2.1.5 Extreme Values of Time Series

In theory [33], the maximum of n arbitrary but independent random variables $S = \{S_1, \dots, S_n\}$ with cumulative distribution functions F_1, \dots, F_N can be simply derived as

$$P(\max(S) \leq t) = \prod_i P(X_i \leq t) = \prod_i F_i(t). \quad (2.4)$$

The minimum can then be found by exploiting the duality between minimum and maximum:

$$P(\min(S) \leq t) = 1 - \prod_i (1 - F_i(t)). \quad (2.5)$$

It is tempting to apply this equation to the probabilistic forecast to directly calculate the distribution of the minimum value over a time period. This equation though, assumes the random variables are independent which certainly does not hold for time series data.

This can be illustrated by examining the equation for the distribution of maxima when all variables are identically distributed. $F(t)^n$ depends on the number of vari-

ables n . Now assume a real, temporally correlated time series has been sampled at n timestamps and then $2n$ timestamps over the same time period. Intuitively, the distribution of the maxima should not change based on the number of samples, but the equations result in $F(t)^n$ and $F(t)^{2n}$. This means the estimation is largely skewed toward the extreme. The same principle holds for the minimum, as ignoring the correlations between random variables will make the extreme values more extreme.

2.2 Related Work

There are numerous studies on the topic of DTLR. A good overview of topics related to DTLR is a review by Karimi [34]. Forecasting DTLR is typically divided into groups by length of the forecasting horizon. There are short term forecasts, typically up to 4 or 6 hours, which are often based on statistical analysis of time series. These forecasts can be used for contingency planning and allow the operator to adjust the N-1 short-term contingency plan based on the long-term emergency capacity available through DTLR [35, 36]. Longer forecasts with a horizon of 24-48 hours are often used with day-ahead markets, which need to know the available capacity in the near future. Methods that produce longer forecasts tend to require a general weather forecast obtained through NWP [35, 36].

A commercially available NWP is used to forecast the safe value of thermal rating for the next day by Uski [37]. In this study, only the ambient temperature forecast is used as the input parameter. The other variables are considered too uncertain and excluded from calculations. Aznarte [38] uses several machine learning methods to estimate the ampacity from a NWP. One of the methods used, Quantile Regression Forest, produces probabilistic predictions. Unlike the majority of research, this study uses machine learning methods to directly estimate the rating from NWP without the use of a thermal model. This makes the ampacity calculation purely a machine learning method. Another study by Taillardat [39] uses machine learning to calibrate the forecast. Ringelband [40] uses Monte Carlo simulation to derive a time series of

probabilistic ampacities from an ensemble NWP calibrated by a probabilistic mixture model.

The work by Dupin [3, 41] is especially notable because it combines several state of the art approaches and new methods into a comprehensive rating system. Four different machine learning methods for converting NWP into a probabilistic rating are compared and Quantile Regression Forest is determined to outperform the others. The authors then present two methods for rating selection from probabilistic prediction. The first method is static quantile selection and the second method is a risk assessment strategy. The risk assessment strategy is a method that allows the operator of the line to select the rating based on how much risk they are willing to accept. It works by finding the rating that makes the expected value of the risk function equal to the accepted risk. The cost function aggregates the costs of lost supply in the case of failure, damages, and accelerated ageing. The evaluation of the cost function requires the temperature of the conductor, which is a value that is impossible to calculate exactly from only the rating. The authors solve this issue by using a lookup table of previous ratings and weather measurements, where they select a row with a similar rating and then use the ambient parameters to estimate the line temperature.

Chapter 3

Derivative Based Prediction with Look Ahead

One of the most important features of wireless sensor networks (WSN) is the ease of their deployment and operation. User want to deploy the network at a deployment site without the need to provide any infrastructure like communication lines or power cables. They do not want to spend time manually recording the data from individual nodes. They want to access the data from the comfort of their office as long as possible, without need for frequent battery replacement.

There have been a number of different techniques developed to meet these requirements. The installation has been simplified using wireless communication and by powering the devices from batteries and energy harvesting devices [42]. There are various self-organizing routing protocols, that minimize the effort for an initial setting of the network, so that user can place the devices at a desired locations and turns them on [43]. The issues of limited range of wireless transmitters have been overcome using multi-hop networks that use short range nodes to relay the packets through the network to enable for longer virtual communication channels. The data from the nodes is collected at a base unit that can be connected to the Internet for convenient access to the data.

The length of the life of a network depends primarily on the energy available to power the nodes and on their energy needs. A typical node is equipped with batteries,

optionally supported by energy harvesting devices. The solar energy is very popular due to its high power density and simplicity of use. Although significant advances have been made in the areas of energy storage and harvesting, any reduction of node energy consumption results in further extension of the network useful life. It has been shown that nodes expend most energy towards wireless transmission. Wireless transceivers are the most power hungry component of WSNs, usually exceeding the power consumption of the remaining parts.

Therefore, it is important to minimize the time when the radio is waiting for, receiving, or transmitting data. There are many levels at which the efficiency of power use by wireless modules can be improved [44, 45]: a well-suited medium access protocol can save energy through duty-cycling the transceiver active time; a smart routing algorithm can forward the data efficiently through the network with a small amount of service data and maintenance communications; and, in the application layer, the volume of data to be transferred can be minimized through an appropriate choice of sampling frequency and avoidance of redundant data transmissions.

There are several approaches to data reduction, including in-network data processing (e.g. event detection triggering data transfer in an event monitoring scenario) and data compression. Another option is data prediction, when a data model is build (either on the node or on the base station) and data queries are retrieved from this model.

Derivative Based Prediction (DBP) [46] lies between compression and prediction. It builds a piecewise linear model of the measured data that is later used to recover the signal. The model is build by the node, and only model parameters are transferred to the base station. As a single model is transferred instead of a number of data samples, a significant data compression can be achieved because the model itself consists only of the parameters for the linear approximator.

Authors of the original DBP model showed that in routed networks, any further reduction of the number of transferred messages would be ineffective [47]. Nodes in

such networks periodically wake up to maintain network layer and forward packages from other nodes, and corresponding service and forwarded packets are responsible for most communication and thus cause the greatest energy expenses.

However, in networks without routing layers (typically networks with the star topology, where each node has a direct connection to the sink), nodes can benefit from any additional decrease of the number of messages that need to be sent.

This is the case of our network where nodes send measured values directly to the sink using a powerful radio capable of transmissions over a kilometer range. The devices wake up periodically, measure the value and transmit the data to a sink. The transmitter current during the transmission can go up to 300 mA, which is very expensive compared to sub mA currents required for actual measurements. In comparison with many other nodes used in wireless sensor networks, the node uses ARM Cortex M4 processor with a floating point calculation unit that can execute relatively advanced algorithms in very short times.

DBP has already been used for temperature prediction by Pötsch et al. [48]. Authors simulated the algorithm on several temperature time series with different sample periods and showed, that the algorithm is able to suppress more than 95 % transmissions.

In this paper, an improvement over the DBP algorithm will be presented. We will describe the basic DBP algorithm and two novel variants that aim to further reduce the number of generated models. Also results from simulation of the three algorithms will be given together with the comparison among them.

3.1 Methods

3.1.1 Artificial Neural Networks

An alternative to the classical models of time series is the use artificial neural networks (ANNs) [49]. A well defined ANN is a powerful system, which can be used

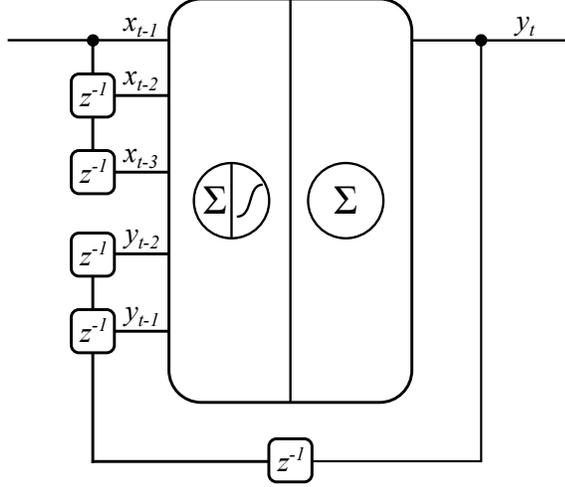


Figure 3.1: A schematic diagram of the NARX network. A block z^{-1} symbolizes one time delay.

to approximate any function if the training algorithm is provided with a sufficient amount of examples.

ANNs can be directly used to approximate function ϕ from data. Consider the following network with linear output layer and L hidden neurons in the middle layer with the activation function f

$$y_t = \sum_{i=1}^I W_i f \left(\sum_{j=1}^p w_{ij} x_{t-j} + \sum_{j=1}^q w'_{ij} y_{t-j} + b_i \right) + B$$

where y_t is the output signal, x_t is an additional exogenous signal, W are weights of the output layer, w_{ij} and w'_{ij} are the weights for the hidden layer, and B and b_{ij} are corresponding biases. A network defined as above, but without any additional variables, is used throughout this paper. The activation function f is hyperbolic tangent.

A network with this particular topology is called Nonlinear Autoregressive network with eXogenous input (NARX) and its block diagram can be seen on Figure 3.1.

One forward step through this network predicts the value of the output one step ahead. If the generated output is taken and fed back to the input of the network, subsequent prediction is obtained. This approach, called multi-step ahead, allows to

generate time series of arbitrary lengths [30].

3.1.2 Genetic Algorithms

Genetic algorithm (GA) is an optimization technique inspired by the genetic evolution of living organisms [50]. There is a population of individuals that represent potential solutions. Individuals can mate using a crossover function which mixes together the properties of parents to create an offspring. They can be also mutated using a mutation operator which changes a portion of the properties of the individual. Both operations happen stochastically with defined rates. After producing a new generation, the fitness of the individuals in the updated population is evaluated and a new population is selected, giving higher chance to better performers. The crossover operator aims to exploit the properties already found by combining them together and the mutation operator introduces new features to the population, thus exploring the solution space.

DBP algorithm and the two proposed variants have several parameters, that have to be tuned for the best performance. Therefore, in this work, GA is used for optimization of the parameters so that the comparison between algorithms is fair. This choice is driven by the flexibility of GAs that place no constraints on the fitness function and on the solution space. In our case, the fitness function is evaluated using simulations of the proposed algorithms, and the solutions are the parameters of the algorithms with defined minimum, maximum, and integral constraints.

The GA used in this work is MI-LXPM [51] algorithm designed for mixed integer problems and implemented in the MATLAB Optimization toolbox. It uses a modified Laplace crossover and Power mutation that satisfy the integer constraint of variables. Offsprings are selected using binary tournament selection function with elitism of 5 % of the population size, which was set to 40.

3.2 Improved DBP Algorithms

3.2.1 Original DBP

DBP is a lossy data compression algorithm that approximates a time series with a sequence of linear models/lines [47, 52]. Measured samples are collected into a buffer. Once the buffer is full, the slope, a , of corresponding line is computed from the first and the last data point in the buffer. This slope, together with the sample, b , and the timestamp, t_0 , form the following linear model

$$y = a \cdot (t - t_0) + b. \quad (3.1)$$

The model output is then compared against the new measurements. As long as the difference between the data predicted by the model and the actual samples does not exceed a predefined threshold, the model remains valid. On the other hand, if the error is higher than the threshold, the old model is invalidated and a new set of model parameters is generated.

It is worth to note that this data compression algorithm works in real time. Once the model is generated, it is transmitted over the communication channel to the sink and used to replicate the time series. This allows to define the compression ratio, C , for real-time communication in terms of the number of transmissions (assuming that one transmission is needed to transfer one sample or one model update)

$$C = \frac{\textit{number of model updates}}{\textit{number of samples}}.$$

DBP uses 4 parameters. First, there is the buffer length, m , which indicates how far into the history to look for the first sample to compute the slope. Second, since the data can have some noise, several samples around the beginning and end of the buffer are averaged to compute the slope. The length of this average is given as parameter l .

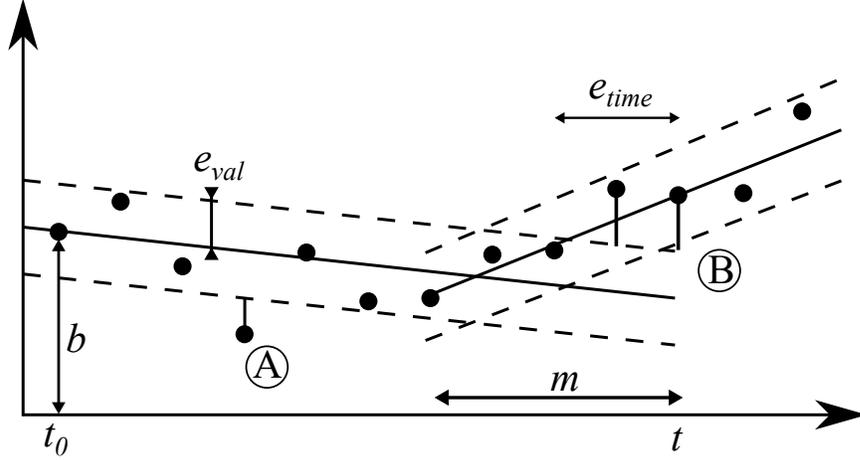


Figure 3.2: An example of DBP fitting of a time series. The circles represent data samples, the solid line is the linear approximation, and the dashed lines are the absolute error margins. The first model was generated at time t_0 with the slope pointing slightly downwards. The event A did not trigger a new model generation, since e_{time} is set to 1 – the algorithm waits for two consecutive errors. At event B , a new model with slope pointing upwards was generated using the sample at time t and $t - m + 1$.

The last two parameters control when a new model is generated. They are the value error, e_{val} , and the time error, e_{time} . The value error has a unit of the measured value and gives the threshold for the maximal difference between the predicted value and the real value after which the model is assumed to be inaccurate. The time error has a unit of time and determines how long an inaccurate model can be tolerated. For instance, if $e_{time} = 1$ then the algorithm will ignore one error and a new model will be generated only after detecting a sequence of two inaccurate values.

To formalize the slope computation, a window is defined over which to generate the slope. The window starts and stops at indexes l_{idx} and r_{idx} . The model is generated for the data sample numbered x_{idx} . For DBP, these values are evaluated as follows

$$l_{idx} = t - m + 1$$

$$r_{idx} = t$$

$$x_{idx} = t$$

The slope, a , of the model at the time x_{idx} is then computed as

$$a = (s(r_{idx}) - s(l_{idx})) / (m - 1)$$

where $s(t)$ is the data sample at time t .

Please note that the averaging of the values around the beginning and the end of the buffer has been omitted from the illustrations and algorithm description. However, it has been implemented and used for every simulation in this paper. The left average takes the first element in the buffer plus $left_average - 1$ elements in front of it, and the right one takes the last element in the buffer plus $right_average - 1$ elements after it. The lengths of both averages were optimized using GA together with other parameters.

3.2.2 Delayed DBP

In the DBP algorithm described above, the prediction and the actual data is compared at the time t . If the invalidity of the model is detected, the algorithm takes the current value as the second point for the computation of the slope. This way, the generated slope is for a line that starts in the past and stops at the current sample. In order to generate a line that better reflects the trend present in the data, it would be better to take the slope of a line which starts in the past and ends in the future.

An example of this behavior is shown in Figure 3.3. After the temperature stops rising, the DBP algorithm keeps assuming the high slope. Delayed DBP *is aware of* the future and generates a more appropriate model. This results in a lower number of model updates.

In order to achieve this, the time series can be delayed by several samples. As before, the current data sample would be used to compute the slope, but for a point in the past. An illustration and comparison of *Delayed DBP* with DBP is shown in Figure 3.4. A new parameter *delay* determines by how many steps is the point (at which the model is evaluated and generated) moved to the past. This modification

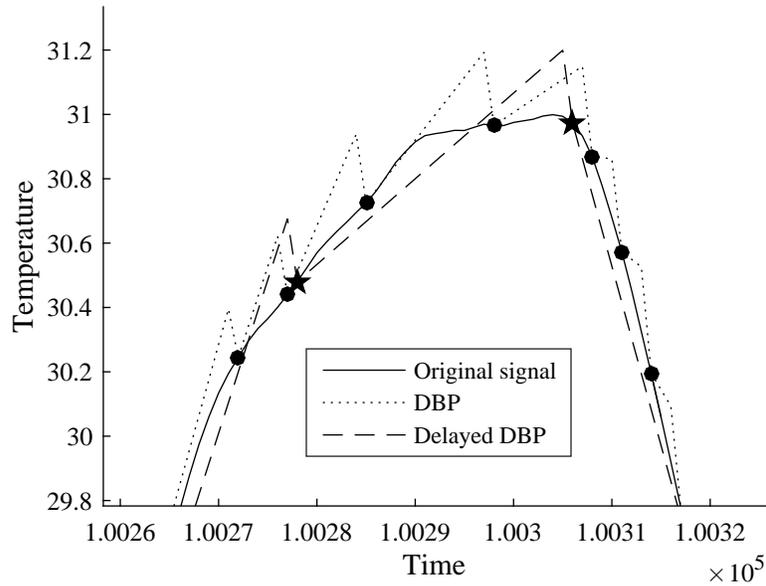


Figure 3.3: Comparison of the output of DBP and Delayed DBP with the same error tolerances. The circles are updates of the model for DBP, while the stars are updates for Delayed DBP.

cannot work in the real-time, as several data points must be collected in advance so that the slope can be computed.

The indexes for the computational window follows

$$l_{idx} = t - m + 1$$

$$r_{idx} = t$$

$$x_{idx} = t - delay$$

3.2.3 DBP with look ahead

The third variant of the algorithm is *DBP with look ahead*. In this case, the model generation point remains in the current time, but the right edge of the slope computing window is moved to the future. The new parameter for this algorithm is *look_ahead* which tells us how far to the right the window should be moved. Since using the true future values is impossible we will use a recurrent neural network to predict number

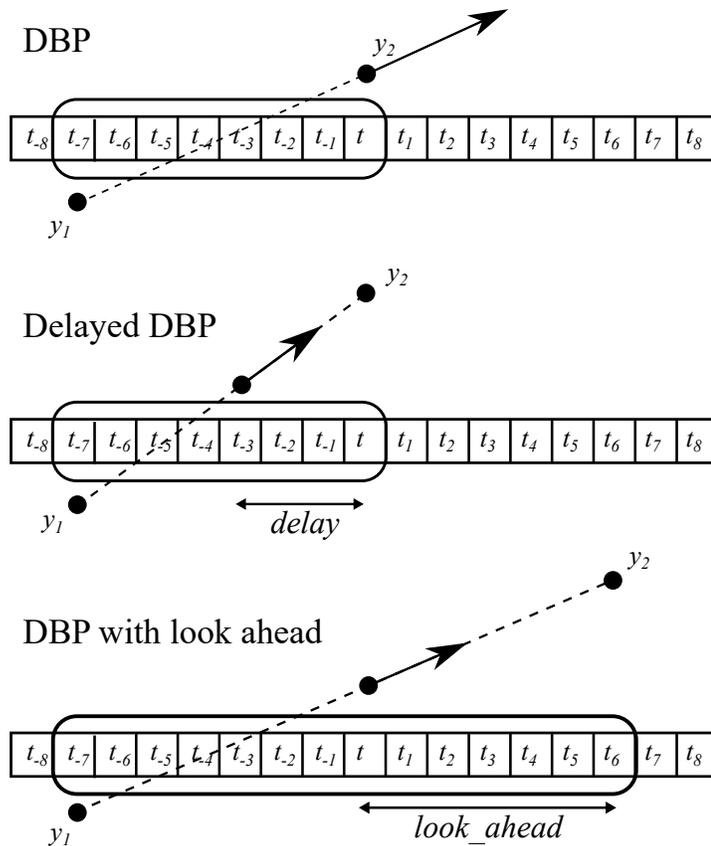


Figure 3.4: All three variants of DBP. Standard DBP uses samples y_1 and y_2 at times $t - m + 1$ and t to compute the slope for the model generated at time t . The delayed DBP uses the same data points to calculate the slope of the model at time $t - delay$. And finally, DBP with look ahead uses points at times $t - m + look_ahead$ and $t + look_ahead$ to compute the model at time t .

of future samples that will be used to perform the slope calculation.

The indexes for the computational window follow

$$l_{idx} = t - m + 1 + look_ahead$$

$$r_{idx} = t + look_ahead$$

$$x_{idx} = t$$

An important observation is that *Delayed DBP* and the *DBP with look ahead* using an ideal prediction give the same results, except for the time shift. This can be shown by letting $look_ahead = delay$ and performing the following substitutions in the equations for *Delayed DBP*

$$l_{idx} = t - m + 1 + (-delay + delay)$$

$$= (t - delay) - m + 1 + look_ahead$$

$$r_{idx} = t + (-delay + delay)$$

$$= (t - delay) + look_ahead$$

$$x_{idx} = (t - delay)$$

Indeed, the results correspond to the equations for DBP with look ahead shifted by $delay$ time steps. Therefore, performance of *Delayed DBP* can be used as an upper limit of compression ratio improvement for the look ahead variant. It performs like *DBP with look ahead* using the *perfect prediction*.

The general procedure is defined in Algorithm 1. It uses parameter p that specifies one of the three variants of the algorithm. If $p < 0$ then the algorithm behaves as *Delayed DBP* with the delay of $-p$, if $p = 0$ then it is the basic DBP, and finally with $p > 0$ the look ahead variant is used with $look_ahead = p$.

Algorithm 1 Universal DBP

```
1: function DBP( $s, e_{val}, e_{time}, m, p$ )
Require: A function  $Predict(n)$  that predicts  $n$  values.
Require: A function  $Compute\_slope(s)$  that returns the slope of the sequence  $s$ .
2:    $delay \leftarrow -1 \cdot \min(0, p)$ 
3:    $look\_ahead \leftarrow \max(0, p)$ 

4:    $start\_pos \leftarrow m - look\_ahead$ 
5:    $a \leftarrow 0, x_0 \leftarrow 0, b \leftarrow \text{inf}$  ▷ Initialize an empty model
6:    $err\_cnt \leftarrow 0$ 

7:   for  $i \leftarrow start\_pos, length(s)$  do
8:      $z \leftarrow a \cdot (i - x_0) + b$  ▷ Generate the value from the model
9:     if  $abs(s[i - delay] - z) \geq e_{val}$  then
10:       $Inc(err\_cnt)$ 
11:    end if
12:    if  $err\_cnt > e_{time}$  then
13:       $seq_l \leftarrow s[i - m + 1, i]$ 
14:       $seq_r \leftarrow Predict(look\_ahead)$ 
15:       $a \leftarrow Compute\_slope(seq_l|seq_r)$ 
16:       $x_0 \leftarrow i - delay$ 
17:       $b \leftarrow x[i - delay]$ 
18:       $err\_cnt \leftarrow 0$ 
19:    end if
20:  end for
21: end function
```

3.3 Evaluation

All three algorithms have been implemented and evaluated on two temperature time series. The first dataset was sampled with 15 second sampling period and has 102401 points. The second dataset has sampling period of 30 seconds and has 128991 points. This covers 17.8 and 44.8 days, respectively. The e_{val} maximal absolute error parameter has been set to 0.25 °C and the e_{time} number of consecutive errors has been set to 0 as we want to have a guarantee on the error.

Because the maximal allowed error is fixed by the parameter e_{val} , standard error-based measures cannot be used to evaluate the performance of the algorithm. Instead, it can be assessed based on the number of models generated to cover the entire time series. The smaller the number of generated models, the better the algorithm performance.

To perform a fair comparison between the algorithms the parameters have been optimized using genetic algorithm. Each individual from the population represents a vector of parameters $(m, averages, look_ahead, delay)$ with $look_ahead$ and $delay$ forced to 0 as appropriate for given algorithm variant. The fitness of an individual is the number of models that have been generated by running the algorithm on the time series.

Figure 3.5 shows how *Delayed DBP* compares to the basic DBP for various values of the delay parameter. As the delayed algorithm practically uses samples from the future, a decrease of the number of generated models is expected. On both time series the *Delayed DBP* was able to decrease the number of generated models to 60 % of the basic algorithm. This performance was achieved with the delay of 15 and 30 time delays (translates to 7.5 minutes) for sampling periods 15 and 30 seconds, respectively. With one third of the delay (5 and 10 time steps) the algorithm yields a reduction to 70 %.

In order to apply the modified algorithm in real-time, a RANN is trained to predict

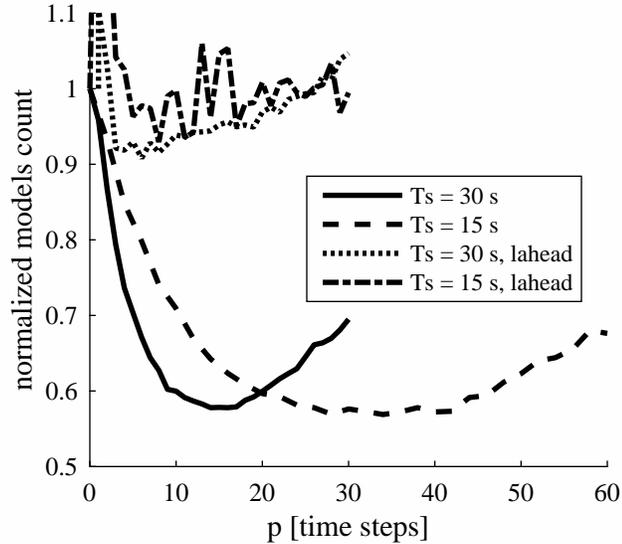


Figure 3.5: The number of models generated by *Delayed DBP* and *DBP with look ahead* on the testing time series relative to the number of models generated by basic DBP depending on the length of the time delay or look ahead. Other parameters of the algorithm have been optimized by GA.

the samples needed to compute the look-ahead. The network has a 15 step input delay and 5 hidden neurons for both time series. The network was trained in an open loop configuration, receiving true data from the time series both as inputs and the desired output (the next data point). After training, the loop was closed and the output of the network was fed back to the input. This allows multi-step time series prediction.

The parameter values found by GA for the delayed DBP have been used for evaluation of the DBP with look ahead, except for the use of predicted values as the look-ahead part of the input data. The results in terms of the model counts can be found in the Figure 3.5. For the sampling period of 30 seconds the neural network aided algorithm was able to reduce the model count to 91 % compared to basic DBP using *look_ahead* of 6. For the 15 second period time series the suppression was to 93 %.

The table 3.1 summarizes the performance of each algorithm on both time series.

Table 3.1: The performance of DBP and its variants on two temperature series.

T= 15 s	# of models	C	m	p
raw time series	102,401			
DBP	2,310	2.23 %	11	
Delayed DBP	1,331	1.30 %	30	-30
DBP with LA	2,147	2.10%	7	8
T= 30	# of models	C	m	p
raw time series	128,991			
DBP	5,686	4.41 %	6	
Delayed DBP	3,288	2.55 %	15	-15
DBP with LA	5,164	4.00 %	5	6

3.4 Conclusions

This paper presents an improvement over DBP, an algorithm for real-time data compression by fitting the time series with linear approximations computed from a close history. The basic DBP prediction has been simulated and evaluated on two temperature datasets collected with 15 and 30 seconds sampling periods. The algorithm reduced the number of data transmissions to 2.2 and 4.4 %, respectively, while keeping the maximal absolute error below 0.25 °C.

The first contribution is *Delayed DBP* that postpones the time series by several samples and then computes the slope of the linear approximator not only from the historical but also using the data points corresponding to the close future. The algorithm was simulated on two temperature data sets with sampling periods of 30 and 15 seconds and the best result was a reduction to 60% of model updates compared to the basic DBP in exchange for introducing a delay of 7.5 minutes. If the delay was set to 2.5 minutes, the reduction dropped slightly to 70%.

The second contribution aims to return the improved algorithm back to its real-

time functionality. Instead of delaying the time series, future data points are predicted using a RNN. Using a network with 5 hidden neurons, the reduction on the dataset sampled with 30 seconds period was to 91% model updates of the basic algorithm when predicting 6 samples.

The future work will focus on improving the prediction accuracy of the network for the DBP with look ahead. The used time series has multiple issues like the seasonality of the data, non-stationarity, and long time dependences, that are decreasing the performance of the network and must be targeted in order to improve the prediction accuracy. The negative effects of seasonality and non-stationarity can be mitigated using proper statistical techniques designed for preprocessing the time series data, while the long time dependencies could be tackled using an advanced neural network architecture like long-term short-term memory (LTSTM), which is specially designed to overcome the vanishing gradient problem found in time series with long time dependencies.

Another possible direction of the future work that will be explored is to replace the linear model with a different function, or possibly with a set of functions. For example the evaluated temperature time series typically do not change in a linear fashion. According to the Newton's law of cooling the instantaneous rate of change of the temperature is proportional to the temperature difference between the object and its surroundings which results in the temperature signal having an exponential form. Therefore the linear model could be replaced with a parametrized model to represent a convex, concave or linear shape depending on a parameter.

Also a more thorough simulation is required. This include testing the algorithms on different types of signals, for instance on relative humidity or soil moisture data, and on signals with different information content – signals that include some high frequency components. In addition, performance under other sampling periods should be examined. Last but not least, the computational, memory, and energy requirements of the algorithms must be evaluated. A more elaborate technique for the slope

prediction might prove to have very high energy demands that would diminish the energy savings gained from the transmission count reduction.

In conclusion, if the user can accept the introduction of a small time delay to the output data stream, the proposed delayed DBP can save a significant amount of data transmissions – up to 40% on the time series used for testing. On the other hand, if the real-time property of the system is required, using a suitable time series prediction technique (such as a recurrent neural network) can result in an improvement of the compression ratio provided by DBP.

Chapter 4

Day-Ahead Dynamic Thermal Line Rating using Numerical Weather Prediction

Transmission networks play a key role in the electrical system, through the delivery of electricity from generators to distributions networks. As the demand for electricity increases and new generators are installed, the transmission lines are required to transport an increasing amount of energy. However, transmission networks have limited capacity and can reach a limit where no further expansion of the system is possible without upgrading the network.

The maximal electrical current that can safely flow through a transmission line is called ampacity or rating. For shorter lines, this limit is determined by the thermal properties of the conductor, which has to be kept under its maximal allowed temperature – the operating temperature. If the operating temperature is exceeded, the conductor can sag or become damaged. Therefore, the ampacity is determined to be the current that would cause the temperature to rise to the operating temperature [16].

Besides electrical current, environmental factors such as air temperature, wind speed, and solar irradiation directly affect the temperature of the conductor. This must be taken into account when determining the ampacity of a transmission line. The details of the thermal behaviour are captured in the standard IEEE-738, which

provides the thermal equations for bare over-head conductors [16].

The ampacity of a transmission line is typically calculated assuming the worst possible ambient conditions. Explicitly stated, the day with the highest temperature and lowest wind speed would be used to calculate the ampacity for the entire season. Since environmental conditions are often more favourable than the worst-case scenario, this approach wastes available capacity. This approach is called Static Line Rating (SLR) and the guideline on selecting appropriate conditions can be found in [17]. Figure 4.1 shows a histogram of the real-time ampacity versus the static limit that is used on the transmission line studied in this paper. It clearly illustrates how a large portion of the ampacity is not used, because the real-time ampacity is usually higher than the static limit.

Dynamic Line Rating (DLR) overcomes the inefficiencies of SLR by modifying the ampacity in real-time according to the current weather conditions. In the ideal DLR scenario, a transmission line is equipped with sensors capable of measuring the weather conditions and temperature of the conductor. The data captured is transmitted to the control center, where the real-time ampacity is calculated and can be applied to power lines for full optimization of capacity. A recent review [34] summarizes various DTLR techniques.

Studies have shown that DLR is superior to SLR with regards to both efficiency and optimal utilization of the transmission line [21, 53]. This more complex system, however, requires some additional considerations. For example, the network of sensors and associated communication infrastructure needs ongoing maintenance, and support must be provided by the Energy Management System (EMS) of the network operator.

A major transmission line operator in Alberta, Canada is interested in applying DLR to some transmission lines to enable installation of more wind power plants. DLR is well suited for this application because of the positive correlation between power output from wind generators and transmission line capacity [54]. However,

the EMS system that is used by the utility operator cannot easily ingest real-time changes of the ampacity.

The proposed method addresses these issues through the use of Numerical Weather Prediction (NWP) to forecast the weather conditions for the next day to calculate an ampacity time series. To be conservative, the minimum ampacity would be selected and applied to the transmission line in the EMS. The reliability of the predicted line rating is assured by quantifying the uncertainty of the weather forecast through statistical modelling of past forecast errors, and subsequent propagation through the thermal model by Monte Carlo simulation. Finally, the ampacity that complies with the selected level of confidence is determined. For example, if an ampacity with 97% confidence is selected, then in 97% of cases the predicted ampacity will be lower than the real ampacity.

Given that the forecast has an acceptable level of reliability, sensors are not required. The ampacity is manually changed daily and remains valid for 24 hours. Thus, the problem of the EMS system not supporting real-time change of ampacity would be overcome. The proposed method is feasible, as the utility operator already uses seasonal line rating [55], which involves inputting ampacities several times per year. NWP was used due to its high forecasting skill for medium range predictions. Other DTLR forecasting methods can be found in a review by Dupin [56].

Related Work

In a similar study by Uski [37], a commercially available NWP is used to forecast the safe value of thermal rating for the next day. Only the temperature forecast with a safety margin, derived from historical predictions is used. Other variables are not considered due to their high uncertainty. Aznarte [38] evaluates several machine learning methods to calculate the ampacity from a NWP. Notably, one of the methods used is Quantile Regression Forest, which is a method capable of probabilistic predictions. The thermal model is used only to generate the training dataset and not for the actual

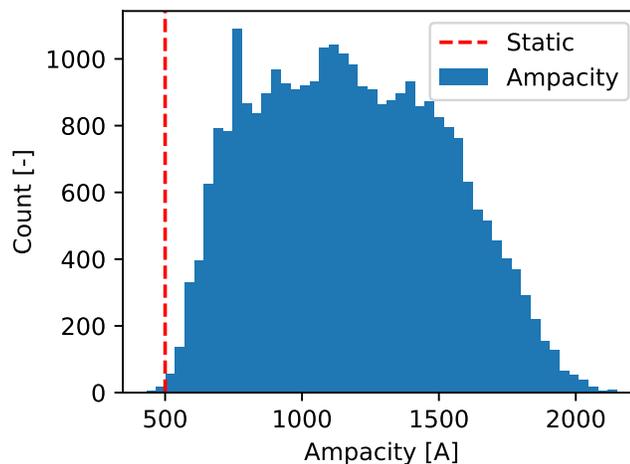


Figure 4.1: Histogram of the real-time ampacity of the studied transmission line in 2016. The static limit is smaller than the actual limit most of the time.

prediction. This makes the ampacity calculation purely a machine learning method. In the work by Ringelband [40], a time series of probabilistic ampacities is derived by Monte Carlo calculations from an ensemble NWP calibrated by a probabilistic mixture model. Ensemble NWP provides information about the uncertainty by running the NWP multiple times with different initial conditions emulating the randomness in forecasts. However, ensemble forecasting is computationally demanding and not easily obtainable. Another study using NWP ensembles was done by Taillardat [39], where machine learning is used to calibrate the forecast as opposed to probabilistic models.

Notation

The following symbols are used for the physical quantities:

T air temperature [$^{\circ}C$],

v wind speed [m/s],

d wind direction [$^{\circ}$],

s solar irradiation [W/m^2],

A ampacity [A].

Moreover the *hat notation* is used for estimated quantities like \hat{T} or \hat{v} for estimated (predicted) temperature or wind speed, respectively.

The following symbols are used for other variables:

N number of days,

D current day (the present),

t forecast lead time [hours],

M number of Monte Carlo samples.

4.1 Methods

The calculations used for analysis will be introduced in this section. First, the thermal model is described. Then the probabilistic ampacity prediction method is presented, along with several alternative ampacity calculation methods for comparison.

4.1.1 Probabilistic Prediction

The NWP system used in this study is the Weather Research and Forecasting (WRF) Model. This model produces a deterministic weather forecast, meaning it provides the forecast in the form of a time series. This forecast does not come with any form of uncertainty and thus, the reliability of the prediction cannot be directly evaluated. To convert the deterministic forecast to a probabilistic forecast Model Output Statistics (MOS) [57] are used. In MOS, historical forecasts together with measurements are used to learn the model of uncertainty and obtain probabilistic forecasts.

The complete algorithm to issue an ampacity forecast for day D is outlined in Algorithm 2. Described here briefly, N past weather forecasts together with the measurements for days $D - N$ to $D - 1$ are loaded from the database and then used to learn the parameters of the probabilistic models for each variable. The NWP model

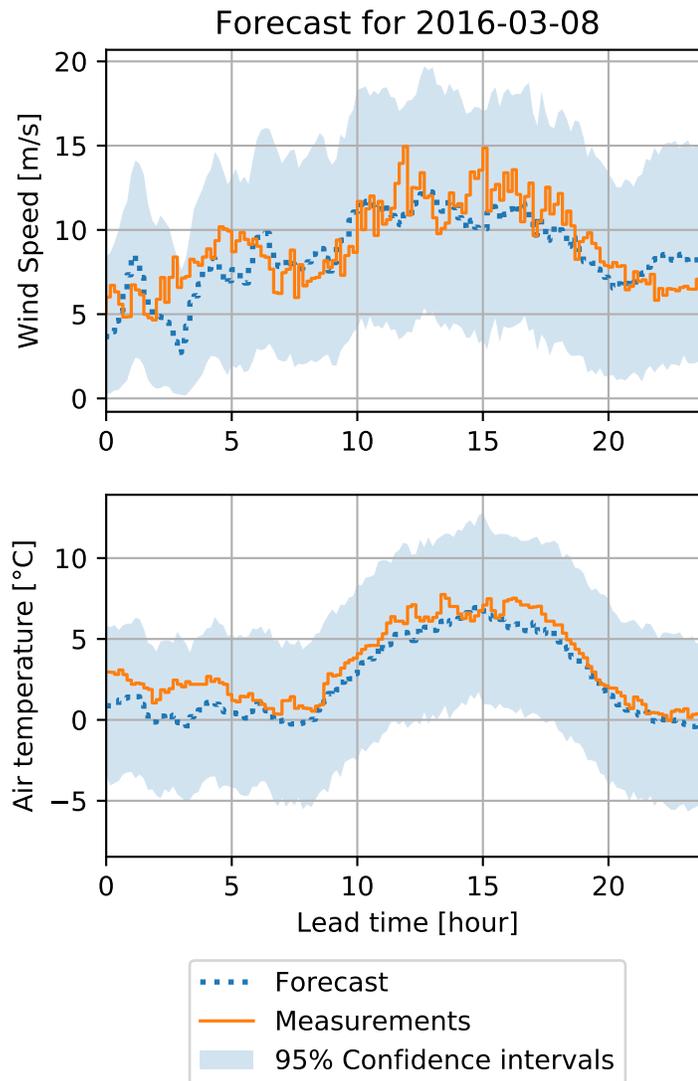


Figure 4.2: Example of a probabilistic forecasts for two different weather variables. The confidence levels were derived from the NWP forecast by the statistical model.

then generates the deterministic forecast for day D . Starting with the first time-step, the models are used to estimate the distribution of predicted variables from which M samples are drawn and then M samples of ampacity are calculated in a Monte Carlo fashion. The final ampacity estimate \hat{A}_t for time-step t with confidence q is taken as $(q \cdot 100)$ th percentile from the samples. After this calculation is repeated for each time-step, the minimal occurring ampacity during the day is found and returned as the daily ampacity.

The parameters of the probabilistic models are found by minimizing the negative log-likelihood expression of individual models on the training dataset consisting of N past forecasts. Limited-memory BFGS optimization algorithm was used to minimize the expression.

Figure 4.2 shows an example demonstrating the result of using the statistical model to convert a deterministic NWP to a probabilistic forecast. The statistical model was trained on forecasts of N previous days and then used to generate the confidence intervals for the forecast.

A separate statistical model is used for each weather variable. The models are adopted from [40], with the exception of wind speed, which is assumed to have truncated normal distribution. The statistical model is updated with each new forecast using N past forecasts and associated measurements. This allows for adaptation to changing weather patterns.

The *ambient temperature* T is modelled as being normally distributed, with the mean depending on the temperature \hat{T}_{NWP} predicted by NWP.

$$\begin{aligned} \hat{T}(t) &\sim N(\mu, \sigma) & (4.1) \\ \mu &= a_0 + a_1 \hat{T}_{NWP}(t) \\ \sigma &= b_0 + b_1 \hat{T}_{NWP}(t) + b_2 t + b_3 \sin \frac{15 + t}{12\pi} \end{aligned}$$

The variance σ is independent of temperature, but depends on lead time t . It

has been observed that the variance increases with lead time and has a periodic component of 12 hours, as modelled by the sinus function.

Wind speed v , as a non-negative variable, is modelled using a truncated normal distribution where the truncation point set at 0.

$$\hat{v}(t) \sim N_{trunc}(\mu, \sigma) \tag{4.2}$$

$$\mu = a_0 + a_1 \hat{v}_{NWP}$$

$$\sigma = b_0 + b_1 \hat{v}_{NWP} + b_2 t$$

In the case of wind speed, both the mean and variance depend on the wind speed \hat{v}_{NWP} predicted by the NWP model. The variance also depends on the lead time t .

For learning the model parameters, censoring is applied to the normal distribution on top of truncating because the data set contains numerous 0 m/s wind speed measurements caused by the non-zero start-up speed of cup anemometers. Censoring is crucial because a high concentration of values close to zero negatively impacts the maximum likelihood optimization. Therefore, all measured wind speeds less than 0.5 m/s are set to 0.5 m/s for learning the parameters, and a compensation term is added to the likelihood function of the model. A wind speed of 0.5 m/s was chosen because it is a standard value that can be assumed for low wind speeds in the thermal calculation [16]. During evaluation, the original values are used.

Wind direction is a circular variable and is modelled using both the von Mises distribution, a probability distribution defined on a circle, and a uniform distribution. The uniform distribution was specifically added to model the random errors caused by measuring wind direction at lower wind speeds

$$\hat{d}(t) \sim \alpha \cdot U(-\pi, \pi) + (1 - \alpha) \cdot \text{VM}(\mu, \kappa), \quad (4.3)$$

$$\mu = \hat{d}_{NWP},$$

$$\alpha = a_0 + a_1 \hat{v}_{NWP}(t),$$

$$\kappa = b_0 + b_1 \hat{v}_{NWP}(t),$$

where α is the mixing ratio and κ is the shape parameter of the von Mises distribution. Both parameters depend on wind speed because the prediction is more accurate as wind speed increases.

Solar irradiation is not modelled due to its relatively small influence on the final ampacity compared to other variables. Additionally, the NWP of solar irradiation alone has a high level of reliability. A probabilistic model was tested, but in the end did not yield a measurable effect on the results.

4.1.2 Naive prediction

A simple baseline method of DLR prediction is defined as the minimum value of daily ampacity over the past several days. More precisely, the ampacity prediction \hat{A}_{naive}^{D+1} for day $D + 1$, issued on day D , is the minimum of N previous days $\{A^{D-N}, \dots, A^D\}$

$$\hat{A}_{naive}^{D+1} = \min(A^{D-N}, \dots, A^D). \quad (4.4)$$

4.1.3 Perfect prediction

The perfect prediction represents the result that could be obtained from daily DLR if access to a perfect forecast was possible. This method is not operationally feasible because the perfect forecast is known only retrospectively.

$$\hat{A}_{perfect}^{D+1} = A^{D+1}. \quad (4.5)$$

Algorithm 2 Probabilistic DTLR prediction algorithm

```
1:  $H \leftarrow$  days  $(D - N) \dots (D - 1)$  from database
2:  $\Theta_v \leftarrow$  EM(wind speed model,  $H$ )
3:  $\Theta_d \leftarrow$  EM(wind direction model,  $H$ )
4:  $\Theta_T \leftarrow$  EM(temperature model,  $H$ )
5:  $W \leftarrow$  NWP prediction for day  $d$ 

6:  $t \leftarrow 1$ 
7: while  $t \leq 24$  do
8:    $samples \leftarrow$  empty list
9:   while  $\text{length}(samples) < M$  do
10:     $v \leftarrow$  sample(wind speed model( $W(t), \Theta_v$ ))
11:     $d \leftarrow$  sample(wind direction model( $W(t), \Theta_d$ ))
12:     $T \leftarrow$  sample(temperature model( $W(t), \Theta_T$ ))
13:     $s \leftarrow W(\text{solar})$ 
14:     $A \leftarrow$  IEEE738( $v, d, T, s$ )
15:    append  $A$  to  $samples$ 
16:   end while
17:    $samples \leftarrow$  sort( $samples$ )
18:    $\hat{A}_t \leftarrow samples[\lfloor (1 - q) \cdot M \rfloor]$ 
19:    $t \leftarrow t + 1$ 
20: end while

21:  $A \leftarrow \min(A_t)$ 
    return  $A$ 
```

This method confers value by defining maximal increase in ampacity possible using daily DLR.

4.1.4 NWP with derating

The deterministic line rating calculation with derating is an alternative method that uses NWP. In this method, the daily rating is calculated from the deterministic forecast, and the thermal rating is then lowered by a fraction. Derating is applied because NWP is not perfect. Specifically, the chance of exceeding the true ampacity is high if the prediction is taken without modification.

$$\hat{A}_{nwp} = \hat{A}^{D+1} - (\hat{A}^{D+1} - A_{static}) \cdot \alpha, \quad (4.6)$$

where α is the derating factor and A_{static} is the static rating. By applying the derating only to the portion of ampacity between the static rating and the forecasted rating, the effect is higher when the prediction is farther from the the static rating.

4.2 Results

In this section the testing data and methodology are first described, and then the results are presented.

Data

The algorithm was tested against real measurements taken at three weather stations installed along a transmission line located in southern Alberta. The weather stations were positioned with one at both ends and one in the middle. This was a relatively short line with a voltage of 138 *kV* and a length of 2 *km*. The dataset is from 2016 and contains measurements for wind speed, wind direction, air temperature, and solar irradiation recorded at 3 minute intervals.

The high resolution weather forecast was generated using the Weather Research and Forecasting Model (WRF), which is a mesoscale numerical weather prediction

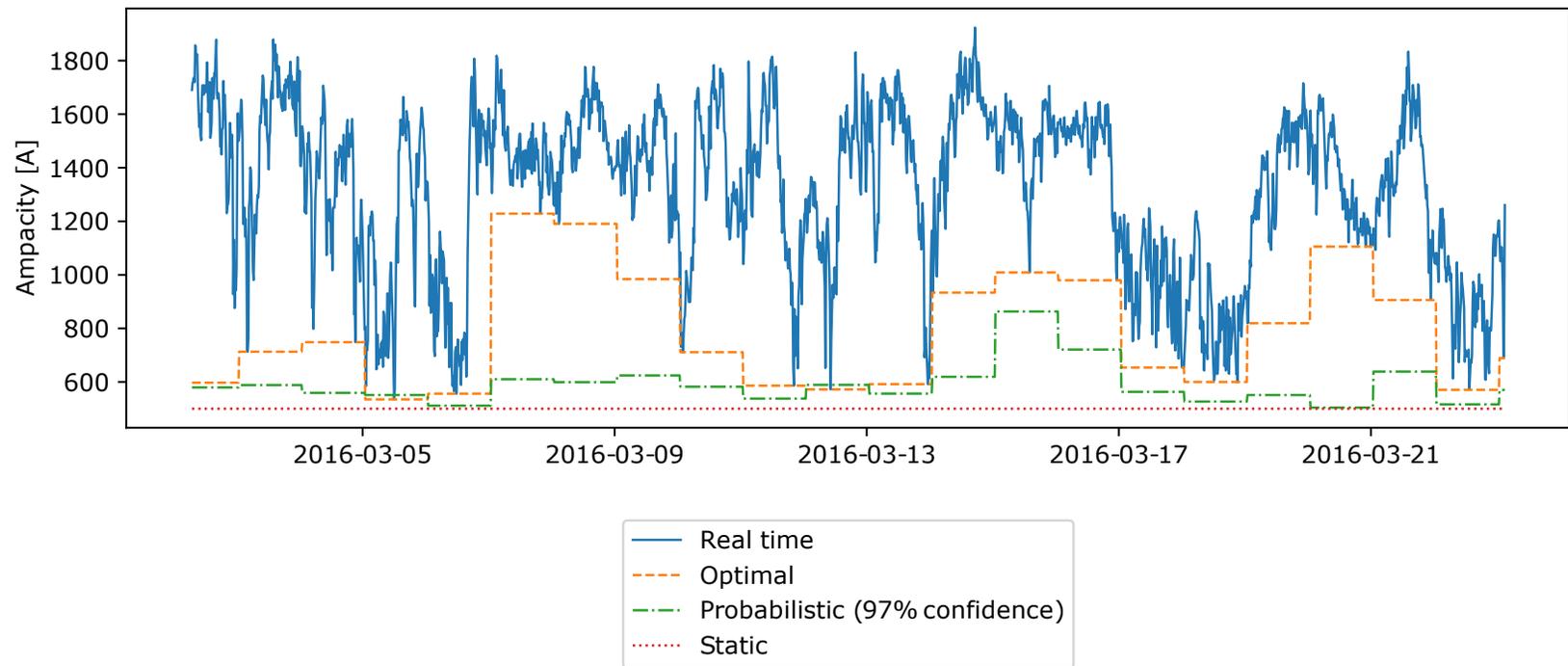


Figure 4.3: Example of the time series of ampacity. The figure shows the real-time ampacity together with the optimal daily ampacity, the static rating and the probabilistic prediction at 97% level of confidence.

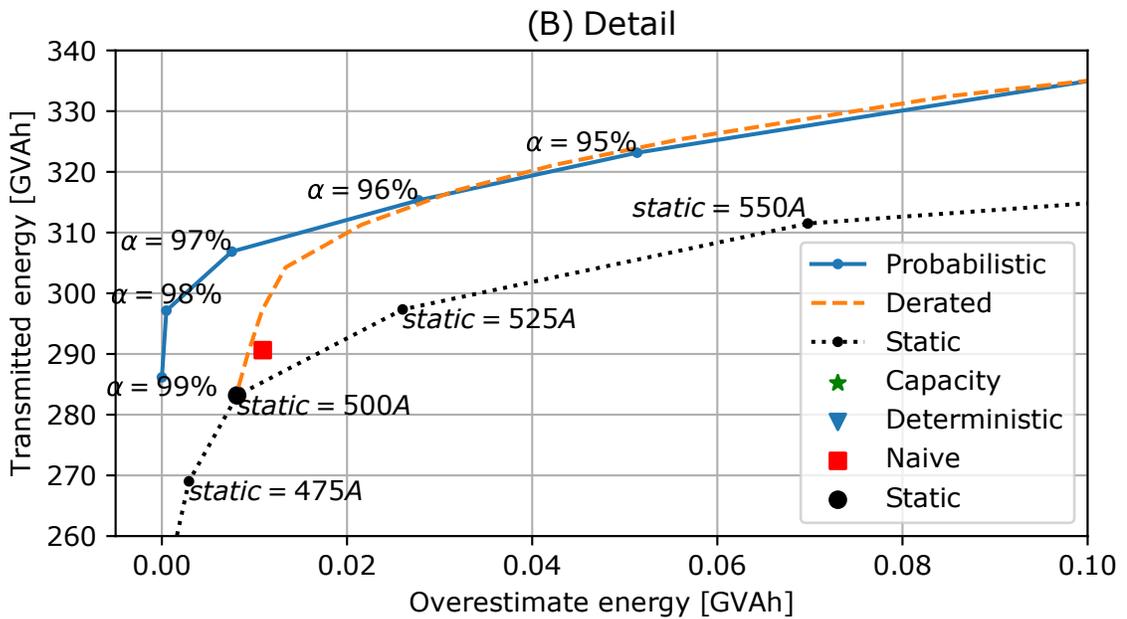
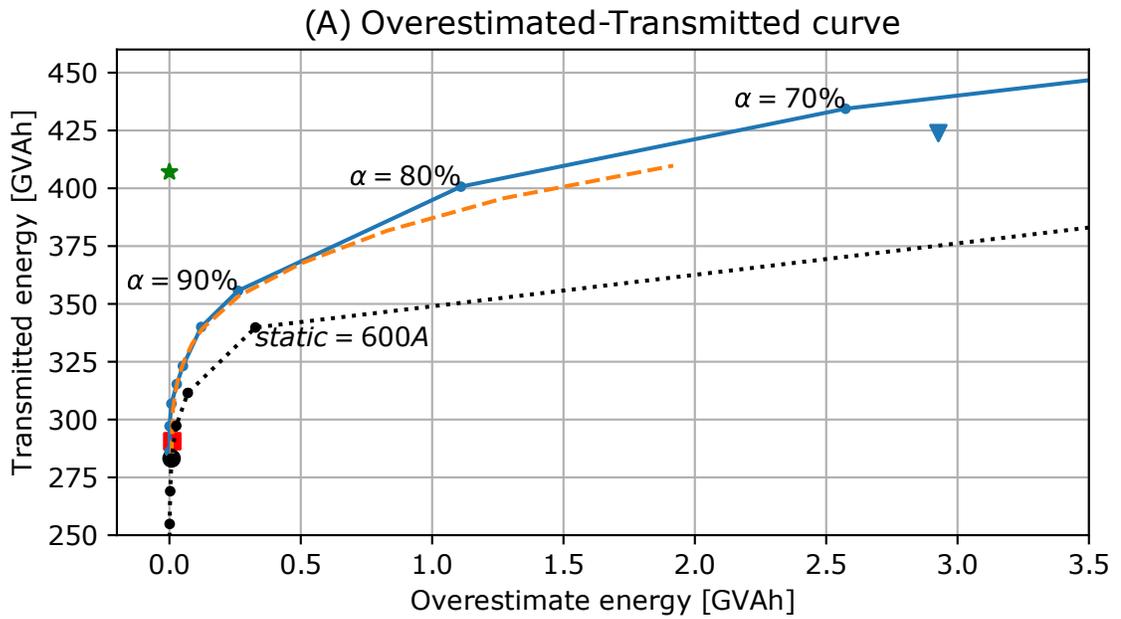


Figure 4.4: Relationship between the overestimated and transmitted energy for varying parameters of used algorithms.

system. The initial and lateral conditions for the model come from the North American Mesoscale Forecast System (NAM) datasets, which are released four times a day in six hours intervals. Each dataset contains data up for to 84 hours in the future with a spatial resolution of 12 km. WRF was used to further increase the time-step to 20 seconds and the spatial resolution to 1 km over southern Alberta.

For the purpose of evaluation, WRF was run in hindcast mode, meaning it was forecasting for past dates. Daily at 23:00 MST, a forecast was issued for the following day and the output was recorded in 5 minute intervals. Both the measurement and the forecast were resampled and aligned to 10 minute time-steps. Due to the time constant of the heating process, a time resolution of 10 minutes is sufficient for calculating the thermal rating [58]. While calculating the averages, special care was given to wind direction, which cannot be averaged as a scalar [59]. Wind speed and direction were converted to perpendicular components, averaged, and then converted back to angles.

Evaluation methodology

All of the methods previously defined are based on the principle of sliding window, where M previous days are used for learning the parameters. Therefore, the evaluation does not include the first 30 days of 2016 because there was no data from 2015 available for this training window.

The ampacity forecast and real-time ampacity were calculated for all three weather station locations. Then, the smallest value from the three locations was used for evaluation. The location with the lowest ampacity is called the critical segment. In DTLR, the ampacity of the line is equal to the ampacity of the critical segment, because it is the weakest part of the transmission line.

The probabilistic ampacity algorithm was evaluated for multiple levels of confidence ranging from 50% to 99% allowing the relationship between confidence and gain in ampacity to be visualized. For similar reasons, the derating algorithm was evaluated with the derating factor α between 0 to 1. Considering the edge values of α , $\alpha = 0$

is equivalent to using the static rating and $\alpha = 1$ is the same as directly using the deterministic value of the forecast.

To evaluate and compare the performance of each method, the following metrics were calculated. The first metric is the amount of energy E that would be transmitted over the testing period if the line were loaded at its rating.

$$E = V \cdot \sum_{n=1}^N \hat{A}(n) \cdot dt, \quad (4.7)$$

where V is the voltage on the line, $\hat{A}(n)$ is the ampacity estimate at interval n , N is the number of time intervals in the testing dataset, and dt is the length of the time interval in hours. The unit of E is $V Ah$.

The second metric is the overestimate E_o of the dynamic rating. In other words, it represents the amount of energy in $V Ah$ that would be transmitted while the ampacity estimate is higher than the true ampacity.

$$E_o = V \cdot \sum_{n=1}^N \max(\hat{A}(n) - A(n), 0) \cdot dt, \quad (4.8)$$

where \hat{A} is the ampacity estimate and A is the true ampacity. Ideally, this metric is $0 V Ah$, meaning that the estimate is always less than or equal to the actual real-time rating.

The third metric is the amount of time T_o that the estimate is higher than the real-time rating.

$$T_o = \sum_{n=1}^N (1 \text{ if } \hat{A} > A \text{ else } 0) \cdot dt \quad (4.9)$$

Discussion

Figure 4.3 shows an example time series of the probabilistic rating, the real-time rating calculated from the measurements, and the static rating. At the end of each

day, a weather prediction is made and a rating is calculated for the following day. This process is represented in Figure 4.3 as discrete steps that happen at midnight.

The overestimated-transmitted energy curves are plotted in Figure 4.4. The plot shows how the amounts of transmitted and overestimated energy change, while varying the parameter of the algorithm. The static rating of the transmission line under study is 500 A, but both transmitted and overestimated energy were calculated for a range of static ratings to provide a curve for comparison. Both panels show the same data, however, the panel on the right provides a more detailed look at the lower values.

Looking at the left panel of Figure 4.4, it can be observed that both the probabilistic and deterministic algorithm provide similar amounts of transmitted energy for higher overestimates. It is worth noting that both algorithms perform better than the static rating.

The right panel contains the curve for the lower values of overestimated energy. This figure shows how the probabilistic algorithm performs better than the deterministic calculation for any confidence higher than 96%. This figure also confirms the assumption that the derated deterministic calculation, with the derating factor of 0, is the same as the static limit.

It is important to note that the static rating of 500 A also overestimates the ampacity by 8 *MVAh*, which means that even with SLR, the thermal limit of the transmission line can potentially be violated. This opens the possibility of using the probabilistic prediction in two ways. Either the transmitted energy can be increased by keeping the overestimate at the same level, or the overestimate can be decreased to nearly zero while keeping the transmitted amount of energy the same.

These results are summarized in Table 4.1. This table also contains the results from the two other weather stations that were not included in Figure 4.4.

	E [GVAh]	$1 - E/E_{static}$ [%]	E_o [MVAh]	T_o [hour]
Real-time capacity	984.2	131.7	0.0	0.00
Daily Capacity, $\hat{A}_{perfect}$	591.5	39.3	0.0	0.00
Naive algorithm	425.9	0.3	10.7	3.75
Static 475A	403.5	-5.0	4.6	2.00
Static 500A	424.8	0.0	17.6	5.50
Static 525A	446.0	5.0	51.1	13.75
Derated $\alpha = 0.01$	426.8	0.5	17.8	5.50
Derated $\alpha = 0.05$	434.9	2.4	18.6	5.50
Derated $\alpha = 0.10$	445.1	4.8	19.8	6.00
Derated $\alpha = 0.15$	455.2	7.2	23.6	8.50
Derated $\alpha = 0.20$	465.4	9.6	33.5	11.25
Derated $\alpha = 0.24$	473.5	11.5	45.9	13.00
Probabilistic $q = 99\%$	425.0	0.1	0.3	0.75
Probabilistic $q = 98\%$	442.5	4.2	7.3	1.75
Probabilistic $q = 97\%$	457.2	7.6	17.6	3.50
Probabilistic $q = 96\%$	469.9	10.6	37.7	7.75
Probabilistic $q = 95\%$	481.3	13.3	68.1	12.75
Deterministic NWP	627.8	47.8	4249.8	248.25

Table 4.1: Simulation results of different transmission line rating techniques during the year 2016.

4.3 Conclusion

DLR is a promising technique that can unlock the additional capacity of power transmission lines. However, its dynamic nature brings additional complexities and requires explicit support for real-time line ratings in the EMS of the grid operator. This paper evaluated a method of DLR where changes are applied daily, as opposed to continuously, to allow DLR to be used even if real-time updates are not supported.

Several methods for calculating daily ratings were compared: a naive method that predicts the minimum ampacity of ratings from several past days, a deterministic method with derating based on NWP, and a probabilistic method based on NWP. These methods were evaluated using real measurements taken from a transmission line in southern Alberta in 2016.

Based on the calculations performed, the static rating was potentially exceeded for 5.5 hours in 2016. This shows that even though the static rating was selected as a conservative estimate, absolute safety is not guaranteed. With this in mind, the performance of probabilistic DLR can be evaluated from two different standpoints. The capacity of the line could be increased by 7.6%, while exceeding the real-time rating for 3.5 hours by using the rating given by the 97% confidence level of the probabilistic ampacity. Alternatively, the reliability could be increased while keeping the same ampacity. When the 99% confidence level was used, the ampacity exceeded the real-time rating only by 0.75 hours, and the transmitted energy remained the same as with SLR.

In conclusion, this experiment showed that daily DLR provide an increase in the capacity of a transmission line regardless of whether implementation of real-time DLR is possible. Unsurprisingly, the impact of daily DLR would never be as substantial as that of real-time DLR – it was calculated, that by having an ideal forecast the daily DLR would increase the capacity by 39.3% over SLR. In comparison, a true real-time DLR achieved a 131.7% increase in capacity on the studied transmission line.

Future work will focus on improving the forecast model so that the increase in capacity can be brought closer to the 39.3% maximum of daily DLR. The probabilistic model employed in this work used only a limited amount of information available from the NWP model. For example, each variable was modelled separately, and the model did not take advantage of temporal correlations. Therefore, a more involved model will be tested. Ensemble NWP or alternative methods, such as pseudo-ensemble [60] and time-lagged ensemble [61], will also be examined. Another area that could be explored is testing different time intervals beside the presented daily DLR. It is possible that a slightly larger or shorter interval may provide better results, while still remaining feasible for implementation without real-time DLR support.

Chapter 5

The Effect of Temporal Discretization on Dynamic Thermal Line Rating

A well functioning power transmission network is transparent to the user. However, when transmission limits are reached, various problems arise as there is insufficient capacity to transport energy to the consumer. These problems include increases in local electricity pricing and curtailment of renewable power resources. While an upgrade to transmission capacity would avoid this, it comes with a high cost in materials, labour, and outages.

Rating (also known as ampacity) is defined as the maximal limit of electrical current that can be carried by any given transmission line. In shorter lines, the determining factor is the maximal allowed temperature of the conductor. Current passing through the line generates heat, thereby increasing the temperature of the conductor. If the temperature surpasses a specific limit the conductor will begin to sag, which may break ground clearance, accelerate conductor ageing, and cause annealing. To prevent these changes the operator sets the thermal rating to guarantee the temperatures are always within a safe range [16].

The simplest method to determine the rating of a line is to find and utilize the highest rating that is safe under all possible conditions. This method is termed Static Line Rating (SLR) and is commonly used in the operation of transmission lines. It

works by identifying the worst case conditions that a line will be subject to and then calculating the maximum electrical current that will reach the temperature limit of the line. The conditions involved in the calculation include the electrical current, ambient temperature, wind speed and direction, and solar irradiation. The advantage of this method is its simplicity, however, it lacks efficiency due to its inherent conservative assumption that only the worst case conditions occur [16].

To improve efficiency and postpone the need for costly upgrades, Dynamic Thermal Line Rating (DTLR) was developed to harness the total potential capacity of the line and is calculated based on actual conditions [34]. Theoretically, it can enable the operator to utilize the full capacity of the transmission line if the parameters of the transmission line and its surroundings are monitored closely and then used to calculate rating in realtime. This ideal case is called Real-Time DTLR (RTLRL) [3].

Currently, a major part of the electrical system relies on preemptive scheduling and planning, such as in the case of the day-ahead electricity market. The rating of the transmission network must be known during the flow calculation, as it affects the pricing in case of network congestion [37]. Therefore, despite the ability of the RTLRL method to provide optimal performance, it would be impractical to implement in the operation of transmission lines due to its realtime nature and reliance on ideal measurements that are not known before hand.

Practically, DTLR is applied as an operator set limit that attempts to mimic RTLRL without exceeding it [3]. In contrast to instantaneous and continuous RTLRL, DTLR is typically applied in advance and in discrete periods where a) DTLR cannot be higher than RTLRL during any given period, and b) DTLR is based on the forecasts of future RTLRL. These implications adversely affect the benefit of DTLR, as in the case of a) DTLR will be lower than RTLRL by definition unless RTLRL is a flat curve, and b) any RTLRL forecast comes with a certain amount of uncertainty that has to be accounted for by decreasing the DTLR by a safety margin [35].

This paper focuses on the effects of *temporal discretization*, which is the process of

converting the continuous RTLR into a discrete series of timesteps by taking the minimum rating during the period. The objective is to describe the relationship between the length of the discretization period and the benefit of DTLR. It is hypothesized that shorter periods of discretization will provide benefit near that of RTLR, while the benefit of longer periods will converge to the value of SLR. Moreover, the function will be expected to decrease monotonically because taking the minimum rating of progressively longer periods will only decrease the benefit. The secondary objective is to identify the longest discretization period where wind speed remains relevant. The expectation is that due to the high temporal variability of wind speed, its effect will drop significantly even with short discretization periods. Ambient-Adjusted Line Rating (AAR) is a simpler rating method that considers only the ambient temperature. Therefore, finding the time scale where wind speed becomes irrelevant allows for the determination of when DTLR transitions to AAR [35, 37].

Wind speed is generally considered to be the major factor affecting DTLR [16] and thus, the temporal discretization of wind speed is analyzed first with other variables held constant. Subsequently, all remaining variables are then discretized and DTLR benefit is observed.

5.1 Temporal Discretization

Definition 1 *Temporal discretization.* Given a time series $\{x_1, \dots, x_n\}$ with equally spaced measurements with a sample period T , a temporally discretized time series $\{\check{x}_1^{D=d}, \dots, \check{x}_n^{D=d}\}$ with a discretization period d is then defined as follows.

$$\check{x}_i^{D=d} \equiv \min(\{x_j; j \in Z, j \div d = i \div d\}), \quad (5.1)$$

where \div is integer division. Without the loss of generalization, the discretization period D is assumed to be in the units of T . Described in words, the time series is divided into segments of length D , and then each value is replaced by the minimum

of the segment to which it belongs.

In Chapter 6 the definition is expanded to:

$$\check{x}_i^{D=d} \equiv f(\{x_j; j \in Z, j \div d = i \div d\}), \quad (5.2)$$

f is a function. In the previous definition, f was assumed to be the minimum function, however, it could be an arbitrary function that summarizes the discretization block. In brief, this process replaces the values of the time series by the summary value of each block. It does not change the number of samples nor the sampling rate. For example, replacing each time-step of a wind speed time series with its daily minimum would be temporal discretization with a length of 24 hours using the minimum function.

Temporal variability affects a discretized time series such that higher variability results in a lower mean value. This is due to the nonlinearity of the minimum function used to calculate the discretized series. This principle is illustrated in the following example. Assume 30 minute measurements of DTLR: 500, 500, 1000, 1000 A. When this rating is discretized to 1 hour periods the result will be 500 A in the first hour and 1000 A in the second hour. Now consider a different series of 30 minute measurements: 500, 1000, 500, 1000 A. This series is transformed to 1 hour periods as 500 A for both hours. Although both series can transfer the same amount of energy on the 30 minute time scale, the second time series transports only two thirds the energy of the first series on the hourly scale.

5.2 Wind speed model

Wind speed is often modelled using Weibull distribution [62], which is a nonnegative distribution that can be specified using two parameters α and β , that represent shape and scale respectively. A useful property of Weibull distribution is that it is reproductive through formation of the minimum [63]. In other words, the minimum of

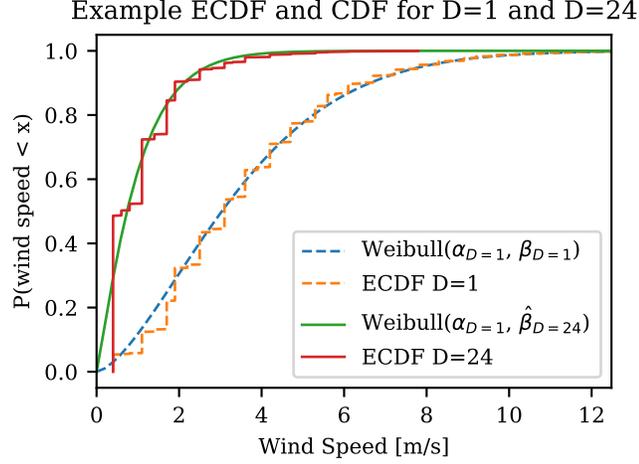


Figure 5.1: Dashed curve of Weibull distribution CDF fits the Empirical CDF (ECDF) of hourly wind speed data well. The solid line shows the CDF of daily minimal wind speed. The parameters for the daily CDF were calculated from the hourly measurements using Eq. 5.3.

several independent random Weibull variables is also distributed according to Weibull distribution. The following equation shows the parameter values for the minimum of n Weibull random variables.

$$\begin{aligned}
 Y &= \min\{X_1, \dots, X_n\}, \quad X_i \sim Weibull(\alpha, \beta_i) \\
 Y &\sim Weibull\left(\alpha, \left[\sum_{i=1}^n \beta_i^{-\alpha}\right]^{-1/\alpha}\right)
 \end{aligned} \tag{5.3}$$

As n increases, the scale parameter β decreases, and the entire distribution shrinks towards 0 while keeping the same origin and shape α . This property of Weibull distribution is used to model the distribution of wind speed after discretization. If $\forall i : \beta_i = \beta$, then Eq. 5.3 can be further simplified to $\beta^{(n)} = \beta / \sqrt[n]{n}$.

Eq. 5.3 assumes that samples $\{X_1, \dots, X_n\}$ are independent, which does not hold in the time series of wind speed data due to temporal correlations. In order to apply this equation, the effective sample size n_{eff} will be used instead of n . Assuming that wind speed is generated from an $AR(1)$ process, n_{eff} can be estimated using the following equation [64].

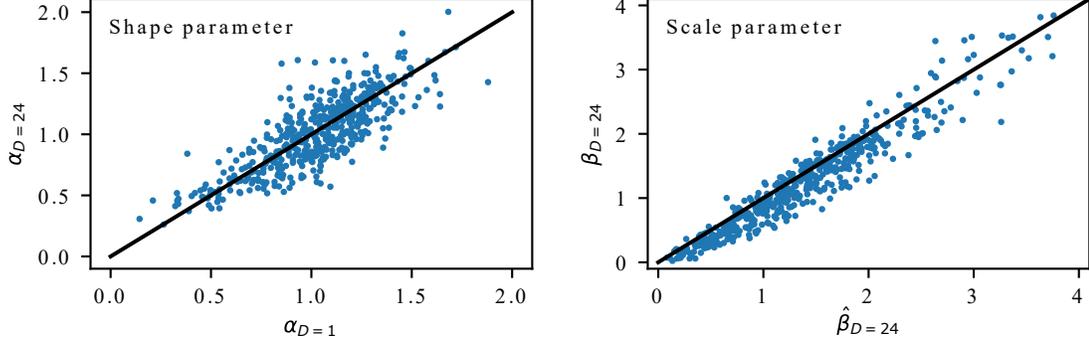


Figure 5.2: Shape and scale parameters of the Weibull distribution of wind speed estimated from hourly ($D = 1$) and daily discretized ($D = 24$) time series. $\hat{\beta}_{D=24}$ is the estimate of $\beta_{D=24}$ calculated from Eq. 5.3. The calculated values are on the horizontal axis and the values estimated from measurements are on the vertical axis.

$$n_{eff} = n \frac{1 - \rho_1}{1 + \rho_1} = n \frac{1}{\tau^2}, \quad (5.4)$$

where ρ_1 is the wind speed time series autocorrelation for the lag of 1, and τ is the autocorrelation time (the number of time steps it takes for the samples to become uncorrelated).

The parameters of the Weibull distribution are estimated through Markov Chain Monte Carlo (MCMC) with positive *HalfCauchy*(1) priors for both α and β . Special care must be given to the left side of the wind speed distribution because most of the used data is collected from cup anemometers, which have high error for low wind speeds due to their non zero start-up speeds. To overcome this problem and improve parameter estimation, all wind speed values less than 1 *m/s* are censored and integrated out of the model [65]. The model is implemented in the software package PyMC3 version 3.8.

5.2.1 Static Line Rating

In this study, SLR is calculated from historical measurements as to satisfy the criteria mentioned in Chapter 2. This selected approach provides the highest estimate of SLR, which in turn produces a conservative estimate of DTLR benefits. Moreover, SLR

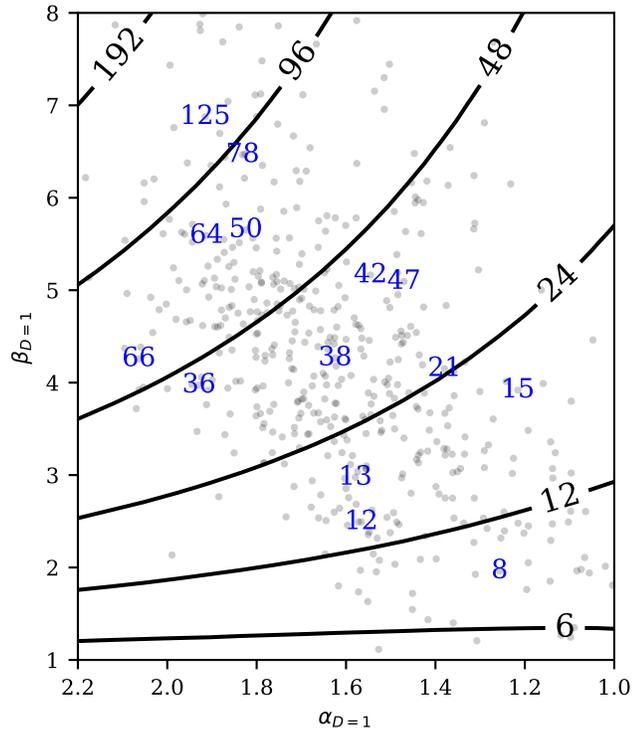


Figure 5.3: The contours represent the length of the discretization periods that results in the rating being within 5% of SLR calculated from Eq. 5.7. The gray points represent the weather stations plotted on their (α, β) coordinates. 15 stations are randomly selected and their discretization lengths are shown in blue numbers.

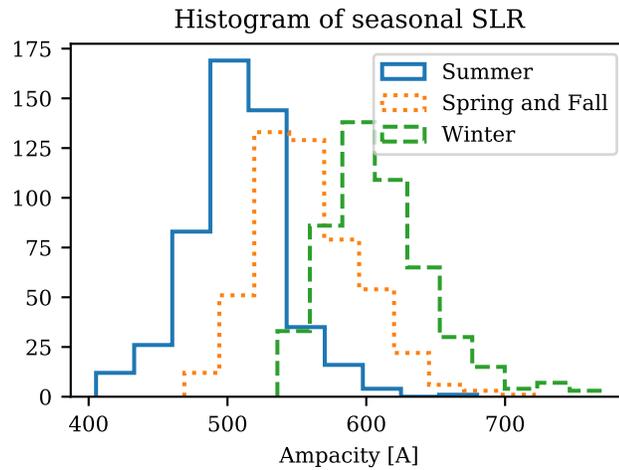


Figure 5.4: Histogram of seasonal SLR for three different seasons. SLR is defined as bottom 1% ampacity during the season across the entire dataset.

used in this study is calculated seasonally, meaning there are three separate limits for winter (Dec-Feb), summer (Jun-Aug), and spring (Mar-May) together with fall (Sep-Nov). Seasonal SLR is used as it is commonly seen in operation and provides a fair baseline for comparison.

SLR is calculated by the following procedure. First, the rating is calculated for the entire time series of hourly weather measurements using the IEE-738 steady state model. The ampacity is then grouped into the three seasonal categories as defined above. Finally, a seasonal ampacity that satisfies the CIGRE-299 requirements is selected.

5.2.2 Dynamic Thermal Line Rating

Dynamic thermal line rating is calculated from hourly weather measurements using Eq. 2.2. The temporal discretizations are then calculated by Eq. 5.1 with the hourly data considered as discretization $D = 1$.

In order to compare different locations, *the benefit of DTLR* is defined as the average of DTLR normalized by the seasonal SLR. For instance, a DTLR benefit of 1.2 means a 20% increase over SLR.

5.2.3 Dataset

This study uses Canadian Weather Energy and Engineering Datasets (CWEEDS) to simulate rating calculations across Canada [66]. This dataset contains historical weather information from 492 weather stations between the period of 1998 and 2014. The data provided in this source include hourly measurements of temperature, wind speed, wind direction, and solar irradiation, which permits the computation of hourly rating for a hypothetical power transmission line segment placed at the location of the weather station.

The dataset contains not only measurements, but also outputs from a numerical weather model that are used to fill in gaps of missing data that were longer than

3 hours. For the purpose of this study, these synthetic values of wind speed and temperature were removed and only time periods of continuous real measurements are kept.

5.3 Discussion

The following will discuss how Weibull distribution is used to estimate the effect of discretization on wind speed and rating when other environmental variables are held constant, and then how a similar analysis is conducted on the rating when all environmental factors are included.

5.3.1 Temporal Discretization of Wind Speed

According to Eq. 5.3, the shape of the wind speed distribution will remain constant after discretization, but the scale will decrease. This is tested on the dataset by a 3-step simulation. First, the parameters $\alpha_{D=1}$, $\beta_{D=1}$, and τ are estimated for hourly wind speed. Then, these parameters are used to calculate $\hat{\beta}_{D=24}$ according to Eq. 5.3. Finally, the true parameters $\alpha_{D=24}$ and $\beta_{D=24}$ are determined from the discretized wind speed.

In theory, $\alpha_{D=1} = \alpha_{D=24}$ and $\hat{\beta}_{D=24} = \beta_{D=24}$. Figure 5.2 shows a scatter diagram of the calculated and measured parameters. The top panel portrays the shape parameter α . While ideally these values should be equal, the plot shows a significant amount of noise. The authors hypothesized two main causes of this problem: (1) estimation of the parameter $\alpha_{D=24}$ is inaccurate because most of the measurements are located close to 0 where the value is dominated by the sensor noise; (2) the assumption of independence is invalidated and there is no correction mechanism like there is for the parameter β . Despite this noise, the correlation is clearly visible with $r^2 = 0.65$ in ordinary least squares (OLS) regression.

The bottom panel of Figure 5.2 shows the relationship between the scale parameter $\hat{\beta}_{D=24}$ that has been calculated from the hourly measurements, and $\beta_{D=24}$ that was

directly estimated from daily discretization. In this case, both Eq. 5.3 and Eq. 5.4 provide an accurate estimate of the parameter with $r^2 = 0.93$ in OLS regression.

The expected wind speed $W_s \sim Weibull(\alpha, \beta)$ is given by

$$\mathbb{E}(W_s) = \beta \Gamma(1 + \alpha^{-1}), \quad (5.5)$$

where Γ is the gamma function. Eq. 5.3 shows that α does not change as D increases and therefore Eq. 5.5 can be combined with Eq. 5.3 to get the expected value of wind speed after discretization $D = d$

$$\mathbb{E}(W_s^{D=d}) = \frac{\beta}{\sqrt[d]{d/\tau}} \Gamma(1 + \alpha^{-1}) = c \cdot d^{-1/\alpha} \quad (5.6)$$

where c is a constant independent of d . As this equation describes the power law, it can be stated that the average wind speed after discretization follows the power law.

To address the secondary objective regarding when the effect of wind speed on DTLR becomes negligible, the expected rating is compared with SLR. SLR is calculated as the steady state rating with an ambient temperature of 35 °C, perpendicular 0.6 m/s wind speed, and 1000 W/m² solar irradiation. DTLR is calculated by the following equation.

$$\begin{aligned} \mathbb{E}(I^{D=d}; \alpha, \beta) &= F(\bar{w}_s; \alpha, \beta^{(d)}) \cdot \bar{I} \\ &+ \int_{\bar{w}_s}^{\infty} f(w_s; \alpha, \beta^{(d)}) I(w_s) dw_s, \end{aligned}$$

where f and F are the probability distribution function (PDF) and cumulative distribution function (CDF) of Weibull distribution, and $\beta^{(d)}$ is calculated according to Eq. 5.3. The following optimization was used to find d that results in $\mathbb{E}(I^{D=d})$ to be within 5% of \bar{I} .

$$\arg \min_d [\mathbb{E}(I^{D=d}; \alpha, \beta) - 1.05\bar{I}]^2 \quad (5.7)$$

This equation is numerically evaluated for a typical range of α and β and the result is shown in Figure 5.3. The contours represent values of D in hours for which the dynamic rating is within 5% of SLR. It can be seen that the benefit of wind profiles with higher β tend to decrease slower, which is in agreement with the observation that higher wind speeds are more stable. To evaluate if Eq. 5.7 is in agreement with the measured data, α , β and D are empirically estimated for the wind speed time series from the weather stations. Fifteen randomly selected results are shown in Figure 5.3. Overall, the experimental values of D agree with the calculated contours.

5.3.2 Temporal Discretization of Rating

The distribution of rating does not lead to an analytical result for the distribution of discretized rating as in the case of wind speed. However, the Fisher-Tippett-Gnedenko theorem can be used to approximate the distribution for a sufficiently large d . The theorem states that the distribution of the maximum of independent identically distributed (iid) random variables converges to either the Gumbel distribution (Type I), the Fréchet distribution (Type II), or the Weibull distribution (Type III).

The selection of distribution type depends on the tail of interest of the modelled distribution. Exponential distributions form Type I, distributions with fat tails (eg Cauchy) form Type II, and limited distributions form Type III extremes. Rating is created as a function of temperature, wind speed, and solar irradiation. Previously, the Weibull distribution was used to model wind speed, therefore it may be tempting to propose that the rating is Type III as Weibull is limited from the bottom. However, it was demonstrated that as d increases the effect of wind speed diminishes and the effect of temperature becomes more prominent. Thus, the distribution of discretized rating will likely converge to the Type I distribution given d is large.

This behaviour is observed in the dataset. Figure 5.5 shows the distribution of hourly ($D = 1$), daily ($D = 24$) and weekly ($D = 168$) rating. The Type III extreme value distribution can be applied to the daily values, but for the weekly values it

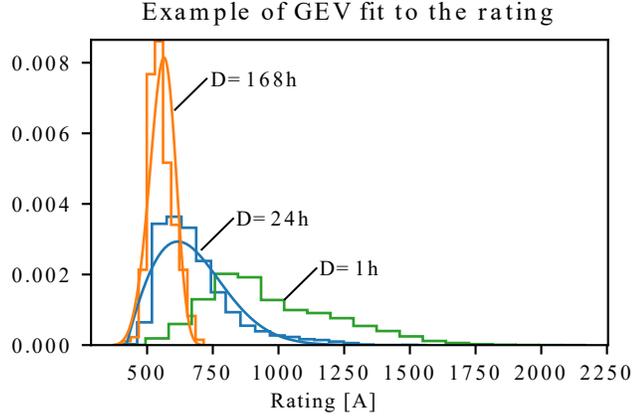


Figure 5.5: Extreme value distribution fitted to the discretized rating. Daily rating $D = 24$ can be fitted with Type III, whereas weekly rating $D = 168$ is fitted with Type I distribution.

clearly does not yield a proper fit. This phenomenon is observed in the majority of datasets where D is sufficiently large.

To evaluate the overall benefit of DTLR on the dataset, the hourly rating is calculated for all stations. The rating is then discretized for $D = 1..168$ and the mean is normalized by SLR to calculate the benefit. The summarized result is plotted in Figure 5.6. Based on the figure, the expectation that rating will decrease as D increases and that the rating tends towards the value of SLR is confirmed. The value of RTLR is not visualized because the hourly rate of the dataset is unable to capture the fastest changes that are occurring in RTLR on approximately a 10 minute scale. It can also be seen that the shape of the curve follows the power law function that was derived in Eq. 5.6.

Table 5.1 quantifies the data from Figure 5.6 for commonly used discretization periods. It can be seen that 90% of locations have an hourly DTLR benefit of 47-103%. This decreases to 11-43% for 24 hour discretization, which means that almost 2/3 of the benefit has been lost going from 1 hour to 1 day.

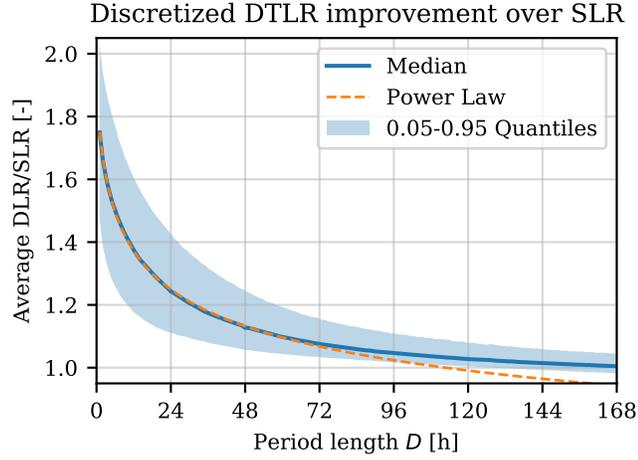


Figure 5.6: DTLR benefit for different temporal discretization periods. The power law approximates the shape of the curve until $D \approx 72$. The divergence for longer periods is expected and is caused by the censoring of wind speed data.

Table 5.1: DTLR benefit for different temporal discretization periods. Values from Figure 5.6 in a tabular form.

Period length [hours]	0.05 Quantile	Median	0.95 Quantile
1	1.47	1.75	2.03
3	1.34	1.61	1.89
6	1.26	1.50	1.77
12	1.18	1.38	1.60
24	1.11	1.24	1.43
48	1.06	1.13	1.25

5.3.3 Limitations

This study has some limitations, the majority of which are a result of the available dataset. Firstly, the dataset was developed for environmental research and thus, the locations of sensors are not representative of typical locations of power transmission lines. Secondly, the resolution of the data is relatively low, as the data is provided in 1 hour timesteps and the required resolution needed to capture RTL_R is around 10 minutes. Therefore, RTL_R cannot be calculated and the 1 hour DTL_R is not representative of it. Lastly, the rating is calculated only for one specific type of conductor. Regardless of these limitations, the authors believe the overall conclusions proposed are valid when acknowledging that the exact results cannot be taken at face value.

5.4 Conclusion

RTL_R is a technique that permits improved utilization of the installed capacity of overhead transmission lines when compared to SL_R. However, the realtime aspect of RTL_R makes it impractical to implement. Alternatively, DTL_R can be applied in discrete periods (eg. daily), which is simpler to incorporate into current systems. This paper examines the effects of temporal discretization to the overall benefit of DTL_R.

The first part focuses on wind speed, as it is the major factor affecting the rating. The commonly used Weibull distribution is fitted to the wind speed. One of the properties of this distribution is that the minimum of a set of iid Weibull random variables also follows Weibull distribution. This property is used to model the discretized wind speed, where the discretized distribution can be derived analytically from the hourly statistics. The validity is illustrated on the experimental data, where the calculated parameters of the discretized distribution are in agreement with the parameters estimated from the measurements.

Effective sample size is used in the calculation of the discretized wind speed distribution parameters to correct the effect of autocorrelated samples. This technique proved successful and the r^2 of a linear regression between the calculated and estimated β parameter was 0.93. The regression of the α parameter yields $r^2 = 63$. Theoretically, α should be constant. However, for reasons not currently understood, it changes. This is a topic for further research.

It was derived that the mean of wind speed behaves according to the power law, as the discretization period D increases (see Eq. 5.6). Although this was derived for the wind speed, it was shown to also have an effect on the overall ampacity, as the power law is visible in the summary in Figure 5.6 of DLR benefit calculated over the entire dataset.

It is shown that for wind speed discretizations longer than 4 days the majority of ampacity time series fall within 5% of the SLR – see Figure 5.3. This agrees with the overall rating in Figure 5.6, where at 4 days 95% of stations report a rating within 5% of SLR.

Overall, the benefit of applying hourly DTLR over SLR is between 47 and 103% and the benefit of daily DTLR (24 hour period) is between 11 and 43% for 90% of analysed locations. However, the exact numbers are only representative of this specific study because the calculation of ampacity is done only for one type of conductor. In an attempt to correct this problem, the authors normalized all results to SLR calculated for the same type of conductor and believe that the results are representative of the general trends. Nevertheless, with different types of conductors the results will inevitably change.

In conclusion, the authors found that in most cases, 96 hours is the maximal time period where wind speed is relevant in RTLR calculations, assuming a *perfect prediction* of wind speed is known in advance. In reality, this is an unrealistic expectation and the uncertainty in wind speed prediction will push the threshold to much lower values. This finding may also prove useful in DTLR prediction system design, because

for longer discretization periods, resources would not need to be spent on wind speed prediction as its effect would be minimal.

In answering the secondary objective, it was identified that discretized wind speed has the same distribution as hourly wind speed with the exception of the scale parameter that changes. It was shown that the scale parameter could be calculated with a high degree of accuracy using the measure of effective sample size and its reciprocal value (autocorrelation time τ). It was also found that the average wind speed as a function of the discretization period follows the power law, and that the distribution of rating tends to the extreme value distribution of Type III for low discretizations and Type I for high discretizations. This information can be used in Bayesian ampacity forecasting models, where it can provide an informative prior for discretized wind speed.

Chapter 6

Probabilistic Forecasting of Dynamic Line Rating with Temporal Correlations

The power transmission network is an essential part of the electrical system. Its purpose is to deliver electricity from producers to consumers. In ideal situation the network is fully transparent – the users of the system need not concern themselves with its existence. However, when the utilization of the network increases and the amount of transmitted power nears the line rating, the power must be rerouted through sub optimal paths increasing electricity costs. If the rerouting of power flow becomes unsolvable the undesirable result is load shedding [7]. Therefore, to avoid this it is more forgiving to have sufficient transmission capacity available in the electrical system at any given time.

While the easiest solution may appear to be upgrading the transmission lines, it is exceedingly expensive when you factor in the cost of labour, materials, and the opportunity lost during outages. Therefore, a more promising approach is to better utilize preexisting transmission equipment. Historically, the thermal rating of transmission lines was calculated for the worst case environmental conditions [17]. This technique is known as Static Line Rating (SLR) [67] and albeit very simple, wastes a lot of potential capacity as the conditions are almost always more favorable.

A more efficient technique is called Dynamic Thermal Line Rating (DTLR) and

sets line rating based on the current environmental conditions [67]. In an ideal case, the factors affecting thermal capacity are measured in real time, and then the rating is calculated and applied on the transmission line in the operator's management system [68]. Theoretically, this should be a continuous process, where the rating is constantly updated. However, as it has been shown that the time constant of the conductor's thermal behaviour is approximately 10 minutes, discrete updates in times steps shorter than this can be assumed to be equivalent to a continuous system [58]. This is known as Real-Time Dynamic Thermal Line Rating (RT-DTLR) and is the process that sets the line rating to the current that brings the temperature of the conductor to its design temperature [3].

RT-DTLR is therefore by definition, the optimal strategy that maximizes the transmission capacity. A current any higher than RT-DTLR would cause the line to overheat and age prematurely [9]. There are however, practical problems with RT-DTLR. Firstly, its realtime nature limits the usefulness, as the majority of planning and scheduling in the electrical system is done in advance [7]. The demand and production is agreed upon in an ahead-of-time market and then the schedules for power plants are created. Larger power plants need time to modify their output and the capacity of the network needs to be known in advance. Secondly, these electrical systems operate in finite granularity, meaning knowledge of continuous rating is not necessary and may even be detrimental. For example, retrofitting some current management systems with RT-DTLR would be impossible. Current systems often expect the rating as a number that changes only a few times a year and at best could accommodate daily changes, but not real time updates.

The objective of the research presented in this paper is to propose a method for deriving a single safe but economically desirable rating for a future period of time. Generally, the time period would be 24 hours, meaning the rating for the next day would be generated and remains static over this period. This proposed method encompasses three elements: weather forecast, probabilistic rating prediction, and rating summa-

rization (ie. condensing the probabilistic rating over time into a single number). An indirect calculation approach is used to estimate the rating, where rating is derived from the weather parameters that affect the thermal behaviour of the conductor [69]. Ambient temperature, wind speed and direction, and solar irradiation are forecast by a numerical weather prediction (NWP) system and then used in the IEEE-738 thermal model of an overhead transmission line to calculate the transmission line rating [16].

The first component of the method is the NWP model that generates the forecast in relatively high resolution in both temporal and spatial domains [70]. The second component of the method is the probabilistic ampacity prediction, which takes the weather prediction and calculates the probability distribution of the rating at any time step. The importance of having a probabilistic prediction, as opposed to a deterministic one, is that the degree of uncertainty is provided in the forecast [35]. For example, assume there are two predictions both with high accuracy, but one is precise and the other is imprecise. Even though both forecasts provide the same deterministic output, the forecast with low precision has a much higher degree of uncertainty that has to be accounted for by increasing the safety margin. Having a probabilistic output allows for objective determination of the safety margin and minimizes the effect of uncertainty on DTLR benefit. The final component of the method, the rating prediction, is to collapse the time series of probabilistic ampacity predictions into one number that is used as a safe rating for the entire time period. Simply put, the expected minimum of the rating over the time period is found in this step. The main challenge is caused by the temporal correlations between close time steps because they affect the estimated extreme value. Calculating the distribution of minima and maxima for multiple dependant random variables is a difficult problem. As such, an analytical solution is known only for a few well behaved distributions, for instance two normal distributions [71].

Similarly, in DTLR when the correlation between samples of the rating is high the

expected minimum of the rating will be higher. Therefore, it is important to take temporal correlation into account when calculating the expected minimum for the time period. The proposed solution to this problem works in a Monte Carlo fashion, where a number of random time series that are compatible with both probabilistic prediction and typical temporal correlations are generated. The distribution of the minimum can then be determined by examining the samples to select the rating through a quantile based method. Alternatively, the samples can be directly used with an objective-based optimization method to select the rating by more complex criteria.

Unlike most other research, the study presented in this paper focuses specifically on temporally discretized rating, i.e. rating that changes at the edges of short periods (hourly, daily) but otherwise stays constant. In the literature, this kind of DTLR is sometimes called Ambient Adjusted DTLR (AA-DTLR) [35], although some authors define AA-DTLR as dependant only on the ambient temperature. The detrimental effects of temporal correlation on rating estimation is described and a sampling method to mitigate this problem is presented. The conductor temperature is used during optimization, as in the study by Dupin [3]. A novel temperature estimation method, however, is developed and provides an alternative approach to the problem of conductor temperature estimation.

6.1 Methods and Procedures

The objective of the rating system is to provide a rating that maximizes the utility of the transmission line all while keeping the temperature related risks at an acceptable level. The system can be classified as an indirect dynamic thermal line rating system, where the rating is derived from weather forecasts. The following section briefly describes the individual processes in the system as presented in Figure 6.2. Detailed descriptions of certain sections follow after the overview.

Numerical Weather Prediction The rating system exploits the dependency between ambient variables and the transmission line rating, as described by the model in Equation 2.2. This allows for capacity estimation by processing a weather forecast. Weather Research and Forecasting (WRF) model [72] is an NWP system that is used to generate high resolution weather forecasts on a 1 km spatial grid with 5 minute temporal resolution. For the experimental evaluation, a new 48 hour forecast is generated daily to calculate the rating for the following day.

Machine Learning Model The weather forecast is processed into a probabilistic rating prediction by a supervised machine learning model that has been trained on historical forecasts and measurements. This is similar to the method used by Dupin [3], where a variety of machine learning models are trained to convert a deterministic weather forecast into a probabilistic rating prediction.

In this work, the rating is discretized before the training of a random forest model, essentially turning the regression problem into a classification problem. This method is called *regression via classification* [73]. The classes are uniformly distributed over the range of the rating and the output is constrained to sum to 1, practically creating a probability mass function over the classes.

Time Series Sampling Due to the problem described in the introductory section, the probabilistic rating is not directly usable to estimate summaries of extremes. If the temporal correlations are not taken into account, the extreme values will be overestimated. To overcome this problem, a Monte Carlo sampling method is used. Samples of time series are drawn from the probabilistic forecast and then used together with a cost function to find a rating that incurs a cost acceptable by the line operator.

The samples are drawn in such a way that the temporal dependencies between neighbouring samples are similar to samples from real measurements. The sampling process uses temporal characteristics that are learned from historical measurements

and thus, produces realistic looking samples of time series.

Line Temperature Estimation To allow for the cost calculation, the temperature of the transmission line has to be estimated. The temperature affects both the sag and conductor ageing calculations. Estimating line temperature from the rating is not straightforward due to its dependence on multiple ambient variables that are unknown at this stage of the process. This paper presents a novel method that uses only the ambient temperature forecast to arrive at a conservative estimate for line temperature. This innovative method is described in detail in a later section.

Over-Temperature Estimation The temperature of the line is calculated for a range of ratings. At each rating, the estimated over-temperature is estimated which creates a monotonically increasing function from rating to the over-temperature.

Rating Selection At the end of the process, the rating is selected from the inverse of the function obtained in the previous step at the maximal over-temperature value that is acceptable to the user.

6.1.1 Machine Learning Model

Machine learning is used to convert the deterministic weather forecast directly into a probabilistic rating prediction. This is in contrast to the authors' previous work [74], where a statistical model was applied to the NWP to get a probabilistic weather forecast. The this alternative approach was selected to avoid complications due to correlations between the weather variables that complicated the Monte Carlo calculation of ampacity. Direct estimation of the rating from NWP is in line with other publications [3, 38].

Unlike other research, the presented approach uses a classification algorithm to convert the NWP to line rating. This technique is termed Regression by Classification [73]. It was chosen because probabilistic prediction is required and there are

many readily available classification algorithms that provide the probabilities for the output classes. Furthermore, the proposed time series sampling method operates in discrete space. In Regression by Classification the continuous target variable is first discretized, a classification algorithm is then trained and applied to the data, and finally the results are converted back to the original continuous space.

The transmission line rating has been discretized into 32 bins of equal width, where the limits have been found on the training dataset. Values outside the limit during evaluation were clipped into the lowest or the largest bin. Other discretization strategies have been tested (equally populated bins and k-means clustering), however, bins with uniform width performed best.

The algorithm used to perform the classification is Random Forest [75]. This algorithm was chosen due to good performance of tree-based models in other publications [3, 38], options for probabilistic output, and due to the availability of the easy to use and mature implementation of this algorithm in scikit-learn library¹.

6.1.2 Conductor Temperature Estimation

In order to evaluate the effects of operating the line at a certain current, the temperature of the conductor has to be estimated. If all ambient parameters are known, this can be easily achieved through the application of the IEEE738 thermal model. However, in the system presented in this paper, the ambient variables are *lost* because only the rating is sampled.

Dupin [3] tackles this problem by keeping a database of historical ambient variables and their associated ratings. When the line temperature is required, a row with the closest rating is pulled from the database and these ambient parameters are then used in the thermal model.

A different method is used in this paper. The goal is to calculate the maximal potential line temperature for a rating value given that the actual rating is lower.

¹<https://scikit-learn.org/>

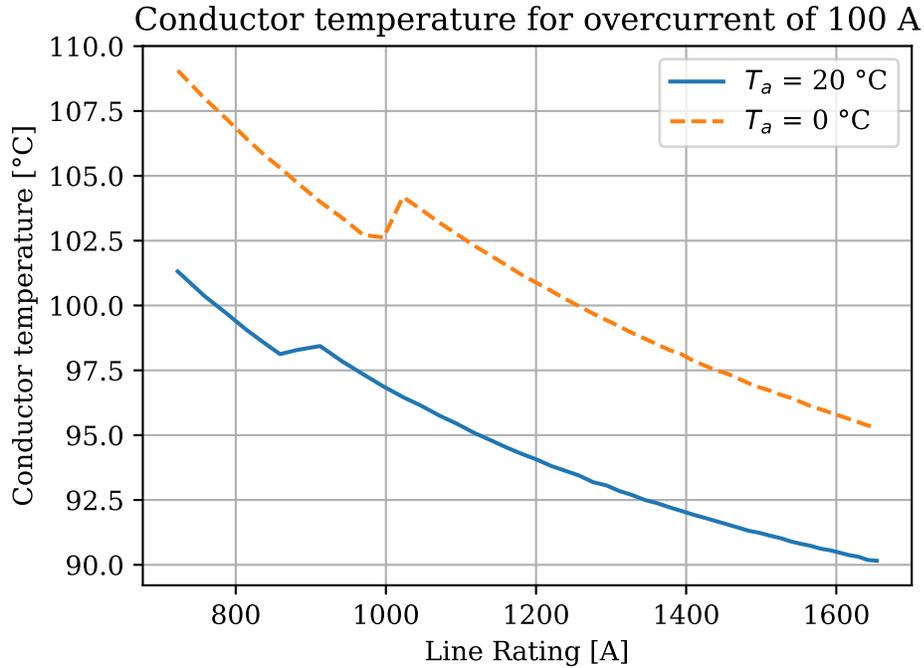


Figure 6.1: This figure shows how the temperature of an overloaded conductor changes not only based on the rating, but also on the ambient air temperature, even though the rating remains the same. The design temperature of the conductor is 80 °C and it has been overloaded by 100 A. The discontinuity in the curves is due to the thermal model switching cooling regimes.

The idea behind this method is that the line temperature at the actual rating is known – it is the design temperature of the conductor. Therefore the wind speed that results in this temperature can be calculated from the thermal model and an ambient temperature forecast. The calculated wind speed can then be used to estimate the temperature at the selected rating.

It is important to evaluate the thermal model at the correct wind speed range because of nonlinearity. The increase in temperature from over-current of 100 A at a true rating of 1200 A is much lower than the temperature increase of the same over-current but with a rating of 600 A. Similarly, given the same rating, the temperature increase varies at different ambient temperatures. Figure 6.1 illustrates this issue.

Define M^I, M^T, M^{ws} to be the IEEE-738 thermal model to calculate the current, conductor temperature and wind speed respectively, from the other variables exclud-

ing the target. Then the cost function f that assigns cost to the conductor temperature can be evaluated for a series of rating samples I_1, \dots, I_n and selected rating I as

$$F([I_1, \dots, I_n], I) = \frac{1}{n} \sum_{i=1}^n f(M^T(T_i, M^{ws}(T_i, I_i)), I). \quad (6.1)$$

This function is evaluated multiple times at gradually increasing values of I until a rating that provides the target cost is found. It is possible to optimize the calculation for cost functions that evaluate to 0 for temperatures lower than the design temperature by skipping samples where $I < I_i$. The same principle could also be employed when evaluating the expected value of the cost function over a part of the probability function of rating $p(I)$:

$$E[F([I_1, \dots, I_n], I)] = \int_0^I p(I') f(M^T(T, M^{ws}(T, I')), I) dI' \quad (6.2)$$

The relation between the ambient temperature T and the current I' in this equation is lost because the calculation treats these as independent variables. For this reason, the numerical evaluation in this paper is done on samples instead of summaries, as samples can be exactly matched to their ambient temperature forecasts.

Example calculation: The rating I selected based on the forecast is 1000 A, but the actual real-time rating I_r is 900 A. What is the temperature T_r of the line under the 1000 A load, assuming the ambient low forecast is 10°C? The temperature at I_R is known, by definition it is the design temperature T of 85°C. The thermal model M^{ws} is used to estimate that a wind speed of 5 m/s results in a conductor temperature of 85°C under 900 A. Now, the thermal model M^T is used to calculate the conductor temperature under the condition of 5 m/s wind speed, 10°C ambient temperature and 1000 A load. The resulting conductor temperature is 94°C. 9°C over the design temperature.

This process is used to calculate the line temperature for each rating with nonzero

predicted probability and the selected rating. This temperature is then used to estimate the cost of operating the line at the selected rating.

6.1.3 Time Series Sampling

The objective of the proposed sampling method is to generate time series samples that would

- mimic typical temporal correlations between time steps of a real-time rating time-series,
- preserve marginal distributions of each random variable as generated by the probabilistic prediction.

The algorithm is based on belief propagation through a graphical model with a chain structure, where the vertices are the random variables and the edges are the factors representing the transition between two time-steps. The algorithm samples the variables in a random order and sends messages with the likelihood to the neighbours. When the neighbours are sampled, the incoming messages from other nodes are incorporated into their probability mass function. Therefore, the samples have a higher probability of being closer to their neighbours in value. As the messages propagate through the chain, the edge transition function is applied to the messages *diluting* the likelihood. The amount of information contained in the messages decreases as the likelihood flattens, meaning the variables further apart have less effect on each other. The propagation of a message stops when an already sampled variable is reached.

Inputs and Outputs

There are two inputs to the algorithm. The first input is a set of n discrete random variables $\{X_1, \dots, X_n\}$, where each variable corresponds to one time step in the rating

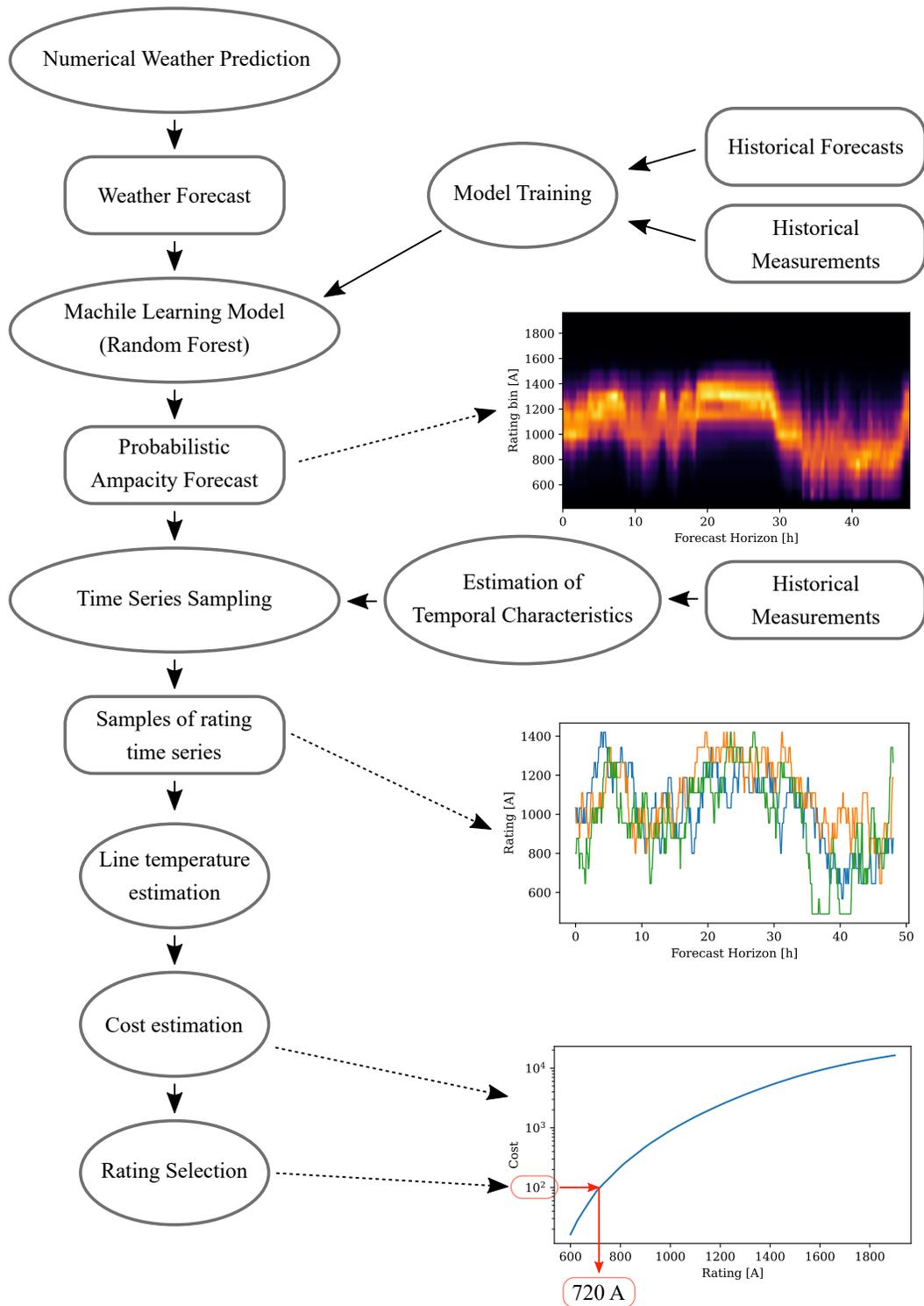


Figure 6.2: Diagram of the rating prediction system

time series. Each random variable is defined by a probability mass function provided by the probabilistic prediction algorithm.

The second input is a conditional probability distribution $E = P(X|Y)$, where X and Y are two neighbouring variables. The distribution describes the probability of X given that the value of Y is known. This distribution is further called the edge function, as it represents the transition probability along one edge. Usually, the probability of similar evaluations of two neighbouring nodes is higher than the probability of more distant ones. The purpose of this input is to provide information about how the value of one variable changes the probability of neighbouring variables. In this algorithm, the probability distribution E is assumed to be stationary with respect to the timesteps. This means all edges have the same transition function regardless of the time-step.

The output of the algorithm is a random sample of a time series $\{x_1, \dots, x_n\}$. The algorithm can be run repeatedly to produce more random samples.

Algorithm Description

In simple terms, the algorithm works by sampling the variables in random order and then propagating the sampled values to its neighbours according to E . The variable X_j is sampled from a PMF defined by the Bayes update rule:

$$P(X_j|X_i; i \neq j) = \frac{\left[\prod_{i \neq j} P(X_i|X_j) \right] P(X_j)}{P(X_i; i \neq j)}, \quad (6.3)$$

where $P(X_j)$ is the prior distribution obtained from the probabilistic prediction, $P(X_i|X_j)$ are the likelihoods of all other variables received through messages, and $P(X_j|X_i; i \neq j)$ is the posterior. Since the structure of the graphical model is a chain, the variable X_i depends only on its neighbours X_{i-1} and X_{i+1} :

$$P(X_j|X_{i-1}, X_{i+1}) = \frac{P(X_i|X_{i-1})P(X_i|X_{i+1})P(X_j)}{P(X_{i-1}, X_{i+1})}. \quad (6.4)$$

Likelihoods are obtained through iterative application of the update rule, assuming the posterior is unknown for variables that have not yet been sampled. The iteration stops at the first observed variable because it stops the propagation of information through the graph:

$$P(X_{n+i}) \leftarrow P(X_n|X_{n+1}) \cdot \dots \cdot P(X_{n+i-1}|X_{n+i})P(X_{n+i}) \quad (6.5)$$

$$= \prod_{j=n}^{i-1} P(X_j|X_{j+1})P(X_{n+i}). \quad (6.6)$$

Recalling that all transition probabilities are identical, the equation can be further simplified to matrix multiplication. Assume that \mathbf{M} is a matrix representing the conditional probability $P(X_j|X_{j+1})$ where PMF are stored in columns. The conditional probability of the chain between X_n and X_m is then \mathbf{M}^d , where $d = |n - m|$ is the number of edges between X_n and X_m :

$$P(X_{n+i}) \leftarrow P(n)\mathbf{M}^iP(X_{n+i}). \quad (6.7)$$

Algorithm 3 shows the implementation of the described procedure. In contrast to Equation 6.6, the the algorithm calculates the likelihood messages when they are transmitted, not when they are received. The transmitted messages are stored in two separate arrays. The messages are propagated through the left and right message arrays, overwriting the previous values until an already sampled location is reached. This ensures that the subsequently sampled location will receive the messages only from the two closest sampled nodes, one message from each side. Besides the message arrays, the algorithm also stores the locations of already sampled variables.

6.1.4 Over-Temperature Cost Function

The average temperature over the design temperature of the conductor has been used to evaluate rating selection methods. Assume a set of n conductor temperature measurements $[T_1, \dots, T_n]$. The average over-temperature, ΔT , is then defined as

Algorithm 3 The sampling algorithm

```
1: input:  $\mathbf{M}^1, X_i, n; \quad i \in [1..n]$ 
2:  $\forall i : \mu_i \leftarrow 1$  ▷ Initialize left messages
3:  $\forall i : \nu_i \leftarrow 1$  ▷ Initialize right messages
4:  $\forall i : \kappa_i \leftarrow \text{false}$  ▷ Initialize sampled flags
5: precalculate  $\mathbf{M}^i : \forall i \in [0..n]$  ▷ Equation 6.7
6:
7: for  $i$  in random permutation of  $1..n$  do
8:    $x_i \leftarrow$  random sample from  $(\mu_i \cdot \nu_i \cdot X_i)$  ▷ Sample from the posterior,
   Equation 6.4
9:   for  $j$  in  $i..n$  do ▷ Propagate messages to the right
10:    if  $\kappa_j == \text{true}$  then ▷ Stop propagation if variable already sampled
11:      Break
12:    end if
13:     $\mu_i \leftarrow \mathbf{M}^{(j-i)} \cdot x_i$  ▷ Update left message array
14:  end for
15:  for  $j$  in  $i..1$  do ▷ Propagate messages to the left
16:    if  $\kappa_j == \text{true}$  then ▷ Stop propagation if variable already sampled
17:      Break
18:    end if
19:     $\nu_i \leftarrow \mathbf{M}^{(i-j)} \cdot x_i$  ▷ Update right message array
20:  end for
21:   $\kappa_i \leftarrow \text{true}$  ▷ Mark sampled
22: end for
    return  $x_i$ 
```

$$\Delta T = \frac{1}{n} \sum_{i=1}^n \max(0, T_i - T_d) \quad (6.8)$$

where T_d is the design temperature of the conductor.

6.1.5 Static Line Rating Overtemperature

The SLR values used in the evaluation are calculated to satisfy these criteria. Moreover, seasonal rating is used to make the comparison more fair. Seasonal rating is a strategy that uses a different SLR for different seasons. Typically, winter months are operated at a higher rating than summer months because the average high temperature over the winter is lower than in the summer. In this study winter SLR is applied between October and March, and summer SLR between April and September.

The criteria for good SLR, according to CIGRE-299 [17], can be used to approximate the expected value of over-temperature for SLR. Good SLR should be lower than the real-time rating with confidence of 99%. Therefore, the temperature of the line is at most 10 °C over the design temperature with probability $p_O = 0.01$. Assuming, that the over-temperature is distributed uniformly between 0 and 10 °C, we can estimate the expected value of the over-temperature, \bar{O}_{extSLR} , as

$$\bar{O}_{SLR} = p_O \cdot E[U(0, 10)] = 0.01 \cdot 5 = 0.05. \quad (6.9)$$

Based on this calculation, the expected over-temperature of SLR, selected according to CIGRE-299, is approximately 0.05 °C.

6.1.6 Rating Estimation Methods

There are three rating estimation methods used in this study. Over-temperature estimation strategy directly searches for a rating that results in the average temperature to equal that of the target over-temperature. This method is implemented by using the over-temperature cost function in Equation 6.8 in the temperature estimation procedure defined by Equation 6.1.

The second method is a simple quantile selection strategy that is commonly used in DTLR studies. The objective of this method is to select a rating that makes the temperature of the conductor less than the design temperature for a pre-selected portion of time. Typically, low quantiles in the range of 1-5% are used with this strategy. In this study, the strategy is applied to the samples obtained through the proposed sampling method. In the diagram, in Figure 6.2, the samples are taken after the step "Time Series Sampling".

The third method is called naive strategy. While somewhat similar to simple quantile selection, the quantile is instead evaluated directly on the probabilistic forecast of rating. This method is included to demonstrate the poor calibration of results that occurs when temporal correlations are ignored throughout the calculation of temporally discretized summaries.

6.2 Results

6.2.1 Dataset

The proposed system is evaluated based on a dataset that consists of 2 years of high resolution measurements from several weather stations located alongside a power transmission line in southern Alberta, Canada. The recorded variables are ambient temperature, wind speed and wind direction, and solar irradiation. The data is recorded every 3 minutes resulting in a dataset of 345,844 measurements.

There are several spans of missing measurements. Gaps smaller than 5 consecutive measurements were filled by linear interpolation, whereas days that contained longer gaps were discarded. There was also several days worth of data with incorrect timestamps that were discarded. In total, there are 543 days of complete data.

The measurement dataset was aligned to the weather prediction dataset so that there were 288 measurements each day with a sampling period of 5 minutes. IEEE738 thermal model was used to calculate the rating from the weather measurements for

”Drake” 26/7 ACSR conductor with a working temperature of 80 °C. The emissivity and solar absorptivity was assumed to be 0.8.

The weather forecast was down-scaled from daily runs of the North American Mesoscale forecast system (NAM) model. This model is released 4 times per day and provides forecasts with an 84-hour horizon with one hour resolution on a 12 km spatial grid. The archived² midnight runs of NAM were used in WRF to down-scale the input model to a 24-hour weather forecast with a temporal resolution of 5 minutes and a grid size of 1 km.

The system was validated on the dataset using 5-fold cross-validation. Days in the dataset were split into 5 groups, where 4 groups were used for training and 1 for testing. The presented results are the means across the testing splits.

6.2.2 Simulation details

The ampacity has been discretized to 32 bins of equal width between 500 and 2000 A. A random forest model was applied to produce probabilistic predictions on these 32 classes. The model used 1000 trees, a maximal depth of 10, and each tree got 5% of training data.

6.2.3 Classification model

Random forest is used to convert the deterministic forecast generated by NWP into a probabilistic forecast. It does this by learning the typical uncertainty in the weather forecast from historical data and then assigning a probability mass function to new forecasts.

There are several aspects of forecast quality. The most important include accuracy, reliability and, sharpness [76]. Accuracy refers to the average difference between individual measurements and forecasts. Reliability is a measure that describes the relationship between observed and predicted values conditioned on the prediction. In

²<https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-mesoscale-forecast-system-nam>

other words, if the forecast predicts probability of an event to be 50%, the event will occur on average in half of the instances. Lastly, sharpness describes the difference between the forecast and climatological value. To elaborate, forecasts too similar to the climatology have no sharpness, whereas sharp forecasts often differ. The goal of a good forecast is to maximize sharpness without sacrificing calibration.

Mean Absolute Error (MAE) will be used to evaluate the accuracy:

$$MAE = \frac{1}{n} \sum |\bar{o}_k - o_k|.$$

The Probability Integral Transform (PIT) histogram assesses the reliability of the forecast. It uses the property of random distributions that states that $u = F_X(x)$ is uniformly distributed if samples x are drawn from a distribution with CDF F_X . Therefore, the predicted CDF will be applied to the the measurements. If the histogram of the result is close to uniform, the prediction is considered reliable.

Sharpness is evaluated using a prediction interval width of 90% [77]. The width will be calculated from the inverse of CDF, as $F^{-1}(0.95) - F^{-1}(0.05)$.

Figure 6.3 shows the distribution of the error over the entire dataset. The distribution of residuals is consistent with a student-t distribution centered around 0. This suggests there is little bias in the model and that it performs well on the data. The MAE of the probabilistic forecasts is 149.6 A and the average width of the 90% prediction interval is 625 A.

The PIT histogram of the model output is in Figure 6.4. The graph shows a 10-bin histogram of PIT together with the optimal shape. It can be seen that the model performance is good and that the true distribution of the data is reflected well by the probabilities assigned by the model.

When using a random forest model it is important to realize that the model cannot be used to extrapolate outside of the space limited by the training data. Random forest directly uses the training dataset to produce results and therefore, the outputs will be clipped to the extent of the values in the training dataset. If using the model

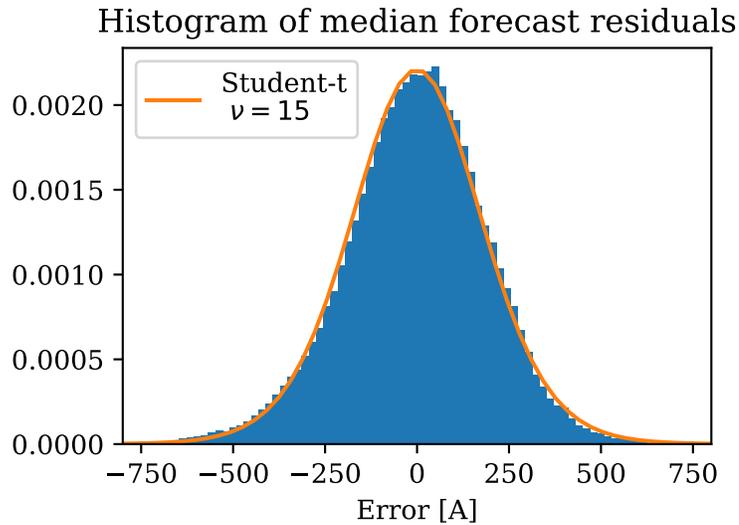


Figure 6.3: A histogram of the error between the median of the probabilistic forecast and the measurements. The error is consistent with the student-t distribution with 15 degrees of freedom.

in production, there has to be a diagnostic procedure that will identify clipping is occurring and either notify the user or replace the output with a value from a different model capable of extrapolation.

6.2.4 Percentile accuracy

Static percentile selection is a simple and commonly used strategy for determining rating. It works by selecting a low percentile, n (typically 1%), and then finding a rating lower than the actual rating $n\%$ of the time.

To validate that the generated samples can be used to reliably determine the rating percentiles, the sampling model will be used to predict percentiles of the minimal rating. Then the predicted value will be compared with the data and the actual portion of minimal rating that exceeds the rating will be determined. Ideally, for a reliable model these proportions should be equal. This means when the model is asked for the 1st percentile prediction of minima, the actual minima should be lower than the predicted in 1% of cases. This simulation was done for a number of prediction window lengths ranging from 1 to 24 hours and percentiles between 0.5% to 10%.

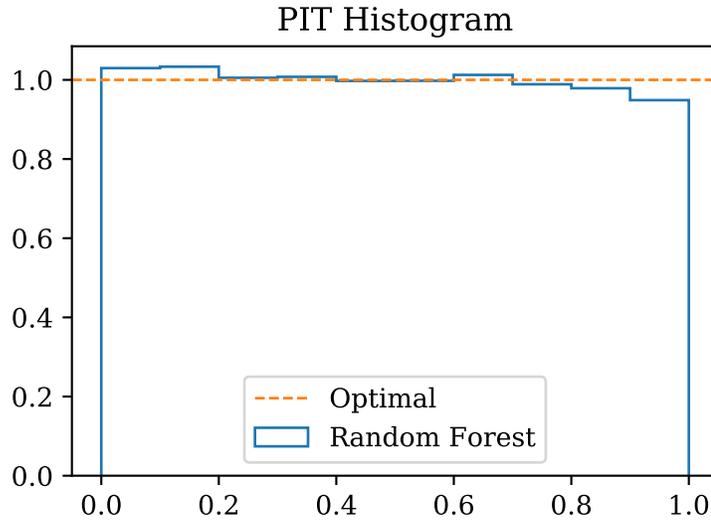


Figure 6.4: PIT Histogram of the probabilistic forecast. Ideally, a reliable forecast should transform into a uniform distribution producing a flat histogram.

Generalized Pareto Distribution (GPD) is used to model the tail of the predicted distribution of minima. The distribution is fitted to the generated samples using Maximum Likelihood (ML) estimation, where the bin samples are treated as interval-censored data. The ML estimation is then performed on $F(x_i + 1) - F(x_i)$ instead of $f(x_i)$ to allow for inter-bin uncertainty.

The model has been used to predict quantiles between 0.01 and 0.1 for discretization periods between 1 and 24 hours over the entire dataset. Figure 6.5 shows the result of this simulation for discretization periods of 2, 6 and 24 hours. The agreement between desired and actual predicted quantiles is generally good, where the true and predicted values are always within 3 standard deviations estimated through bootstrapping (shown for only one curve for clarity).

6.2.5 Distribution of sampled minimum

Figure 6.6 shows the PIT histogram of minima of 6 hour rating windows calculated using Equation 2.5 and estimated from the sampled time series. This figure confirms that disregarding the independence assumption and naive application of the equation

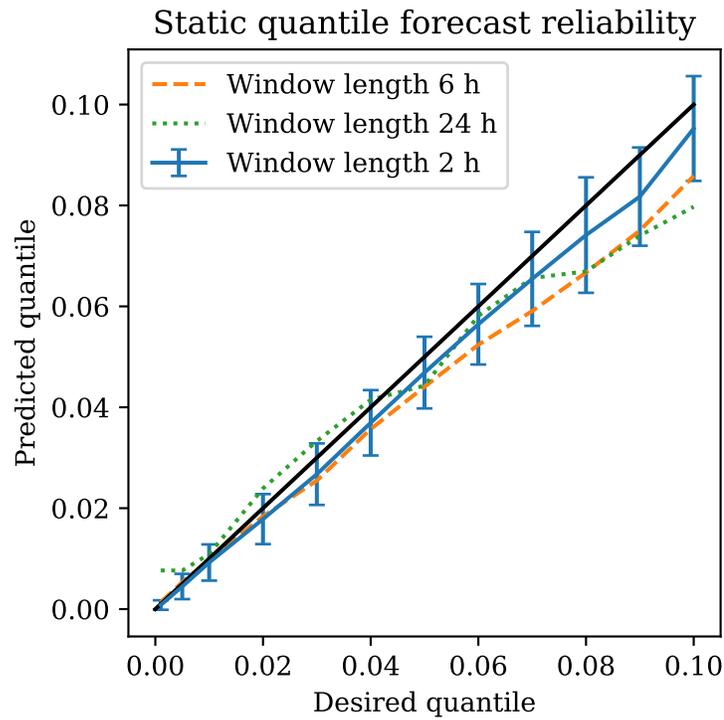


Figure 6.5: The desired quantile against the forecasted quantile. The desired quantile is input into the forecast algorithm, whereas the forecasted quantile is estimated on the data after the fact. Ideally, these values should be equal. The error bars represent 3 standard deviations generated through bootstrapping.

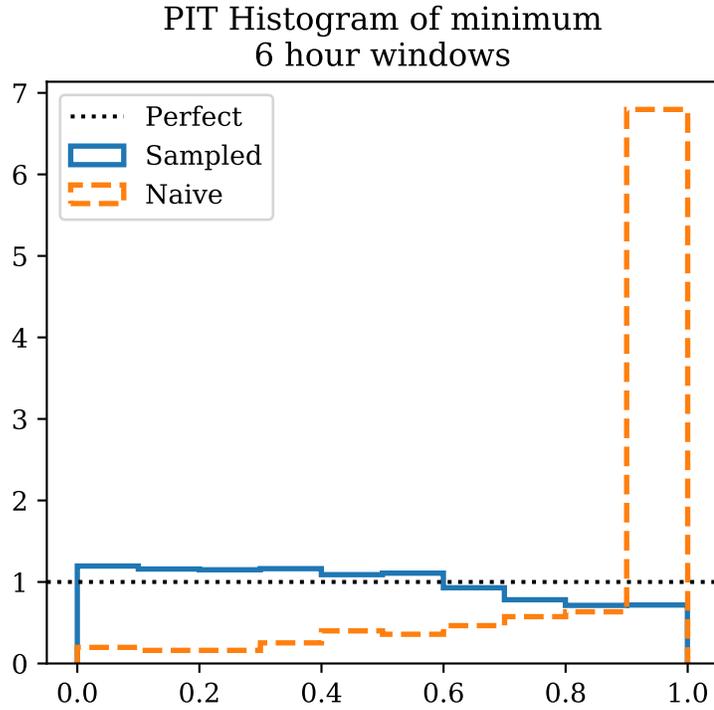


Figure 6.6: PIT of minimal values estimated by applying Equation 2.5 and calculated from sampled data.

results in a probability distribution that is consistently shifted towards lower values. On the other hand, the PIT of the probabilistic prediction of minima produced by the proposed method is sufficiently well calibrated and provides a reliable result with a slight bias toward higher values.

6.2.6 Comparison of rating strategies

In this section, the SLR, naive quantile selection, static quantile selection strategy, and over-temperature estimation strategy will be compared in terms of average rating and over-temperature.

All three algorithms are executed on the output from NWP to generate the rating forecast. The actual measurements are then used to calculate the over-temperature that would occur while operating the line at the forecast rating. Each evaluation is repeated for different lengths of discretization periods. The periods tested in this

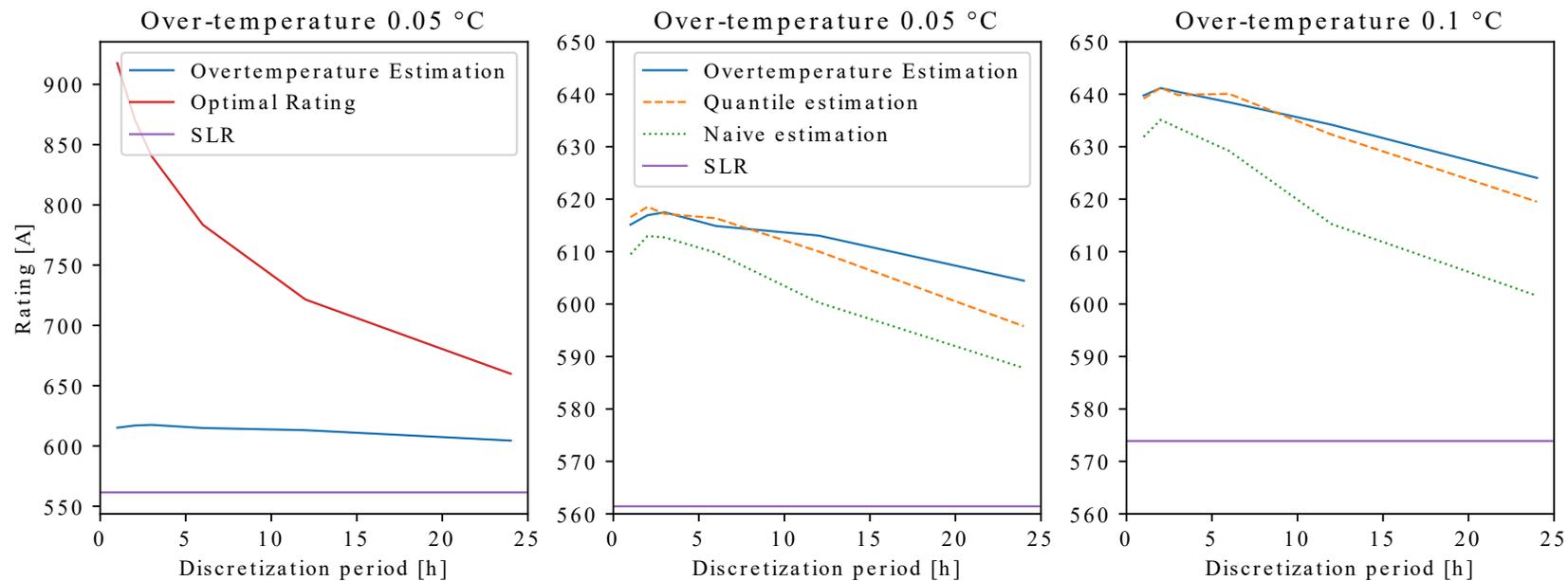


Figure 6.7: This figure shows the dependence of the line rating on the length of the discretization period. The left panel shows the optimal DTLR, the predicted rating by the over-temperature strategy, and the static rating all aligned to an over-temperature of 0.05 °C. The purpose of this figure is to illustrate the large disparity between the optimal possible rating and the predicted rating. Ideally, the predicted rating would be equal to the ideal rating. The middle and right panel compare three different rating strategies under constant over-rating of 0.05 and 0.1 °C.

Average overtemperature for different rating strategies

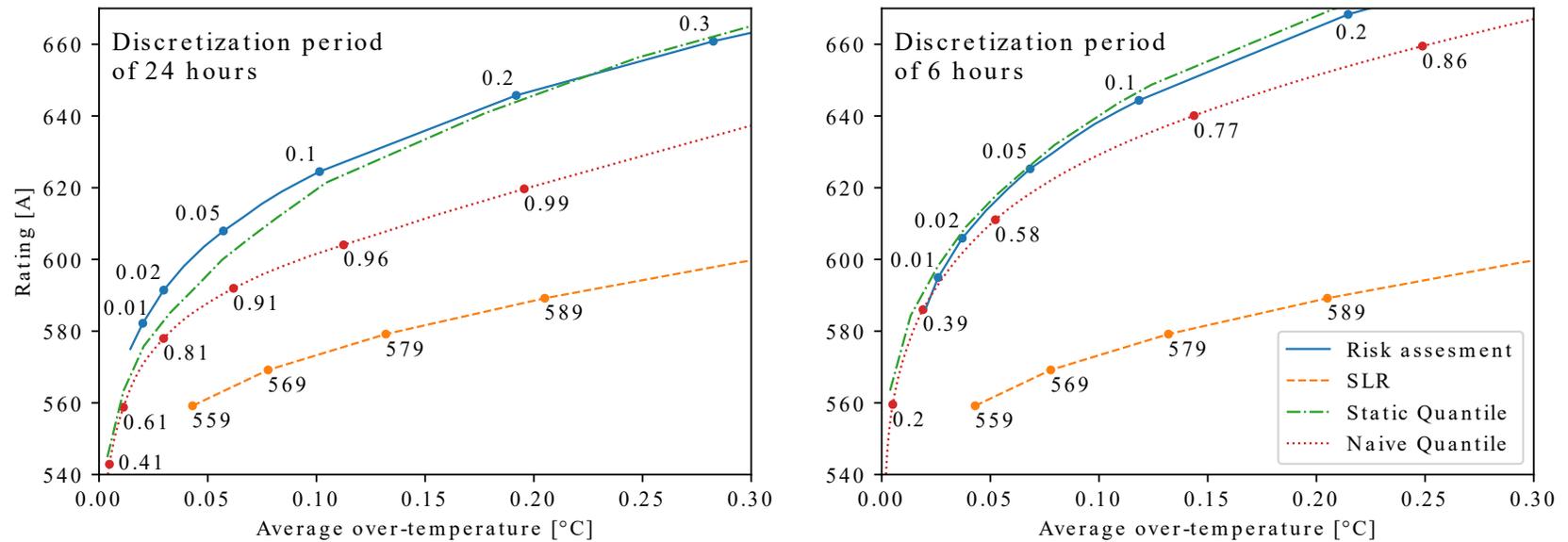


Figure 6.8: The dependency between the average over-temperature and the rating is visualized in this figure for discretization periods of 24 and 6 hours. The numbers on the plot show the input parameter for the rating strategy. This input parameter is the target over-temperature for the over-temperature estimation, the quantile for the naive quantile strategy, and the rating for SLR.

study are 1, 2, 3, 6, 12, and 24 hours. The system uses the high resolution day-ahead NWP to produce a series of discrete ratings covering the next 24 hours. For example, the forecast for the next day discretized to 6 hours would produce a series of 4 ratings, each valid for 6 consecutive hours.

To form a point of reference, the optimal rating is calculated from the measurement data as the rating that maximizes the average rating without ever exceeding the temperature limit. The left panel of Figure 6.7 shows the optimal rating for different discretization periods and also the SLR and rating selected by over-temperature estimation. The optimal rating is 918 and 660 A for 1 and 24 hour discretization periods respectively. As anticipated, the rating fall-off follows the power law, where the rate of the decline decreases as the line approaches the SLR. The optimal rating provides a major increase in capacity compared to SLR, where a 63.4% and 17.5% increase over SLR can be achieved at 1 and 24 hour discretization periods, respectively.

The middle and right panels of Figure 6.7 visualize the effects of temporal discretization on the predicted rating. As expected, the benefit of DTLR is greater at shorter discretizations, but the difference is not large. Especially in comparison with the optimal rating shown in the left panel, where the forecast rating looks almost flat. This is a very important observation, as it suggests that there is only a small benefit in increasing the temporal resolution of rating prediction for these forecasting strategies. The difference in rating between 24 hour (604 A) and 2 hour (617 A) discretization periods is 12.4 A and between 6 (615 A) and 2 hours is only 2 A for an average 0.05 °C over-temperature.

The cause of this effect is currently unknown and should be the focus of further research. Improving the scaling of DTLR benefit with the length of discretization would unlock shorter discretization periods for practical use. The authors believe the limiting factor is currently the low precision of the probabilistic rating forecast. The width of the 90% prediction interval has been established at the beginning of this section as 625 A. This large range forces the value of the high reliability forecast to

be rather low to ensure safety regardless of the size of discretization. Improving the forecast precision would permit the forecasting strategy to produce bolder predictions and allow it to follow the shape of the optimal DTLR more closely.

This data also suggests that discretization is affecting the naive approach more than the other two strategies that use samples that include temporal dependencies. The authors hypothesize this is caused by the collapse of CDF of the estimated minima, where the majority of probability mass is concentrated only on a small range of values. The discrete CDF is incapable of capturing this narrow distribution and thus the estimated rating is a result of simple interpolation between two bin edges for most forecasts.

Figure 6.8 provides a different view of the results. The figure shows the relationship between the average over-temperature and the rating for the three different rating strategies. The left and right panel show the result of the rating strategies applied to the same data, but under different discretization lengths (24 hours on the left panel and 6 hours on the right panel). This data has been obtained by varying the input parameter of the rating algorithm and recording the resulting over-temperature and rating. It also shows the over-temperature of SLR when the static rating is artificially increased from the base rating of 559 A in steps of 10A.

The figure confirms that all three strategies are more efficient than SLR. It also shows that the marginal benefit of DTLR diminishes with respect to increasing the over-temperature. The left panel shows the relationship when discretization of 24 hours is applied. In this time-frame, both over-temperature and quantile estimation clearly perform better than naive strategy. This behaviour is expected because the longer the discretization period, the more error that is introduced by omission of temporal correlations into the calculation of the extreme value. This effect is substantially lower on the data with 6 hour discretization, as can be seen on the right panel of Figure 6.8.

The difference between the over-temperature and quantile selection strategy is not

large. The over-temperature selection performs better on low over-temperature targets under longer discretization periods, but in other cases the results are similar. The over-temperature strategy directly optimizes the temperature, which might contribute to its performance under this metric. It also takes into account the actual value of the forecast over-temperature, whereas the quantile strategy uses only the information that rating is higher or lower than the actual measurement, regardless of the amount over the limit.

The parameters of over-temperature estimation and naive quantile strategy are annotated on the figure next to the appropriate data point. The parameter for over-temperature estimation is the target over-temperature, and for naive strategy it is the quantile. By projecting the data-points on the X axis and comparing the parameters it can be seen that there is good correspondence between the target and the actual value, as calculated on the test data. This relationship also holds at different discretization period lengths, which can be seen by comparing the left and right panels representing 24 and 6 hour periods, respectively. On the other hand, naive strategy requires the selection of an arbitrary quantile through a calibration process to achieve the results at a selected over-temperature level. Furthermore, a different quantile is required to produce the same over-temperature for different discretization periods, as can be seen by comparing the two panels.

6.3 Conclusion

Temporal discretization of transmission line rating is a process of establishing a single constant rating that is safe to use throughout the entire time period. In other words, this rating has to be lower than the actual rating with high probability to ensure the safe operation of the transmission line. The concept of temporal discretization of correlated data and its effect on DTLR has been discussed in this paper and a computational method to overcome this problem has been proposed.

At the beginning of the paper, the effects that temporal correlations have on the

calculation of extreme values were introduced. It was explained that temporal correlations make the calculated extreme values *more extreme* and result in poorly calibrated estimates. This effect was illustrated by an experiment on the measurement data, where a minimum of 6-hour windows was estimated from the probabilistic forecast using a naive method that assumed samples are independent. It is clear that the naive calculation underestimates the minimum and that the majority of samples fall into the 90th percentile or higher.

The system uses an indirect method to estimate the rating through a NWP model that predicts the weather conditions. A machine learning model is then used to convert the deterministic NWP to a probabilistic rating prediction by learning the typical uncertainty of the model on a training dataset. To mitigate the problem of correlated data, a sampling procedure was developed. The proposed method generates samples of an entire time series that has similar temporal correlations as the input data. These samples can now be used to estimate the rating by sorting the samples and selecting the appropriate percentile according to the chosen reliability requirement, or proceed by optimizing cost function over the samples to select the rating based on more sophisticated criteria. ¹¹ The entire system is based on an approach known as *regression-via-classification*. In this framework, continuous inputs are converted to discrete variables. The computation then proceeds in the discrete domain using discrete algorithms such as classification. At the end, the results are converted back into the continuous domain. The advantage of this approach is that it opens the possibility to use powerful machine learning methods like random forest with probabilistic output, graphical models, classification deep neural networks, and others, in regression problems. The positive results in the evaluation demonstrate that regression-via-classification is a viable option for DTLR forecasting and should be considered in further research as an alternative to more common regression techniques.

There were certain challenges with converting discrete outputs back into continu-

ous space. Two techniques that assisted in the process were dithering and fitting a Generalized Pareto distribution (GP) to the discrete data treated as interval-censored data. Dithering was used to smooth the output from discrete Monte Carlo simulation, where a small amount of noise was added to the sampled data before optimizing the cost function. This removed the visible clusters of data around the edges of the bins and greatly improved the output statistics. GP distribution was used when a low quantile was required to estimate from a PMF. GP is often used to model tails of other distributions and is therefore, a good fit for this purpose. The problem of fitting a GP distribution, which is continuous, to a discrete PMF was solved by treating the PMF as an interval-censored dataset. In other words, the actual values are unknown, but the number of values between two edges of a bin can be used in a maximum likelihood estimation to fit the distribution. Again, this technique provides smooth output without any sign of the bin edges in the data.

The purpose of the sampling method is to generate random time series that are compatible with both the probabilistic prediction from NWP and the temporal statistical properties of the real measurements. The proposed method works by sampling each random variable in random order and modifying the PMF of neighbouring variables so that the samples of the neighbours are closer together according to the correlations seen in measurement data. It was demonstrated that this method is capable of producing samples that can be used to estimate the PMF of extreme values. The PIT histogram of the minima prediction over 6 hour blocks of real time rating is nearly flat, indicating that the predicted probability is reliable. This is in sharp contrast to the PIT histogram of the raw probabilistic prediction. Moreover, the calibration works well when the samples are used in the over-temperature estimation strategy. This is supported by the results on the test set being very close to the targets of the algorithm across a wide range of targets and different discretization periods.

Over-temperature estimation uses a novel method to approximate the conductor temperature from the predicted rating. This method is particularly useful in cost

optimization, where the expected conductor temperature for a certain rating has to be known in order to calculate the expected cost of operating the line. This problem is not straightforward, as the thermal model is highly nonlinear and different combinations of ambient temperature and wind speed produce different conductor temperatures for the same current over the limit. The proposed method is able approximate temperature from the predicted rating, the evaluated rating, and the forecast temperature. The nature of this method guarantees that the approximation error will be always conservative, which means the approximated temperature will be higher than the actual temperature given a conservative forecast of ambient temperature.

The proposed method together with two other simpler methods has been tested on two years of high resolution weather data. All training for the involved machine learning methods was done using 5-fold cross validation over daily batches of measurement data and NWP forecasts. There was no overlap between the individual folds. The MAE of the probabilistic forecast of rating obtained from the random forest classification model was 149.6 A and the average width of the 90% prediction interval was 625 A. The errors are equally distributed around zero, which suggests little bias in the model. PIT histogram shows the reliability of the probabilistic forecast is good.

Although this research is focused on temporal discretization, it is important to note the similarity between temporal and spatial discretization of transmission line rating. In the spatial domain, the final rating is dictated by the smallest rating of all segments. Typically, the critical span of the line is identified and then the rating for the entire line is derived from this segment. The technique proposed in this paper could also be applied to spatial correlations, where the correlated rating alongside the length of the line is generated and then directly used to estimate the rating applicable to the entire transmission line.

The relationship between the length of the discretization period and the benefit of the DTLR forecast was also examined. It was expected that the shorter the discretization period, the more beneficial DTLR would be. The relationship generally

holds, however, the effect was much lower than anticipated. The difference in rating between 24-hour (604 A) and 2-hour (617 A) discretization was 12.4 A and between 6-hour (615 A) and 2-hour only 2A, given a 0.05 °C over-temperature. This was a very small difference, especially between 6 and 2 hours. The authors believe this may be caused by the low precision of the probabilistic forecast, where the rating estimation procedure must compensate for the uncertainty by increasing the safety margin thereby nullifying the benefit of DTLR. In the conducted experiment, there was only minimal improvement when the discretization window was decreased to less than 6 hours. Interestingly, the results showed there was actually a small decrease in benefit when using the shortest window of 1 hour, compared to 2 hours.

An area for further research on this topic would be to thoroughly examine the underlying reasons as to why scaling the benefit with length of the discretization period produced a DTLR benefit that was less than expected. The sources of error should be isolated by generating synthetic forecasts where the amount of precision and accuracy is controlled. The effects of various types of uncertainty could then be evaluated. Another possible direction for this research is to explore alternative methods of generating random samples of time series. Most notably, Generative Adversarial Networks (GAN) seems promising given the latest advancements in generating poetry, music, or images. GANs could be tested with DTLR as there have already been applications noted in the literature where they were used to generate time series data [78, 79].

Chapter 7

Conclusion and Final Remarks

DTLR is a promising technology that works to optimize the capacity of a transmission system. To allow for wider practical use, reliable systems capable of forecasting DTLR up to several days into the future are necessary.

Two distinct DTLR forecasting systems were developed throughout this thesis. Both systems utilize NWP, although differ in that one system uses a statistical model to quantify forecast uncertainty, while the other employs a machine learning method.

In this thesis, temporal discretization was defined and its effects on DTLR deduced. The effects were then validated on DTLR measurements and forecasts.

The importance of temporal correlation of time series in conjunction with temporal discretization was elucidated and its effect on DTLR analyzed. A method to incorporate temporal correlations into a DTLR forecasting system was proposed to improve result calibration.

The contributions of the work presented in this theses can be summarized in the terms of the research objectives:

DTLR forecasting system:

- The available literature was analyzed and it was concluded that an indirect NWP driven forecasting system with a probabilistic output is currently the best choice for the practical implementation of a DTLR prediction system.

- A time-series compression method termed DBP with Look Ahead has been proposed for use in data collection by a Wireless Sensor Network. Evaluation of this method on typical sensor data revealed that up to 98% of data transmissions can be saved. This method greatly reduces communication costs of widespread DTLR sensors deployment.
- Two different DTLR systems were developed and evaluated on experimental measurement data. Both systems utilize a NWP weather forecast system to derive the prediction, however, the systems differ in rating estimation technique. The first system uses a custom statistical model, that is simple to implement, but provides an output that requires additional calibration. The second system uses machine learning and a more involved probabilistic sampling method to produce a realistic time series output that can be used with a cost function to generate calibrated predictions.
- Probabilistic predictions are provided by both systems. The ratings estimated from such values result in increased capacity when compared with SLR and have a higher degree of reliability.

Effect of temporal discretization on DTLR:

- A dataset containing 16 years of measurements from 492 weather stations across Canada was used to estimate the effects of temporal discretization.
- It was concluded that the benefit of DTLR decreases with increasing length of temporal discretization. In the test scenario, the majority of locations exhibited an hourly DTLR increase in capacity when compared to SLR by between 103 and 47%. For daily DTLR, the increase was only by 43 and 11%.
- It was derived that the average wind speed decreases following the power law. It can therefore, be argued that the line rating approximately follows the same curve. This was experimentally validated on the measurement data.

Forecasting temporally discretized DTLR with temporal correlations:

- It was concluded, that temporal correlations have detrimental effects on DTLR if temporally discretized ratings are predicted.
- The issue manifests in poorly calibrated predictions where the probability functions of the results are skewed towards the extreme value. This effect increases with longer temporal discretization lengths.
- A sampling method has been proposed that can be used in a Monte Carlo simulation to mitigate the problem. The sampling method generates samples of time series that emulate the temporal correlations in historical data. These samples can be used to optimize a cost function to estimate an objective rating.
- The methodology was tested on measurement data and the results revealed that this method provides results that are well calibrated.

7.1 Future Research Direction

Several areas that warrant future ongoing investigation and research are discussed below.

An algorithm to generate random samples for a Monte Carlo evaluation was proposed in Chapter 6. The purpose of this algorithm is to generate random samples that imitate the temporal correlations in measurement data. As this was the first attempt made to overcome this issue, the author believes alternative methods should be explored. For instance, Generative Adversarial Networks (GANs) have been recently used to generate time series [78, 79]. Given the success of GANs in generating dependant data in other areas, the author suspects GANs are a promising candidate for future testing in DTLR application.

In Chapter 6, the benefit of DTLR prediction was unexpectedly low for lower discretization period lengths. The author expected the benefit to grow with decreasing

discretization period length, however, the result showed that the benefit was relatively insensitive to the variable. This finding was intriguing and a detailed study regarding the source of this behaviour should be conducted.

Related to the previous paragraph, the results of Chapter 6 suggest a discretization of 1-hour performs worse than a 2-hour one. This counter-intuitive result should be explored further to determine if this finding was merely an artifact of the used evaluation and forecasting method or if a longer discretization could have been truly more beneficial.

There is a remarkable similarity between the effects of temporal discretization discussed in this thesis and the problem of critical span selection. The purpose of both concepts is to select the smallest rating over a series of data points. In temporal discretization these points are in time and in critical span selection they are in space. It might be beneficial to merge both of these techniques and unify the rating selection into a single framework. Theoretically, it should be possible to use a two dimensional method to perform the discretization in both time and space to predict a safe rating throughout the time period and over the entire length of the line.

Bibliography

- [1] Canadian Electricity Association, *History of electricity*, <https://electricity.ca/learn/history-of-electricity/>, [Online; accessed 2-November-2020], 2020.
- [2] The World Bank, *Sustainable energy for all dataset*, <https://datacatalog.worldbank.org/dataset/sustainable-energy-all>, [Online; accessed 5-November-2020], 2018.
- [3] R. Dupin, G. Kariniotakis, and A. Michiorri, “Overhead lines dynamic line rating based on probabilistic day-ahead forecasting and risk assessment,” *International Journal of Electrical Power & Energy Systems*, vol. 110, pp. 565–578, 2019, ISSN: 0142-0615. DOI: <https://doi.org/10.1016/j.ijepes.2019.03.043>.
- [4] A. von Meier, “Integration of renewable generation in california: Coordination challenges in time and space,” in *11th International Conference on Electrical Power Quality and Utilisation*, 2011, pp. 1–6. DOI: 10.1109/EPQU.2011.6128888.
- [5] A. Meier, *Electric power systems : a conceptual introduction*. Hoboken, N.J: IEEE Press Wiley-Interscience, 2006, ISBN: 978-0-471-17859-0.
- [6] P. S. R. Murty, *Electrical power systems*. Cambridge, MA: Butterworth-Heinemann, an imprint of Elsevier, 2017, ISBN: 9780081012451.
- [7] S. Stoft, *Power system economics : designing markets for electricity*. Piscataway, NJ New York: IEEE Press Wiley-Interscience, 2002, ISBN: 9780471150404.
- [8] X. Zhang, *Flexible AC transmission systems : modelling and control*. Berlin New York: Springer, 2012, ISBN: 978-3-642-28241-6.
- [9] P. Musilek, J. Heckenbergerova, and M. M. I. Bhuiyan, “Spatial analysis of thermal aging of overhead transmission conductors,” *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1196–1204, 2012. DOI: 10.1109/TPWRD.2012.2199523.
- [10] C. R. Bayliss, *Transmission and distribution electrical engineering*. Oxford: Newnes, 2011, ISBN: 9780080969138.
- [11] I. E. Agency, *World Energy Outlook 2020*. 2020, p. 464. DOI: <https://doi.org/https://doi.org/10.1787/557a761b-en>. [Online]. Available: <https://www.oecd-ilibrary.org/content/publication/557a761b-en>.

- [12] R. Martins, H. Hesse, J. Jungbauer, T. Vorbuchner, and P. Musilek, “Optimal component sizing for peak shaving in battery energy storage system for industrial applications,” *Energies*, vol. 11, no. 8, 2018, cited By 20. DOI: 10.3390/en11082048.
- [13] P. Menck, J. Heitzig, J. Kurths, and H. Schellnhuber, “How dead ends undermine power grid stability,” *Nature Communications*, vol. 5, 2014, cited By 210. DOI: 10.1038/ncomms4969.
- [14] A. Khodaei and M. Shahidehpour, “Microgrid-based co-optimization of generation and transmission planning in power systems,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1582–1590, 2013, cited By 167. DOI: 10.1109/TPWRS.2012.2224676.
- [15] C. Bayliss and B. Hardy, “Chapter 26 - high voltage direct current transmission,” in *Transmission and Distribution Electrical Engineering (Fourth Edition)*, C. Bayliss and B. Hardy, Eds., Fourth Edition, Oxford: Newnes, 2012, pp. 1027–1057, ISBN: 978-0-08-096912-1.
- [16] “Ieee standard for calculating the current-temperature relationship of bare overhead conductors,” *IEEE Std 738-2012 (Revision of IEEE Std 738-2006 - Incorporates IEEE Std 738-2012 Cor 1-2013)*, pp. 1–72, Dec. 2013. DOI: 10.1109/IEEESTD.2013.6692858.
- [17] “Guide for the selection of weather parameters for bare overhead conductor ratings,” International Council on Large Electric Systems, CIGRE, Standard, 2006.
- [18] B. S. Howington and G. J. Ramon, “Dynamic thermal line rating summary and status of the state-of-the-art technology,” *IEEE Transactions on Power Delivery*, vol. 2, no. 3, pp. 851–858, 1987. DOI: 10.1109/TPWRD.1987.4308190.
- [19] M. Dabbaghjamanesh, A. Kavousi-Fard, and S. Mehraeen, “Effective scheduling of reconfigurable microgrids with dynamic thermal line rating,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1552–1564, 2019. DOI: 10.1109/TIE.2018.2827978.
- [20] M. Schneider, A. Hoffrichter, and R. Puffer, “Theoretical potential of dynamic line ratings for congestion management in large-scale power systems,” in *2019 IEEE Milan PowerTech*, 2019, pp. 1–6. DOI: 10.1109/PTC.2019.8810835.
- [21] A. K. Kazerooni, J. Mutale, M. Perry, S. Venkatesan, and D. Morrice, “Dynamic thermal rating application to facilitate wind energy integration,” in *2011 IEEE Trondheim PowerTech*, Jun. 2011, pp. 1–7.
- [22] A. Michiorri, R. Currie, P. Taylor, F. Watson, and D. Macleman, “Dynamic line ratings deployment on the orkney smart grid,” 2011.
- [23] T. Ringelband, M. Lange, M. Dietrich, and H. Haubrich, “Potential of improved wind integration by dynamic thermal rating of overhead lines,” in *2009 IEEE Bucharest PowerTech*, 2009, pp. 1–5. DOI: 10.1109/PTC.2009.5281897.

- [24] N. Viafora, S. Delikaraoglou, P. Pinson, and J. Holbøll, “Chance-constrained optimal power flow with non-parametric probability distributions of dynamic line ratings,” *International Journal of Electrical Power & Energy Systems*, vol. 114, p. 105 389, 2020, ISSN: 0142-0615. DOI: <https://doi.org/10.1016/j.ijepes.2019.105389>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0142061519309317>.
- [25] K. Bubenchikov, A. Gonzalez-Castellanos, and D. Pozo, “Benefits of dynamic line rating for the russian power corridor between the european and siberian zones,” in *2020 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, 2020, pp. 1–6. DOI: 10.1109/REEPE49198.2020.9059177.
- [26] B. P. Bhattarai, J. P. Gentle, T. McJunkin, P. J. Hill, K. S. Myers, A. W. Abboud, R. Renwick, and D. Hengst, “Improvement of transmission line ampacity utilization by weather-based dynamic line rating,” *IEEE Transactions on Power Delivery*, vol. 33, no. 4, pp. 1853–1863, 2018. DOI: 10.1109/TPWRD.2018.2798411.
- [27] F. Teng, R. Dupin, A. Michiorri, G. Kariniotakis, Y. Chen, and G. Strbac, “Understanding the benefits of dynamic line rating under multiple sources of uncertainty,” *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3306–3314, 2018. DOI: 10.1109/TPWRS.2017.2786470.
- [28] C. W. G. B2.42, “Guide for thermal rating calculations of overhead lines,” *Technical Brochure 601*, 2014.
- [29] A. Arroyo, P. Castro, R. Martinez, M. Manana, A. Madrazo, R. Lecuna, and A. Gonzalez, “Comparison between ieee and cigre thermal behaviour standards and measured temperature on a 132-kv overhead power line,” *Energies*, vol. 8, no. 12, pp. 13 660–13 671, 2015, ISSN: 1996-1073. DOI: 10.3390/en81212391.
- [30] E. Diaconescu, “The use of narx neural networks to predict chaotic time series,” *WSEAS Trans. Comp. Res.*, vol. 3, no. 3, pp. 182–191, Mar. 2008, ISSN: 1991-8755.
- [31] R. Shumway, *Time series analysis and its applications with R examples*. New York: Springer, 2011, ISBN: 978-1-4419-7865-3.
- [32] J. Connor *et al.*, “Recurrent neural networks and robust time series prediction,” *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 240–254, Mar. 1994, ISSN: 1045-9227. DOI: 10.1109/72.279188.
- [33] S. Coles, *An introduction to statistical modeling of extreme values*. London New York: Springer, 2001, ISBN: 9781447136750.
- [34] S. Karimi, P. Musilek, and A. M. Knight, “Dynamic thermal rating of transmission lines: A review,” *Renewable and Sustainable Energy Reviews*, vol. 91, no. C, pp. 600–612, 2018. DOI: 10.1016/j.rser.2018.03.10.

- [35] D. A. Douglass, J. Gentle, H. Nguyen, W. Chisholm, C. Xu, T. Goodwin, H. Chen, S. Nuthalapati, N. Hurst, I. Grant, J. A. Jardini, R. Kluge, P. Traynor, and C. Davis, “A review of dynamic thermal line rating methods with forecasting,” *IEEE Transactions on Power Delivery*, vol. 34, no. 6, pp. 2100–2109, Dec. 2019, ISSN: 1937-4208. DOI: 10.1109/TPWRD.2019.2932054.
- [36] I. Albizu, E. Fernandez, A. J. Mazon, K. J. Sagastabeitia, M. T. Bedialauneta, and J. G. Olazarri, “Overhead line rating forecasting for the integration of wind power in electricity markets,” in *2015 International Conference on Clean Electrical Power (ICCEP)*, 2015, pp. 382–388. DOI: 10.1109/ICCEP.2015.7177652.
- [37] S. Uski, “Dynamic line rating forecastability for conservative day-ahead line rating values,” in *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, Nov. 2015, pp. 003 738–003 742. DOI: 10.1109/IECON.2015.7392683.
- [38] J. L. 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, Feb. 2017, ISSN: 0885-8977. DOI: 10.1109/TPWRD.2016.2543818.
- [39] M. Taillardat, O. Mestre, M. Zamo, and P. Naveau, “Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics,” *Monthly Weather Review*, vol. 144, no. 6, pp. 2375–2393, 2016. DOI: 10.1175/MWR-D-15-0260.1.
- [40] T. Ringelband, P. Schäfer, and A. Moser, “Probabilistic ampacity forecasting for overhead lines using weather forecast ensembles,” *Electrical Engineering*, vol. 95, no. 2, pp. 99–107, Jun. 2013, ISSN: 1432-0487. DOI: 10.1007/s00202-012-0244-8.
- [41] R. Dupin, A. Michiorri, and G. Kariniotakis, “Optimal dynamic line rating forecasts selection based on ampacity probabilistic forecasting and network operators risk aversion,” *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2836–2845, 2019. DOI: 10.1109/TPWRS.2018.2889973.
- [42] F. K. Shaikh and S. Zeadally, “Energy harvesting in wireless sensor networks: A comprehensive review,” *Renewable and Sustainable Energy Reviews*, vol. 55, pp. 1041–1054, 2016, ISSN: 1364-0321. DOI: <http://dx.doi.org/10.1016/j.rser.2015.11.010>.
- [43] N. Pantazis *et al.*, “Energy-efficient routing protocols in wireless sensor networks: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 551–591, Feb. 2013, ISSN: 1553-877X. DOI: 10.1109/SURV.2012.062612.00084.
- [44] G. Anastasi *et al.*, “Energy conservation in wireless sensor networks: A survey,” *Ad Hoc Netw.*, vol. 7, no. 3, pp. 537–568, May 2009, ISSN: 1570-8705. DOI: 10.1016/j.adhoc.2008.06.003.

- [45] T. Rault *et al.*, “Energy efficiency in wireless sensor networks: A top-down survey,” *Computer Networks*, vol. 67, pp. 104–122, 2014, ISSN: 1389-1286. DOI: <http://dx.doi.org/10.1016/j.comnet.2014.03.027>.
- [46] U. Raza *et al.*, “What does model-driven data acquisition really achieve in wireless sensor networks?” In *2012 IEEE Int. Conference on Pervasive Computing and Commun. (PerCom)*, Mar. 2012, pp. 85–94. DOI: 10.1109/PerCom.2012.6199853.
- [47] A. Bogliolo *et al.*, “Towards a true energetically sustainable wsn: A case study with prediction-based data collection and a wake-up receiver,” in *2014 9th IEEE Int. Symp. on Ind. Embedded Syst. (SIES)*, Jun. 2014, pp. 21–28. DOI: 10.1109/SIES.2014.6871181.
- [48] T. Potsch *et al.*, “Model-driven data acquisition for temperature sensor readings in wireless sensor networks,” in *2014 IEEE Ninth Int. Conference on Intelligent Sensors, Sensor Networks and Inform. Process. (ISSNIP)*, Apr. 2014, pp. 1–6. DOI: 10.1109/ISSNIP.2014.6827658.
- [49] J. M. P. Menezes Jr. and G. A. Barreto, “Long-term time series prediction with the narx network: An empirical evaluation,” *Neurocomput.*, vol. 71, no. 16-18, pp. 3335–3343, Oct. 2008, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2008.01.030.
- [50] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd. Chichester, England: Wiley Publishing, 2007, ISBN: 0470035617.
- [51] K. Deep *et al.*, “A real coded genetic algorithm for solving integer and mixed integer optimization problems,” *Applied Mathematics and Computation*, vol. 212, no. 2, pp. 505–518, 2009, ISSN: 0096-3003. DOI: <http://dx.doi.org/10.1016/j.amc.2009.02.044>.
- [52] U. Raza *et al.*, “Practical data prediction for real-world wireless sensor networks,” *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 8, pp. 2231–2244, Aug. 2015, ISSN: 1041-4347. DOI: 10.1109/TKDE.2015.2411594.
- [53] A. Bergström, U. Axelsson, and V. Neimane, “Dynamic capacity rating for wind cooled overhead lines,” in *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*, Jun. 2013, pp. 1–4. DOI: 10.1049/cp.2013.0849.
- [54] E. Fernandez, I. Albizu, M. Bedialauneta, A. Mazon, and P. Leite, “Review of dynamic line rating systems for wind power integration,” *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 80–92, 2016, ISSN: 1364-0321. DOI: <https://doi.org/10.1016/j.rser.2015.07.149>.
- [55] J. Heckenbergerova, P. Musilek, and K. Filimonenkov, “Assessment of seasonal static thermal ratings of overhead transmission conductors,” in *2011 IEEE Power and Energy Society General Meeting*, Jun. 2011, pp. 1–8. DOI: 10.1109/PES.2011.6039393.

- [56] R. Dupin and A. Michiorri, “13 - dynamic line rating forecasting,” in *Renewable Energy Forecasting*, ser. Woodhead Publishing Series in Energy, G. Kariniotakis, Ed., Woodhead Publishing, 2017, pp. 325–339, ISBN: 978-0-08-100504-0. DOI: <https://doi.org/10.1016/B978-0-08-100504-0.00013-5>.
- [57] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences, Volume 100 (International Geophysics)*. Academic Press, 2011, ISBN: 0123850223.
- [58] J. Hosek, P. Musilek, E. Lozowski, and P. Pytlak, “Effect of time resolution of meteorological inputs on dynamic thermal rating calculations,” *IET Generation, Transmission Distribution*, vol. 5, no. 9, pp. 941–947, Sep. 2011, ISSN: 1751-8687. DOI: 10.1049/iet-gtd.2011.0039.
- [59] S. Grange, “Technical note: Averaging wind speeds and directions,” Tech. Rep., Jun. 2014. DOI: 10.13140/RG.2.1.3349.2006.
- [60] S. Al-Yahyai, A. Gastli, and Y. Charabi, “Probabilistic wind speed forecast for wind power prediction using pseudo ensemble approach,” in *2012 IEEE International Conference on Power and Energy (PECon)*, Dec. 2012, pp. 127–132.
- [61] C. Lu, H. Yuan, B. E. Schwartz, and S. G. Benjamin, “Short-range numerical weather prediction using time-lagged ensembles,” *Weather and Forecasting*, vol. 22, no. 3, pp. 580–595, 2007. DOI: 10.1175/WAF999.1.
- [62] J. Seguro and T. Lambert, “Modern estimation of the parameters of the weibull wind speed distribution for wind energy analysis,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 85, no. 1, pp. 75–84, 2000, ISSN: 0167-6105. DOI: [https://doi.org/10.1016/S0167-6105\(99\)00122-1](https://doi.org/10.1016/S0167-6105(99)00122-1).
- [63] H. Rinne, *The Weibull Distribution: A Handbook*. Chapman and Hall/CRC, Nov. 2008.
- [64] M. B. Thompson, *A comparison of methods for computing autocorrelation time*, 2010. arXiv: 1011.0175 [stat.CO].
- [65] A. Gelman, *Bayesian Data Analysis: Texts in Statistical Science*. CRC Press, 2013.
- [66] R. Morris, *Updating CWEEDS Weather Files*. Jul. 2016. [Online]. Available: https://climate.weather.gc.ca/prods%5C_serve/engineering%5C_e.html.
- [67] A. Deb, *Powerline Ampacity System: Theory, Modeling and Applications*. CRC Press, 2017, ISBN: 9781482274127. [Online]. Available: <https://books.google.ca/books?id=TbDNBQAAQBAJ>.
- [68] C. Black and W. Chisholm, “Key considerations for the selection of dynamic thermal line rating systems,” *IEEE Transactions on Power Delivery*, vol. 30, no. 5, pp. 2154–2162, 2015, cited By 35. DOI: 10.1109/TPWRD.2014.2376275.
- [69] E. Carlini, C. Pisani, A. Vaccaro, and D. Villacci, “A reliable computing framework for dynamic line rating of overhead lines,” *Electric Power Systems Research*, vol. 132, pp. 1–8, 2016, cited By 14. DOI: 10.1016/j.epsr.2015.11.004.

- [70] N. Woodhouse and P. Musilek, “Computational acquisition of meteorological data for applications in electric power systems,” in *2020 IEEE Electric Power and Energy Conference (EPEC)*, Edmonton, AB, Canada, Nov. 2020, pp. 1–5.
- [71] S. Nadarajah and S. Kotz, “Exact distribution of the max/min of two gaussian random variables,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 2, pp. 210–212, 2008. DOI: 10.1109/TVLSI.2007.912191.
- [72] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, Z. Liu, J. Berner, W. Wang, J. G. Powers, M. G. Duda, D. M. Barker, and X.-Y. Huang, “A description of the advanced research wrf model version 4,” en, Tech. Rep., 2019. DOI: 10.5065/1DFH-6P97. [Online]. Available: <https://opensky.ucar.edu/islandora/object/opensky:2898>.
- [73] L. Torgo and J. Gama, “Regression by classification,” in *Advances in Artificial Intelligence*, D. L. Borges and C. A. A. Kaestner, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 51–60, ISBN: 978-3-540-70742-4.
- [74] T. Barton and P. Musilek, “Day-ahead dynamic thermal line rating using numerical weather prediction,” in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, 2019, pp. 1–7.
- [75] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.
- [76] D. Wilks, *Statistical methods in the atmospheric sciences : an introduction*. Amsterdam: Elsevier, 2019, ISBN: 9780128165270.
- [77] T. Gneiting, F. Balabdaoui, and A. E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 2, pp. 243–268, 2007. eprint: <https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2007.00587.x>. [Online]. Available: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2007.00587.x>.
- [78] C. Esteban, S. L. Hyland, and G. Rätsch, *Real-valued (medical) time series generation with recurrent conditional gans*, 2017. arXiv: 1706.02633 [stat.ML].
- [79] J. Yoon, D. Jarrett, and M. van der Schaar, “Time-series generative adversarial networks,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32, Curran Associates, Inc., 2019, pp. 5508–5518. [Online]. Available: <https://proceedings.neurips.cc/paper/2019/file/c9efe5f26cd17ba6216bbe2a7d26d490-Paper.pdf>.